Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Departamento de Estatística

"Modelos multiníveis para resposta binária: um enfoque na má classificação"

Sandra Rêgo de Jesus

Orientação: Profª Drª Viviana Giampaoli

Co-orientação: Profª Drª Maria Cristina Falcão Raposo

Área de Concentração: Estatística Aplicada

Dissertação apresentada ao Departamento de Estatística da Universidade Federal de Pernambuco para obtenção do grau de Mestre em Estatística

Recife, janeiro de 2005

Universidade Federal de Pernambuco
Mestrado em Estatística

21 de janeiro de 2005
(data)

Nós recomendamos que a dissertação de mestrado de autoria de

Sandra Rêgo de Jesus

intitulada

Modelos multinívois para resposta binária: um enfoque na má classificação

seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

Maria Cristina Falcão Raposo — orientador

Dalton Francisco de Andrade (UFSC)

Cristiano Ferraz

Este documento será anexado à versão final da dissertação.

"Modelos multiníveis para resposta binária: um enfoque na má classificação"

Sandra Rêgo de Jesus

Dissertação apresentada ao Departamento de Estatística da UFPE, para a obtenção do grau de Mestre em Estatística, Área de Concentração: Estatística Aplicada.

Recife/PE
Janeiro/2005

"O importante é estar pronto, a qualquer momento, a sacrificar aquilo que somos em favor do que podemos vir a ser". (Charles Dubois).

Aos meus pais, Antônio e Disdete, e aos meus irmãos, Antônio e Osvaldino.

AGRADECIMENTOS

À Deus, inteligência suprema e causa primária de todas as coisas, luz do meu caminho, força na minha fraqueza, ânimo no meu cansaço, companhia na minha solidão, perseverança nos obstáculos, presença constante no meu coração.

Aos meus pais, que com amor e compreensão sempre me apoiaram e me incentivaram durante todos os acontecimentos de minha vida e com quem aprendi a não desistir diante às dificuldades. À vocês todo o meu amor.

À professora Viviana Giampaoli pela orientação paciente, dedicação, competência, compreensão, carinho, amizade e por ter acreditado em mim.

À professora Maria Cristina Falcão Raposo pela co-orientação competente com que forneceu ótimas sugestões para realização deste trabalho e também pela compreensão, amizade, carinho e preocupação.

Aos meus irmãos, Antônio e Osvaldino, pelo amor, carinho e amizade.

Ao querido e amado, Ariel Victor, por existir e pelos inúmeros momentos de felicidade.

À minha tia Bernadete pelo incentivo, carinho e muito apoio, nos momentos delicados.

Ao amigo Nove, pelos conselhos maravilhosos, carinho, preocupação e, que mesmo de longe contribuiu de forma tão especial para a finalização deste trabalho.

À Matildes pela atenção, carinho e amizade, e acima de tudo por sempre acreditar em meu trabalho e ser uma verdadeira mãe para mim.

À Nicinha e Amélia pelo apoio e carinho com que me receberam em Recife.

A Gilson, pela confiança, companheirismo, irmandade, atenção, presteza, e acima de tudo uma amizade sincera.

À Silvia pelo carinho e pela convivência compartilhada no primeiro ano.

À Patricia Leal pela irmandade, atenção e presteza.

À Tatiene e Moises pela atenção e apoio.

Aos colegas da minha turma de mestrado: Tatiane, uma das pessoas mais prestativas que já conheci, ombro amigo no desespero e sorriso imenso nas alegrias, muito obrigada amiga, aprendi muito com você; Sandra Pinheiro, por sua seriedade, amizade, respeito e compreensão; Júnior, por estar sempre disposto a ajudar e pelos momentos de descontração; Andreia, por ser meiga e atenciosa; Gecinalda, por sua alegria e por sua enorme confiança;

André, por sua tranquilidade; Cherubino (*in memorian*) pelas idéias compartilhadas.

A todos os colegas de mestrado, Renata, Lenaldo, Daniela, Polyane, Luz Milena, Kátia, Themis, Tiago, Luiz Hernando, Francisco, Juan Camilo, Carlos Tomé, Carlos Gadelha, Denis e Artur.

A professora Claudia, por seu enorme carinho e apoio.

A todos os professores do Programa de Mestrado em Estatística da UFPE.

A secretária Valéria, pela competência, disponibilidade e atenção.

A minha amiga e eterna professora Giovana.

A todos que, de forma direta ou indireta, contribuíram para a realização deste trabalho.

Aos participantes da banca examinadora pelas sugestões.

À CAPES pelo suporte financeiro.

Resumo

Modelos multiníveis para resposta binária são uma ferramenta poderosa para análise de dados binários que tem uma estrutura hierárquica ou de agrupamento. Esta estrutura é caracterizada pela presença de observações individuais (pessoas ou objetos em estudo) que são considerados agrupados em um nível mais alto. Este agrupamento pode ser devido a subamostragem em unidades amostrais primárias. Assim, as unidades de um mesmo nível pertencentes a uma unidade de nível mais alto raramente são independentes. Por este motivo os modelos mutiníveis são apropriados para avaliar dados desta natureza. Eles permitem medir como as variáveis de vários níveis afetam a variável resposta, além de quantificar quanto da variabilidade da resposta se deve a cada nível.

Os modelos multiníveis para dados binários assumem que a variável resposta é medida sem erro. O objetivo deste trabalho é apresentar a teoria dos modelos multiníveis de 2 níveis para resposta binária com um enfoque na má classificação bem como os métodos de implementação na análise de grupo-específico, isto é, na análise dos modelos mistos lineares generalizados, quando as probabilidades de má classificação são conhecidas. Estes métodos mostram que ignorar erros na resposta conduz a perda substancial de informações sobre os efeitos da covariáveis. Foram realizados estudos de simulações que confirmam estas conclusões. Uma aplicação na área de saúde é também apresentada. Para as análises foi utilizado o programa computacional `R`.

Palavras-chave: Modelos multiníveis, dados binários, modelos mistos lineares generalizados, resposta binária má classificada.

Abstract

Multilevel models for binary response are a powerful tool for the analyses of binary data with hierarchical or clustering structure. This structure is characterized by the presence of individual observations (people or objects under study), considered as a cluster in a higher level. This clustering may be due to subsampling primary sampling units. Under such a structure, units of the same level, belonging to a higher level unit, are rarely independent. For this reason, multilevel models are appropriate to evaluate data of this nature. They not only allow to measure how covariates affect the outcome at different levels, but also allow to detect how much of the outcome variability is due to each level.

The multilevel models for binary data assume that the response variable are measured without error. The objective of this work is to introduce the theory of two-level multilevel models for binary response focusing on misclassification. Implementation methods for cluster-specific analyses are also introduced, featuring the analyses of generalized linear mixed models when the probabilities of misclassification are known. These methods show that ignoring errors in the response variable can lead to substantial losses of information about covariate effects. Simulation studies illustrate the findings. An application in the health area is also presented. Statistical analyses were conducted using the computational package R.

Key-words: Multilevel models, binary data, generalized linear mixed models, binary misclassified response.

SUMÁRIO

Página

1 Introdução

1.1 Considerações inicias

Os modelos multiníveis para resposta binária são uma ferramenta poderosa para análise de dados que tem uma estrutura hierárquica ou de agrupamento. Estes modelos têm sido utilizados para descrever processos epidemiológicos, demográficos, econômicos e sociais, dentre outros, que estão interessados em explicar e predizer fenômenos que podem ser caracterizados por uma variável resposta binária.

A estrutura hierárquica de agrupamento é definida pela presença de observações individuais (pessoas ou objetos em estudo) que são consideradas agrupadas ou aninhadas em um nível mais baixo, os quais estão agrupados dentro de um nível mais alto e assim sucessivamente. Dados com essa estrutura são encontrados em várias áreas de estudo. Por exemplo, em estudos sociais (de instituições) em que o pesquisador pode estar interessado em investigar como as características do local de trabalho influenciam na produtividade do trabalhador. Os dados amostrais deste estudo são vistos como uma estrutura hierárquica, com as unidades de trabalhadores agrupados dentro das firmas. Dessa forma, as variáveis poderão ser mensuradas em ambos os níveis: trabalhador (nível 1) e firma (nível 2). Outro exemplo são estudos da área de educação; o pesquisador pode estar interessado em saber como a estrutura organizacional de uma escola influencia no desempenho do aluno. Nesta situação, os estudantes compõem as turmas, as turmas compõem as escolas e as escolas compõem os distritos escolares (órgãos administradores). Assim, teremos quatro níveis hierárquicos associando-se os estudantes ao nível 1, as turmas ao nível 2, as escolas ao nível 3 e os distritos escolares ao nível 4.

Quando os dados são estruturados em hierarquias, as unidades de um mesmo nível pertencentes a uma unidade de nível mais alto, raramente são independentes. Isso acontece porque essas unidades compartilham de um mesmo ambiente ou apresentam características semelhantes. Portanto, a suposição de independência é violada e passa a existir correlação entre as observações de uma mesma unidade. Ignorar o risco desta relação, omitindo a importância de efeitos de grupo, pode tornar inválida muitas das técnicas de análise estatística tradicionais utilizadas para estudar a relação de dados.

Este argumento indica que uma análise de modelos de regressão tradicionais, que exige a suposição de independência entre as observações, não seria a mais adequada para essa estrutura de dados, pois implicaria na superestimação dos erros padrão do modelo em estudo.

Para isto, foram desenvolvidos modelos multiníveis, uma extensão direta dos modelos lineares generalizados de McCullagh e Nelder (1989), os quais contemplam uma estrutura hierárquica dos dados, considerando todas as correlações existentes entre as observações, nos diferentes níveis de hierarquia, e possíveis influências aleatórias. Assim, a principal diferença desses modelos reside nas estimativas dos erros padrão, que serão mais precisas, uma vez que a correlação intra-classe é considerada no ajustamento dos modelos.

Nos modelos multiníveis, as unidades de cada nível são vistas como tendo um efeito aleatório, ou seja, são amostras aleatórias de uma população dessas unidades. Estes efeitos aleatórios são responsáveis por coeficientes aleatórios que levam em conta a variabilidade intra unidades através da variabilidade nos interceptos e/ou através da variabilidade nas inclinações. Além disso, os modelos multiníveis com efeitos aleatórios possuem a vantagem de acomodarem dados hierárquicos (Bergamo, 2002).

Segundo Guo e Zhao (2000), as principais vantagens dos modelos multiníveis, em relação aos modelos de regressão tradicionais, são que eles permitem: medir como os indivíduos, as variáveis de vários níveis e as interações entre as variáveis de diferentes níveis afetam a variável resposta; estimar os parâmetros com seus respectivos erros padrão e intervalos de confiança precisos que são resultantes da estrutura de agrupamento; decompor a variância total da variável resposta em partes associadas a cada nível, que é obtida da estimativa da covariância dos efeitos aleatórios de vários níveis, além de fornecerem testes de significância.

1.2 Revisão de literatura

Os modelos lineares multiníveis são também conhecidos como modelos lineares hierárquicos (MLH), modelos com coeficientes aleatórios, ou regressões aparentemente não relacionadas ("*seemingly unrelated regressions*").

Segundo Bryk e Raudenbush (1992), o termo modelo linear hierárquico foi introduzido por Lindley e Smith (1972) e Smith (1973) como parte das suas contribuições em estimações bayesianas de modelos lineares, mas a utilização destes modelos requeria a estimação dos componentes da variância na presença de dados não balanceados e nenhuma aproximação de estimação geral era possível no início dos anos 70, pois não havia nenhum método aproximado para estimar os componentes da variância na presença de dados não balanceados.

Benett (1976) com base num estudo com crianças do ensino fundamental realizado nos anos 70 na Inglaterra supôs que as crianças expostas à uma maneira formal de ensinar a ler, mostravam um aprendizado maior do que as não expostas. A análise foi feita utilizando técnicas de regressão múltipla, considerando apenas as crianças individualmente como sendo as unidades de análise e ignorando ignorando os agrupamentos para professores e classes

escolares. Os resultados obtidos foram estatisticamente significantes. Dempster *et al.* (1977), desenvolveram um procedimento inovador para a estimação dos componentes da variância chamado de algoritmo EM (Esperança Maximizada). Após o trabalho de Dempster *et al.* (1977), Aitkin *et al.* (1981) demonstraram que as diferenças significantes encontradas no trabalho de Benett (1976) desapareciam ao considerar a análise com as crianças agrupadas em classes e as que se submeteram ao ensino formal, não apresentaram qualquer diferença em relação às demais (Goldstein, 1995).

O trabalho clássico subsequente de Aitkin e Longford (1986) iniciou uma série de procedimentos, resultando nas idéias centrais dos modelos multiníveis e nos softwares desenvolvidos. Vários pesquisadores como Mason *et al.* (1983), Goldstein (1986), Raudenbush e Bryk (1986) e Longford (1987), desenvolveram técnicas e programas de computadores para ajustar modelos lineares hierárquicos que incorporam a condição dos erros distribuídos normalmente em vários níveis. Mas, a partir da década de 90 houve um interesse substancial na estimação de modelos hierárquicos não lineares ou modelos hieráquicos com variáveis dependentes limitadas; por exemplo, Goldstein (1991), Schall (1991), Breslow e Clayton (1993), Wolfinger (1993), Longford (1994), propuseram e implementaram procedimentos de estimação aproximados para os modelos multiníveis com variáveis discretas.

Os modelos lineares multiníveis com 2 e 3 níveis para aplicações específicas a dados educacionais e também para experimentos com medidas repetidas podem ser encontrados em Bryk e Raudenbush (1992). Longford (1993a) apresenta uma discussão mais teórica dos modelos multiníveis para análise fatorial, modelos com respostas categorizadas e modelos multivariados. Gibbons e Hedeker (1997) exploram modelos logístico e probit com efeitos aleatórios para dados com 3 níveis. Yau (2001) apresenta uma discussão e aplicação para modelar dados de sobrevivência com estrutura hierárquica. Neuhaus (2002), examina o efeito da resposta má classificada para dados binários longitudinais utilizando os procedimentos média-populacional, conhecido como modelo marginal, e grupo-específico, conhecido como modelo de efeito aleatório.

1.3 Objetivos

- Apresentar a teoria dos modelos multiníveis de 2 níveis para resposta binária com um enfoque na má classificação bem como os métodos de implementação na análise de grupo-específico, isto é, na análise dos modelos mistos lineares generalizados, quando as probabilidades de má classificação são conhecidas;
- Apresentar os resultados de simulação do modelo multinível para resposta binária má classificada, para o caso de duas funções de ligação: logit e probit, em duas situações:

na primeira, em que as probabilidades de má classificação são conhecidas e na segunda, em que são desconhecidas;

- Uma aplicação na área de saúde é também apresentada quando as probabilidades de má classificação são desconhecidas.

1.4 Organização dos capítulos

A presente dissertação encontra-se dividida em 7 capítulos. No primeiro capítulo, como já visto, tem-se uma breve revisão da idéia básica sobre os dados com estrutura hierárquica, mostrando as características desses dados, bem como a necessidade do uso de modelos multiníveis em dados com estrutura hierarquica. O capítulo 2 apresenta os principais conceitos de modelos lineares multíniveis e uma breve revisão dos modelos mistos lineares para variáveis respostas do tipo contínua e, modelos lineares generalizados para variáveis respostas do tipo discreta. O capítulo 3 trata de modelos mistos lineares generalizados para dados discretos, com ênfase em variáveis respostas binárias. O capítulo 4 discute os modelos mistos lineares generalizados para respostas binárias má classificadas. O capítulo 5 apresenta o comportamento das estimativas dos parâmetros dos modelos mistos lineares generalizados para respostas binárias má classificadas através das funções de ligação logit e probit, com base em simulações. No capítulo 6 encontra-se o resultado de uma aplicação na área de saúde considerando os modelos mistos lineares generalizados para respostas binárias má classificadas. Por fim, o último capítulo contém uma discussão sobre os resultados encontrados.

1.5 Plataformas computacionais

Existem vários programas disponíveis para modelos lineares multiníveis dos tipos: comercialmente distribuídos e gratuitos. Os programas comercialmente distribuídos são:

- O HLM 5 (Bryk e Raudenbush, 2002) que modela dados com 2 e 3 níveis. A estimação dos parâmetros é realizada através dos algoritmos EM (Esperança Maximizada) e QVP (Quase-verossimilhança Penalizada). O algoritmo EM calcula as estimativas de máxima verossimilhança restrita com alguns passos especiais para acelerar a convergência (acelerador de Aitkin). A rotina QVP resolve as equações de máxima verossimilhança aplicando a aproximação da integral de Laplace. Maiores detalhes podem ser encontrados no endereço http://www.ssicentral.com/hlm/hlm.htm;
- O MLn e MLwiN foram escritos por pesquisadores do projeto de modelos multiníveis do Instituto de Londres de Educação. O método de estimação dos parâmetros é realizado

através dos algoritmos: MQGI (Mínimos Quadrados Generalizados Iterativos), MQGIR (Mínimos Quadrados Generalizados Iterativos Restritos), QVP (Quase-verossimilhança Penalizada) e QVM (Quase-Verossimilhança Marginal). Para mais detalhes ver http: //statcomp.ats.ucla.edu/mlwin;

- O BMDP foi planejado por Jennrich e Schluchter (1986) e tem uma série de módulos que permite modelar dados multiníveis. O módulo BMDP-3V ajusta modelos com efeitos aleatórios agrupados e com covariáveis de efeitos fixos, os módulos BMDP-4V e BMDP-5V são para analisar dados com medidas repetidas com especial ênfase para dados não balanceados. As estimativas dos parâmetros através do método de máxima verossimilhança são realizadas por três algoritmos: Newton- Raphson, Scoring de Fisher e o algoritmo EM. (http://www.statsol.ie/bmdp/bmdp.htm);
- O VARCL contêm dois programas, VARL3 e VARL9, escritos por Longford (1993b). VARL9 permite modelar dados estruturados hierarquicamente com até 9 níveis, mas é restrito para modelos em que todas as inclinações são fixas, considerando o VARL3 sob 3 níveis ele permite inclinações aleatórias. O VARCL foi desenvolvido usando o algoritmo Scoring de Fisher para implementar as análises de coeficientes aleatórios sem permitir termos de interação entre variáveis de níveis diferentes. Além disso, ele pode ser utilizado para analisar dados multivariados. (http://www.assess.com/Software/VARCL.htm);
- O EGRET foi desenvolvido na escola de Saúde Pública da Universidade de Washington, USA. Este programa modela dados hierárquicos com 2 níveis, mas é limitado para os seguintes tipos de dados: categóricos, dados de caso-controle e dados de sobrevivência, muito comuns em estudos epidemiológicos e biomédicos. Este software não ajusta modelos para resposta Gaussiana e nem modelos com efeito aleatório para a variável resposta Poisson, porém a estimativa da máxima verossimilhança marginal pode ser obtida por quatro algoritmos: Newton-Raphson, Newton-Modificado, Quase-Newton e Nelder-Mead. (http://www.cytel.com/products/egret);
- O aML foi desenvolvido pelos economistas Lillard e Panis (2003). A estrutura do aML permite modelar dados hierárquicos com 2 e 3 níveis para dados categóricos e contínuos e, os parâmetros do modelo são estimados utilizando a Quadratura de Gauss-Hermite. (http://www.applied-ml.com);
- O SYSTAT é distribuído pelo SYSTAT Software Incorporation. Para ajustar dados com estrutura hierárquica de 2 níveis, utiliza o algoritmo EM para estimação da máxima verossimilhança marginal sendo um software que ajusta modelos multiníveis apenas para variável resposta Gaussiana. (http://www.systat.com);
- O SAS permite ajustar modelos de efeitos aleatórios de 2 níveis utilizando o procedimento MIXED, para variável resposta contínua, através do algoritmo de Newton-Raphson e o procedimento NLMIXED, para variável resposta discreta, através do método Quadratura

de Gauss-Hermite. (http://www.sas.com);

- O SPSS é um programa de análises estatísticas que permite modelar dados multiníveis e utiliza o módulo VARCOMP. O SPSS ajusta dados hierárquicos com até 2 níveis. (http://www.spss.com);
- O Stata também é um programa de análises estatísticas que permite a estimação de alguns modelos multiníveis para dados hierárquicos com 2 níveis, o módulo *loneway* ("long oneway") dá estimativas para o modelo nulo, o módulo de séries *xt* são utilizados para análises de dados longitudinais, o comando *xteg* estima o modelo com intercepto aleatório, os comandos *xtpois* e *xtprobit* estimam os modelos de regressão probit e Poisson, respectivamente. (http://www.stata.com).

Os programas distribuídos gratuitamente são :

- O GenStat, que foi desenvolvido em 1960 na Estação Experimental de Rothamsted para analisar dados de experimentos agrícolas e, a partir de 1989 foram adicionados ao programa algumas macros para análisar dados multiníveis balanceados sem limitações para os níveis de hierarquia. A estimação dos parâmetros é realizada através dos algoritmos MVR (máxima verossimilhança restrita), QVP e QVM. (http://www.biometris.nl/GenStat);
- O Mixed-Up é um conjunto de programas que ajusta modelos mistos de 2 níveis, incluindo modelos para respostas binárias, nominal, ordinal e para dados de contagem. O conjunto de programas Mixed foi desenvolvido por Hedeker e Gibbson (1996a, 1996b) na Universidade de Illinois em Chicago. O Mixed-Up utiliza dois algoritmos para a estimação da verossimilhança marginal, o método Scoring de Fisher modificado, para variável resposta Normal, através do programa MIXREG e o método Quadratura de Gauss-Hermite, para variável resposta binária, nominal, ordinal e Poisson, através dos programas MIXOR, MIXNO, MIXOR e MIXPREG, respectivamente. (http://tigger.uic.edu/%7Ehedeker/mix. html);
- O `R` é baseado na linguagem de alto nível S desenvolvido para análise, manipulação e apresentação gráfica de dados, foi inicialmente desenvolvido por Ihaka e Gentleman (1996). Este programa pode ser utilizado para modelar dados categóricos e contínuos estruturados hierarquicamente sem limitações para a quantidade de níveis e utiliza duas rotinas, EM com repetições por Newton-Raphson e QVP. (http://www.r-project.org).

A plataforma computacional utilizada para a realização deste trabalho foi o pacote estatístico `R`. Este pacote pode ser obtido em versões Windows, Linux, Unix e Macintosh. Uma vantagem dessa plataforma é sua flexibilidade em permitir a criação de novas funções e a possibilidade de modificações de muitas funções internas. A versão utilizada foi 2.0.

2 Conceitos básicos

2.1 Modelos lineares multiníveis com 2 níveis

O modelo linear multinível assume que existe um conjunto de dados hierárquicos, com uma única variável resposta que é medida no nível mais baixo da hierarquia e variáveis explicativas nos níveis existentes. Conceitualmente o modelo pode ser visto como um sistema hierárquico de equações de regressão.

Na abordagem dos modelos lineares hierárquicos ou modelos lineares multiníveis com 2 níveis, refere-se as unidades de nível 1 como "indivíduos" e as unidades de nível 2 como "grupo", (Snijders e Bosker, 1999). A quantidade de grupos é denotada por J e, o número de indivíduos no grupo, que pode variar de grupo para grupo é denotado por n_j para cada grupo j ($j = 1, 2, \ldots, J$). Neste caso, os dados não precisam ser balanceados, uma vez que o número de indivíduos pode variar de grupo para grupo.

Os modelos para o nível 1 podem ser desenvolvidos separadamente em cada grupo j, levando em consideração a variação dos interceptos bem como a variação das inclinações. Suponha que Y é uma variável resposta e X é uma única variável explicativa, o modelo para a análise no nível 1 é definido da seguinte forma

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij}, \tag{2.1}$$

com $i = 1, 2, \ldots, n_j$ e $j = 1, 2, \ldots, J$ onde

Y_{ij} é a variável resposta do i-ésimo indivíduo do nível 1 para o j-ésimo grupo do nível 2;

X_{ij} é a variável explicativa em sua medida original ou centrada na média geral $\overline{X}_{..}$ ou centrada na média de um grupo $\overline{X}_{.j}$, ou centrada em outro valor, medida no i-ésimo indivíduo e agrupado para o j-ésimo grupo;

β_{0j} é o intercepto para o j-ésimo grupo;

β_{1j} é a mudança esperada na variável resposta quando X_{ij} aumenta em uma unidade;

e_{ij} é o erro aleatório associado ao i-ésimo indivíduo agrupado para o j-ésimo grupo em que os erros são independentes e normalmente distribuídos, $N(0, \sigma^2)$.

A interpretação dos parâmetros do modelo, particularmente os interceptos β_{0j} para todo j, depende da forma como as variáveis explicativas são consideradas no nível 1. Pode-se afirmar que:

- Quando a variável explicativa do nível 1 é considerada em sua medida original (X_{ij}) o intercepto β_{0j} é o valor esperado da variável resposta Y_{ij} quando X_{ij} for igual a zero;
- Quando a variável explicativa do nível 1 estiver centrada na sua média geral ($\overline{X}_{..}$) o intercepto β_{0j} é o valor esperado da variável resposta Y_{ij} quando X_{ij} for igual a $\overline{X}_{..}$;
- Quando a variável explicativa do nível 1 estiver centrada na média de cada grupo ($\overline{X}_{.j}$) o intercepto β_{0j} é o valor esperado da variável resposta Y_{ij} quando X_{ij} for igual a $\overline{X}_{.j}$;
- E, finalmente, quando a variável explicativa do nível 1 estiver centrada em qualquer outro valor específico o intercepto β_{0j} é interpretado como o valor esperado da variável resposta Y_{ij} para esse valor específico.

Utilizando a variável resposta Y_{ij} e a variável explicativa X_{ij}, no modelo para o grupo, têm-se diferentes coeficientes de interceptos β_{0j}'s e inclinações β_{1j}'s com $j = 1, 2, \ldots, J$, iguais aos da equação (2.1). Portanto os coeficientes de regressão para os modelos no nível 2 podem ser escritos como:

$$\beta_{0j} = \gamma_{00} + u_{0j}\,, \tag{2.2}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}\,, \tag{2.3}$$

onde

γ_{00} é o valor esperado dos interceptos nos grupos;

γ_{10} é o valor esperado das inclinações na população do grupo;

u_{0j} e $u_{1j'}$ são variáveis aleatórias independentes com as seguintes suposições:

$$\bullet\, u_{0j} \sim N(0, \sigma^2_{u0})\ ;\ u_{1j} \sim N(0, \sigma^2_{u1}) \text{ e } cov(u_{0j}, u_{1j}) = \sigma_{u01}; \tag{2.4}$$

em que

u_{0j}'s e u_{1j}'s são independentes dos e_{ij}'s ;

u_{0j} é o efeito aleatório do j-ésimo grupo no intercepto β_{0j};

u_{1j} é o efeito aleatório do j-ésimo grupo na inclinação β_{1j};

σ^2_{u0} é a variância populacional dos interceptos;

σ^2_{u1} é a variância populacional das inclinações;

σ_{u01} é a covariância entre β_{0j} e β_{1j}.

Portanto, a partir de (2.2) a (2.4) temos que

$$\beta_{0j} \sim N(\gamma_{00}, \sigma^2_{u0}),\ \ \beta_{1j} \sim N(\gamma_{10}, \sigma^2_{u1}) \ \text{ e } \ cov(\beta_{0j}, \beta_{1j}) = \sigma_{u01}.$$

Das equações (2.2) e (2.3), nota-se que os interceptos e as inclinações variam aleatoriamente de grupo para grupo, logo os interceptos e as inclinações são diferentes e os seus

efeitos aleatórios, u_{0j} e u_{1j}, ajudam a esclarecer estas possíveis diferenças. Além disso, é possível predizer a variação entre as unidades do nível 2 introduzindo variáveis explicativas. Portanto, considere a variável explicativa Z_j. Assim, o modelo no nível 2 será da seguinte forma:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j}\,, \tag{2.5}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j}\,, \tag{2.6}$$

onde

γ_{00} é o valor esperado dos interceptos quando $Z_j = 0$;

γ_{10} é o valor esperado das inclinações quando $Z_j = 0$;

γ_{01} é o coeficiente de regressão associado à variável explicativa do nível 2 relativo ao intercepto;

γ_{11} é o coeficiente de regressão associado à variável explicativa do nível 2 relativo à inclinação;

u_{0j} e $u_{1j'}$ são variáveis aleatórias com as mesmas suposições (2.4), em que

u_{0j}'s e u_{1j}'s são independentes dos e_{ij}'s ;

u_{0j} é o efeito aleatório do j-ésimo grupo no intercepto β_{0j};

u_{1j} é o efeito aleatório do j-ésimo grupo na inclinação β_{1j};

σ^2_{u0} é a variância populacional dos interceptos encontrada após a predição do coeficiente β_{0j} na presença da variável do nível 2, Z_j;

σ^2_{u1} é a variância populacional das inclinações encontrada após a predição do coeficiente β_{1j} na presença da variável do nível 2, Z_j;

σ_{u01} é a covariância entre β_{0j} e β_{1j}.

No modelo para o nível 2, os β's são tratados como variáveis respostas em modelos de regressão com variável explicativa Z_j para a população de grupos (Snijders e Bosker, 1999). Pode-se considerar também Z_j em sua medida original ou centrada na média geral ($\overline{Z}_{..}$) ou centrada na média de cada grupo ($\overline{Z}_{.j}$), da mesma forma visto anteriormente para a variável explicativa no nível 1 (2.1).

Substituindo as equações (2.5) e (2.6) em (2.1), tem-se o modelo combinado

$$Y_{ij} = \gamma_{00} + \gamma_{01} Z_j + \gamma_{10} X_{ij} + \gamma_{11} Z_j X_{ij} + u_{0j} + u_{1j} X_{ij} + e_{ij}. \tag{2.7}$$

O modelo combinado (2.7) envolve o segmento $\gamma_{00} + \gamma_{01} Z_j + \gamma_{10} Z_{ij} + \gamma_{11} Z_j X_{ij}$, que é a parte fixa do modelo, por conter todos os coeficientes fixos e o segmento $u_{0j} + u_{1j} X_{ij} + e_{ij}$,

que é a parte aleatória ou complexa. Note que X_{ij} e Z_j são variáveis explicativas dos níveis 1 e 2, respectivamente e $Z_j X_{ij}$ é o termo de interação entre os níveis que aparece no modelo como consequência de modelar a variação do coeficiente de inclinação β_{1j}. Ainda tem-se que os erros nas unidades de nível 1 (indivíduos) não são independentes, ou seja, há uma dependência entre os indivíduos agrupados dentro de cada um dos grupos em termos de u_{0j} e u_{1j}. Além disso, as variâncias dos erros podem ser heterogêneas se u_{0j} e u_{1j} assumirem diferentes valores dentro do grupo.

O grau de dependência dentro dos grupos é medido pelo coeficiente de correlação intra-classe, $\rho(Y_{ij}, Y_{i'j})$. O coeficiente de correlação intra-classe pode ser estimado por meio dos modelos lineares multiníveis a partir do modelo intercepto, ou seja, por um modelo que não inclui variáveis explicativas. Este modelo é obtido removendo todos os termos que contém as variáveis X ou Z no modelo (2.7). O modelo nulo será da forma

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}. \tag{2.8}$$

A partir do modelo nulo (2.8), podemos obter a partição básica da variabilidade nos dados entre os dois níveis. A variância total de Y pode ser decomposta como a soma das variâncias dos níveis 1 e 2,

$$Var(Y_{ij}) = Var(u_{0j}) + Var(e_{ij}) = \sigma^2_{u0} + \sigma^2,$$

a covariância entre 2 indivíduos (i e i', com $i \neq i'$) no mesmo grupo j é igual a variância da contribuição de u_{0j} que é compartilhada por estes indivíduos,

$$cov(Y_{ij}, Y_{i'j}) = var(u_{0j}) = \sigma^2_{u0},$$

e sua correlação intra-classe é

$$\rho(Y_{ij}, Y_{i'j}) = \frac{\sigma^2_{u0}}{\sigma^2_{u0} + \sigma^2}.$$

A correlação intra-classe é uma estimativa da proporção da variância populacional explicada pela estrutura do agrupamento. Assim, o modelo nulo (2.8) estabelece simplesmente que a correlação intra-classe estimada é igual a proporção estimada da variância no nível grupo quando comparada com a variância total estimada.

2.2 Métodos de estimação dos modelos lineares multiníveis

Os dois principais métodos de estimação dos modelos lineares multiníveis são os métodos de máxima verossimilhança e de máxima verossimilhança restrita, (Snijders e Bosker, 1999).

Os dois métodos pouco diferem com relação as estimativas dos coeficientes de regressão, mas eles diferem com relação as estimativas dos componentes da variância. O estimador de máxima verossimilhança restrita estima os componentes da variância levando em conta os números de graus de liberdade usados na estimativa dos efeitos fixos, enquanto o estimador de máxima verossimilhança não considera essa informação. O resultado disso é que o estimador de máxima verossimilhança restrita produz estimativas menos viesadas para os componentes da variância no caso de amostras pequenas, em relação ao estimador de máxima verossimilhança .

Vários algoritmos são disponíveis para determinar estas estimativas. Os principais algoritmos são: EM (esperança maximizada), Scoring Fisher, MQGI (mínimos quadrados generalizados iterativos) e MQGIR (MQGI restrito). Estes algoritmos utilizam procedimentos iterativos e produzem, em geral, estimativas idênticas, mas com diferente número de iterações. Contudo, algumas vezes ocorrem problemas computacionais (um algoritmo pode convergir e outro não) e o tempo de execução do algoritmo pode ser diferente. Os principais problemas de não convergência são tamanhos de amostras pequenos, má especificação dos modelos e a estimação de muitos componentes aleatórios, que na verdade são próximos ou iguais a zero.

O algoritmo EM é o método geral para calcular as estimativas de máxima verossimilhança em casos em que há dados faltantes. O EM é construído de tal forma que a convergência é garantida, mas frequentemente lenta. O algoritmo de Newton-Rapson pode ser utilizado tanto para os estimadores de máxima verossimilhança quanto para os de máxima verossimilhança restrita. O estimador de mínimos quadrados generalizados iterativos é utilizado para calcular a máxima verossimilhança e o estimador de mínimos quadrados generalizados iterativos restritos é utilizado para calcular a máxima verossimilhança restrita. Quando se supõe conhecidos os componentes da variância, os coeficientes de regressão podem ser estimados diretamente por MQG. Convencionalmente, quando se consideram conhecidos os coeficientes de regressão podemos estimar os componentes da variância. Estes dois processos de estimação podem ser alternados: podemos utilizar valores provisórios para os parâmetros da parte aleatória para estimar os coeficientes de regressão e depois utilizar a última estimativa do coeficiente de regressão para estimar os parâmetros da parte aleatória até que este processo iterativo convirja.

2.3 Modelo misto linear

O modelo multinível linear pode ser visto como um caso particular do modelo misto linear ou modelo de efeitos aleatórios (Renard, 2002). Considere o modelo misto linear padrão de

2 níveis com a seguinte estrutura

$$y_{ij} = \mathbf{x}_{ij}^T\beta + \mathbf{z}_{ij}^T\mathbf{u}_j + e_{ij}, \tag{2.9}$$

com $i = 1, \ldots, n_j$ e $j = 1, \ldots, J$ onde

y_{ij} é o valor observado do i-ésimo indivíduo no j-ésimo grupo, $\mathbf{x}_{ij} = (1, x_{1ij}, \ldots, x_{p\,ij})^T$ é o vetor de regressores associado ao efeito fixo, $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$ e $\mathbf{z}_{ij} = (1, z_{1ij}, \ldots, z_{qij})^T$ é o vetor de regressores associado ao efeito aleatório, $\mathbf{u}_j = (u_{0j}, u_{1j}, \ldots, u_{qj})^T$. Os efeitos aleatórios, $\mathbf{u_j}$, seguem uma distribuição normal multivariada com média $\mathbf{0}$ e matriz de covariância $\mathbf{D}(\theta)$ que depende de um vetor de parâmetros de dispersão θ e e_{ij} é o erro residual que segue uma distribuição normal com média 0 e variância σ^2 com a suposição de que são independentes dos $\mathbf{u}_j$.

Alternativamente, podemos escrever (2.9) em termos matriciais como:

$$\mathbf{y}_j = \mathbf{x}_j\beta + \mathbf{z}_j\mathbf{u}_j + \mathbf{e}_j$$

com $j = 1, \ldots, J$ em que,

- $\mathbf{y}_j$ é um vetor de dimensão $n_j \times 1$ para as observações no j-ésimo grupo;
- $\mathbf{x}_j$ é a matriz de regressores de dimensão $n_j \times (p+1)$ associada aos efeitos fixos para as observações no j-ésimo grupo;
- β é o vetor de coeficientes fixos de dimensão $(p+1) \times 1$;
- $\mathbf{z}_j$ é a matriz de regressores de dimensão $n_j \times (q+1)$ associada aos efeitos aleatórios para as observações no j-ésimo grupo;
- $\mathbf{u}_j$ é o vetor de efeitos aleatórios de dimensão $(q+1) \times 1$ para o j-ésimo grupo, onde $\mathbf{u}_j$ segue uma distribuição normal multivariada com média $\mathbf{0}$ e matriz de covariância $\mathbf{D}(\theta)$ de dimensão $(q+1) \times (q+1)$ que depende de um vetor de parâmetros de dispersão θ;
- $\mathbf{e}_j$ é o vetor de erros residuais de dimensão $n_j \times 1$ para o j-ésimo grupo que segue uma distribuição normal multivariada com média $\mathbf{0}$ e matriz de covariância $\sigma^2\mathbf{I}$, onde $\mathbf{I}$ é a matriz identidade de dimensão $n_j \times n_j$. A especificação do modelo é completa com a suposição de que os $\mathbf{e}_j$ são independentes dos $\mathbf{u}_j$.

A estimação dos parâmetros pode ser feita maximizando a função de verossimilhança. Para esta finalidade, pode ser executado o algoritmo de Newton-Raphson ou a Esperança Maximizada (EM). Um procedimento equivalente, chamado mínimos quadrados generalizados iterativos (MQGI), foi proposto por Goldstein (1986). Este algoritmo iterage entre a estimação dos parâmetros dos efeitos fixos e aleatórios utilizando o princípio de mínimos quadrados generalizados padronizados, sendo um procedimento atrativo e computacionalmente eficiente comparado com a maximização direta da verossimilhança, além disso este

algoritmo pode ser modificado para obter as estimativas por máxima verossimilhança restrita (EMVR) que são não viesadas para os parâmetros aleatórios e pode ser referido como (MQGIR).

A maioria dos softwares, R, SAS, Stata ou SPSS, como comentado anteriormente, têm "macros" ou procedimentos para ajustar modelos mistos lineares. Nos softwares especializados em modelos multiníveis como HLM e MLwiN também são disponíveis as "macros" referidas. Eles permitem ajustar modelos mistos lineares (utilizando o algoritmo MQGI ou MQGIR), manipular dados com resposta discreta, além de oferecem os métodos de estimação por bootstrap paramétrico, não paramétrico, Monte Carlo de Cadeia de Markov (MCCM) para ajustar modelos Bayesianos, bem como várias "macros" para trabalhar com outros tipos de respostas como, por exemplo, resposta multi-categóricas.

2.4 Modelo linear generalizado

O modelo misto linear apresentado na seção anterior assume variável dependente contínua com erros independentes e normalmente distribuídos. Quando a natureza da variável resposta é discreta não se deve levar em consideração para a análise uma distribuição contínua, pois isto poderá levar a conclusões errôneas. Os métodos de regressão linear não devem ser aplicados para estes tipos de variáveis por duas razões.

A primeira é que para respostas dicotômicas, que podem ser representadas como tendo valores 0 e 1 (por exemplo, fracasso e sucesso, respectivamente) os valores estimados por estes métodos podem ser negativos ou maiores do que 1. Similarmente, para dados de contagem, os valores estimados podem ser negativos. A segunda razão é de natureza mais técnica, o fato é que as variáveis discretas frequentemente tem uma relação natural entre a média e variância da distribuição e em termos de modelos multiníveis, isto conduz a uma relação entre os parâmetros da parte fixa e os parâmetros da parte aleatória. Os modelos mais conhecidos para estes tipos de dados são: o modelo de regressão logística para dados dicotômicos e o modelo de regressão de Poisson para dados de contagem. Na literatura, estes modelos são casos particulares da família de Modelos Lineares Generalizados (MLG), conforme refere McCullagh e Nelder (1989).

Alguns modelos estatísticos que são casos especiais de MLG são os seguintes: modelo clássico de regressão com erro normal; modelo clássico de análise de variância e de covariância com erro normal; modelo de análise de variância com efeitos aleatórios, sendo um caso particular do modelo gama; modelo de regressão de Poisson; modelo logístico para proporções.

O modelo linear generalizado é definido a partir de uma distribuição de probabilidades para a variável resposta, membro da família exponencial de distribuições, um preditor linear e uma função de ligação entre a média da variável resposta e a estrutura linear.

Assim, o MLG consiste dos seguintes componentes:

a) O componente aleatório.

Suponha $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias independentes cada uma com densidade da forma exponencial, isto é,

$$f(\theta_i, \phi, y_i) = exp\left[\frac{(y_i\theta_i - b(\theta_i))}{a(\phi)} + c(\phi, y_i)\right], \tag{2.10}$$

onde $b(.)$ e $c(.)$ são funções conhecidas; ϕ é o parâmetro de dispersão e θ é o parâmetro canônico ou natural que caracteriza a distribuição em (2.10). A função $a(\phi)$ tem a forma $a(\phi) = \phi/w_i$, onde w_i é um peso conhecido que pode variar de observação para observação. A expressão (2.10) acima é conhecida como família exponencial uniparamétrica com parâmetro canônico θ_i, se ϕ é conhecido, mas se ϕ é desconhecido então (2.10) pode ou não pertencer a família exponencial biparamétrica (McCullagh e Nelder, 1989).

Para a determinação dos dois primeiros momentos da variável Y, escrevemos a log-verossimilhança como $l(\theta_i, \phi, y_i) = \log f(\theta_i, \phi, y_i)$ que corresponde a contribuição da i-ésima unidade de investigação para o logaritmo da função de verossimilhança. Assim, tem-se de (2.9) que

$$l(\theta_i, \phi, y_i) = \frac{(y_i\theta_i - b(\theta_i))}{a(\phi)} + c(y_i, \phi), \tag{2.11}$$

e, derivando em relação a θ_i

$$\frac{\partial l}{\partial \theta_i} = \frac{(y_i - b'(\theta_i))}{a(\phi)}, \qquad \frac{\partial^2 l}{\partial \theta_i^2} = -\frac{b''(\theta_i)}{a(\phi)}.$$

Considerando as relações

$$\mathrm{E}\left(\frac{\partial l}{\partial \theta_i}\right) = 0 \quad \text{e} \quad \mathrm{E}\left(\frac{\partial^2 l}{\partial \theta_i^2}\right) + \mathrm{E}\left(\frac{\partial l}{\partial \theta_i}\right)^2 = 0$$

que sob condições de regularidade satisfaz a família exponencial.

Então, a partir da expressão (2.11) tem-se que a média e a variância são dadas por:

$$\mathrm{E}(Y_i) = b'(\theta_i) = \mu_i \quad \text{e} \quad \mathrm{Var}(Y_i) = b''(\theta_i)a(\phi).$$

A variância pode ser expressa como uma função de sua média, isto é,

$$\mathrm{Var}(Y_i) = a(\phi)V(\mu_i) = a(\phi)V_i,$$

com $V(.)$ representando a função de variância, que caracteriza a distribuição da família exponencial e $a(\phi)$, fator escala, é uma constante conhecida para alguns membros da família

de modelos lineares generalizados, enquanto para outros é um parâmetro adicional a ser estimado.

b) O componente sistemático é o preditor linear.

Considere a estrutura linear nos parâmetros, que relaciona η_i com o vetor de variáveis explicativas $\mathbf{x_i} = (x_{i1}, \ldots, x_{ip})^T$, dado por

$$\eta_i = \mathbf{x}_i^T \beta, \;\; i = 1, \ldots, n$$

e portanto,

$$\eta_i = \sum_{j=1}^{p} x_{ij}\beta_j,$$

onde η_i é a i-ésima componente do vetor $\eta = (\eta_1, \ldots \eta_n)'$ de preditores lineares; β_j é o j-ésimo componente do vetor $\beta = (\beta_1, \ldots \beta_p)^T$ de parâmetros, e $\mathbf{x}_i^T = (x_{i1}, \ldots, x_{ip})$ é a i-ésima linha da matriz $\mathbf{x}$ do modelo, sendo esta conhecida. A função linear (η_i) dos parâmetros desconhecidos (β_j) é chamada de preditor linear.

c) A função de ligação.

A função de ligação relaciona o valor esperado (μ_i) da variável y_i ao preditor linear (η_i), através de uma função monótona e diferenciável, $g(.)$, através de

$$g(\mu_i) = \eta_i = \mathbf{x}_i^T \beta.$$

Para construir um MLG é necessário decidir pela escolha da distribuição de probabilidades para a variável resposta, o preditor linear e a função de ligação. No modelo linear clássico tem-se uma ligação do tipo identidade $(\eta = \mu)$ sendo plausível para modelar dados normais no sentido em que η e μ podem assumir qualquer valor na linha real. Entretanto, certas restrições surgem quando se trabalha, por exemplo, com a distribuição de Poisson em que $\mu > 0$, neste caso, a função de ligação identidade não deve ser utilizada, pois $\hat{\eta}$ poderá assumir valores negativos dependendo dos valores obtidos para $\hat{\beta}$. Já para modelos que asumem a distribuição binomial, onde temos que $0 < \mu < 1$, existe a restrição de que o domínio da função de ligação esteja no intervalo (0,1) enquanto seu contradomínio é a linha real. Neste caso, se Y^* é a proporção de sucessos em n ensaios independentes, cada um com probabilidade de ocorrência μ e, assumindo que $nY^* \sim B(n, \mu)$, a densidade de nY^* é dada por

$$f(\mu, ny^*) = \binom{n}{ny^*}\mu^{ny^*}(1-\mu)^{n-ny^*} = \exp\Big\{\Big[\log\binom{n}{ny^*}\Big] + ny^*\log\Big(\frac{\mu}{1-\mu}\Big) + n\log(1-\mu)\Big\},$$

ou equivalentemente,

$$f(\mu, ny^*) = \exp\Big\{n\Big[y^*\log\Big(\frac{\mu}{1-\mu}\Big) + \log(1-\mu)\Big] + \log\binom{n}{ny^*}\Big\},$$

onde $0 < y^* < 1$ e $0 < \mu < 1$. Logo, $\theta = \log\{\mu/(1-\mu)\}$, $b(\theta) = \log(1+e^{\theta})$, $a(\phi) = 1/n$, $w = 1$, $\phi = 1/n$, $c(\phi, y^*) = \log(\binom{n}{ny^*})$ e a função de variância é dada por $V(\mu) = \mu(1-\mu)$.

A função de ligação canônica, ocorre quando o parâmetro natural (θ) e o preditor linear (η) coincidem, isto é, quando

$$\theta_i = \eta_i = \sum_{j=1}^{p} x_{ij}\beta_j, \;\; i = 1, \ldots, n.$$

Sempre que estas funções existirem, obtem-se estimativas de máxima verossimilhança únicas para os parâmetros $\beta_1, \ldots, \beta_p$, ou seja, as ligações canônicas garantem a concavidade de $L(\beta; y)$ e conseqüentemente muitos resultados assintóticos são obtidos mais facilmente.

Ao se considerar a distribuição binomial na família exponencial, tem-se que a função de ligação canônica é:

$$\textbf{Logit} : \eta_1 = \log\{\frac{\mu}{1-\mu}\},$$

chamado modelo logístico. Outras funções de ligação que podem ser consideradas são:

- **Probit:** $\eta_2 = \Phi^{-1}(\mu)$, onde $\Phi(.)$ é a função distribuição Normal acumulada;
- **Complemento log-log:** $\eta_3 = \log\{-\log(1-\mu)\}$.

3 Modelo misto linear generalizado

3.1 Considerações iniciais

O modelo misto linear generalizado (MMLG) é uma generalização do modelo misto linear (MML) proposto por Laird e Ware (1982). O MMLG também é considerado um MLG com termos aleatórios no preditor linear. Segundo Breslow e Clayton (1993), os MMLG's são utilizados para:

- tratar a superdispersão observada entre os resultados que tem distribuição Binomial ou Poisson;
- modelar a dependência entre as observações em estudos longitudinais;
- produzir estimativas não viesadas em problemas multiparamétricos.

Segundo Breslow (2003), o MMLG é um modelo adequado para ajustar um conjunto com N observações de uma variável resposta univariada, y_i, junto com um vetor $\mathbf{x}_i$ de regressores $(p+1)$-dimensional associado ao efeito fixo, β e um vetor $\mathbf{z}_i$ de regressores $(q+1)$-dimensional associado ao efeito aleatório, $\mathbf{u}$, em que $i = 1, \ldots N$. Supondo que dado um vetor q-dimensional de efeitos aleatórios, $\mathbf{u}$, os $\mathbf{y}_i$ são condicionalmente independentes com média e variância especificado por um MLG, temos que

$$\mathrm{E}(y_i \mid \mathbf{u}) = \mu_i^{\mathbf{u}} \text{ e } \mathrm{Var}(y_i \mid \mathbf{u}) = a_i(\phi)v(\mu_i^{\mathbf{u}}), \tag{3.1}$$

onde $v(.)$ é a função de variância especificada e a função $a(\phi)$ tem a forma $a_i(\phi) = \phi/w_i$, onde w_i é um peso conhecido e ϕ é o parâmetro de dispersão que pode ou não ser conhecido. A média condicional é relacionada com o preditor linear,

$$\eta_i^{\mathbf{u}} = \mathbf{x}_i^T\beta + \mathbf{z}_i^T\mathbf{u}, \tag{3.2}$$

pela função de ligação, $g(\mu_i^{\mathbf{u}}) = \eta_i^{\mathbf{u}}$, com inversa $h = g^{-1}$, onde β é um vetor de efeitos fixos, portanto, a média condicional é dada por $\mathrm{E}(y_i \mid \mathbf{u}) = h(\mathbf{x}_i^T\beta + \mathbf{z}_i^T\mathbf{u})$.

A especificação do modelo é completa com a suposição de que $\mathbf{u}$ segue uma distribuição normal multivariada com média $\mathbf{0}$ e matriz de covariância $\mathbf{D}(\theta)$, que depende de um vetor de parâmetros de dispersão θ.

Segundo Schall (1991), a análise de um modelo misto linear generalizado pode proceder-se da seguinte forma: especificar uma distribuição da classe parametrizada para os efeitos aleatórios ($\mathbf{u}$) e em seguida, estimar os parâmetros da parte fixa (β) e os da parte aleatória

($\mathbf{u}$) pela máxima verossimilhança baseado na distribuição marginal das observações y_i. Este procedimento é normalmente utilizado no modelo misto linear onde as distribuições dos efeitos aleatórios e a distribuição condicional de $\mathbf{y_i}$ são assumidas serem normais. Anderson e Aitkin (1985) e Im e Gianola (1988) utilizam a estimação de máxima verossimilhança em modelos logístico e probit, onde os efeitos aleatórios são assumidos ter uma distribuição normal e a distribuição de Y é binomial.

Contudo, em MMLG, exceto em MLG, há dois problemas com o procedimento acima: primeiro, frequentemente é difícil justificar uma distribuição particular para os efeitos aleatórios; segundo, a estimação de máxima verossimilhança baseada na distribuição marginal das observações envolve integrais sem soluções de forma fechada e portanto a estimação dos efeitos aleatórios devem ser feitas por procedimentos que incluam integrações numéricas.

3.2 Algoritmos de estimação

Os modelos lineares hierárquicos são modelos apropriados para dados normalmente distribuídos e as estimações dos parâmetros são feitas utilizando uma mistura de estimação empírica de Bayes e de máxima verossimilhança marginal. Quando a variável resposta é discreta as equações de verossimilhança não têm uma solução simples de forma fechada.

A utilização da verossimilhança aproximada tem sido sugerida por vários autores. Green (1987), baseando-se nos trabalhos de Laird (1978) e de Stiratelli e Ware (1984), explorou o método chamado de Quase-verossimilhança Penalizada (QVP), como uma aproximação do método de Bayes, estando disponível para inferências em modelos hierárquicos onde o foco principal é a estimação do efeito aleatório. Um método alternativo ao algoritmo da QVP desenvolvido por Schall (1991) utiliza uma linearização da média condicional como uma função dos efeitos fixos e aleatórios.

Breslow e Clayton (1993), mostraram que utilizando o método de Laplace para a aproximação da integral, obtem-se o algoritmo QVP. Eles também utilizam a Quase-verossimilhança Marginal (QVM) que é o nome que eles dão ao procedimento proposto por Goldstein (1991). Goldstein (1991) propôs um procedimento alternativo (QVM) para a estimação de modelos multíniveis não lineares, incluindo o modelo logístico com efeitos aleatórios. QVP e QVM podem ser vistas como procedimentos iterativos que requerem o ajuste do modelo misto linear por meio de uma expansão da série de Taylor de $1^{\underline{a}}$ ordem. Rodriguez e Goldman (1995) mostraram que estes métodos aproximados podem ser viesados sob certas circunstâncias, por exemplo, com resposta binária.

Para reduzir a extensão dos vieses, alguns autores têm sugerido a introdução do termo correção de viés (Breslow e Lin, 1995; Lin e Breslow, 1996) ou a utilização de bootstrap iterativo (Kuk, 1995). Goldstein e Rasbash (1996) mostraram que a inclusão da expansão

da série de Taylor de 2ª ordem reduz consideravelmente os vieses descritos por Rodriguez e Goldman (1995). Breslow (2003), mostrou que o método da QVP apresenta, ainda, um bom desempenho em comparação com procedimentos mais elaborados, como por exemplo, os métodos da quadratura Guassiana adaptativa, Viés Corrigido de Quase-verossimilhança Penalizada (VCQVP) e Monte Carlo de Cadeias de Markov (MCCM), em muitas situações práticas.

3.2.1 Estimação por quase-verossimilhança penalizada

O método de quase-verossimilhança foi proposto por Wedderburn (1974) e examinado extensivamente por McCullagh & Nelder (1989). Este método não requer a especificação da forma exata da distribuição, mas requer a especificação de relações funcionais para os seus dois primeiros momentos, isto é, a especificação de relacionamento entre a média e uma função de variância que expresse a dependência da variância com a média. Assim, este método pode ser usado com uma variedade de variáveis respostas, e a distribuição da variável resposta ficará determinada quando a função de variância escolhida coincidir com a função de variância de alguma distribuição da família exponencial (Paula, 2003). O critério penalizador, penaliza a medida da bondade de ajuste do modelo, sendo representado por um termo aditivo.

A função de quase-verossimilhança penalizada para estimar os parâmetros (β, θ) do MMLG apresentada em Breslow (2003) é dada por

$$L = \frac{1}{\sqrt{(2\pi)^q |\mathbf{D}(\theta)|}} \int_{R^q} \exp\Big[-\frac{1}{2\phi} \sum_{i=1}^{N} d_i(y_i, \mu_i^{\mathbf{u}}) - \frac{1}{2}\mathbf{u}^T \mathbf{D}^{-1}(\theta)\mathbf{u} \Big] d\mathbf{u}, \tag{3.3}$$

onde

$$d_i(y, \mu) = -2 \int_y^{\mu} \frac{y - b}{a_i v(b)} db,$$

é o desvio ponderado. Se a variável resposta Y é Gaussiana e a função de ligação, $g(.)$, é a identidade, a integral em (3.3) pode ser avaliada em forma fechada. Por outro lado, como mencionado anteriormente, se a variável resposta Y é discreta, a maximização desta expressão deve ser feita por procedimentos que incluam integrações numéricas a cada ciclo da iteração.

Breslow (2003), apresenta o processo de estimação do vetor de parâmetros para os efeitos fixos (β) e para os efeitos aleatórios $(\mathbf{u})$ de um modelo multinível de 2 níveis, utilizando o método da QVP, seguindo o procedimento proposto por Goldstein (1991) que se apresenta a seguir.

Considere o modelo de 2 níveis com J grupos tendo n_j observações , $j = 1, \ldots J$, com os efeitos aleatórios $(\mathbf{u}_j)$ assumidos independentes entre os grupos. A i-ésima observação para

o j-ésimo grupo pode ser escrita como

$$y_{ij} = \mu_{ij}^{\mathbf{u}} + e_{ij} = h(\mathbf{x}_{ij}^T\beta + \mathbf{z}_{ij}^T\mathbf{u}_j) + e_{ij}, \tag{3.4}$$

com $i = 1, \ldots, n_j$ e $\text{Var}(e_{ij}) = a_j(\phi)v(\mu_{ij}^{\mathbf{u}})$, onde h é a inversa da função de ligação , $\mathbf{x}_{ij}$ e $\mathbf{z}_{ij}$ são os vetores de regressores associados aos efeitos fixos (β) e aleatórios ($\mathbf{u}_j$), respectivamente.

O objetivo é linearizar o modelo para associá-lo com o modelo teórico com distribuição normal, ou seja, a idéia é linearizar a função de ligação inversa utilizando uma expansão da série de Taylor de 1ª ou de 2ª ordem em torno dos valores de β e $\mathbf{u}_j$ inicialmente definidos a partir dos métodos de Quase-verossimilhança Penalizada de primeira ordem (QVP1) ou Quase-verossimilhança Penalizada de segunda ordem (QVP2). Os passos para a linearização da função da ligação são apresentados abaixo

- **Algoritmo da QVP de primeira ordem (QVP1)**

Primeiro passo: Expandir h sobre as estimativas atuais de $(\hat{\beta}, \tilde{\mathbf{u}}_j)$. A partir da equação (3.4) pode-se escrever a função como:

$$f(\beta, \mathbf{u}_j) = h(\mathbf{x}_{ij}^T\beta + \mathbf{z}_{ij}^T\mathbf{u}_j), \tag{3.5}$$

as primeiras derivadas da função com relação aos parâmetros são dadas por

$$\frac{\partial f(\beta, \mathbf{u}_j)}{\partial \beta} = h^{'}(\mathbf{x}_{ij}^T\beta + \mathbf{z}_{ij}^T\mathbf{u}_j)\,\mathbf{x}_{ij}^T, \qquad \frac{\partial f(\beta, \mathbf{u}_j)}{\partial \mathbf{u}_j} = h^{'}(\mathbf{x}_{ij}^T\beta + \mathbf{z}_{ij}^T\mathbf{u}_j)\,\mathbf{z}_{ij}^T.$$

A aproximação da série de Taylor de 1ª ordem da função em torno dos parâmetros estimados $(\hat{\beta}, \tilde{u}_j)$ é construída da seguinte forma

$$\begin{aligned} f(\beta, \mathbf{u}_j) &\approx f(\hat{\beta}, \tilde{\mathbf{u}}_j) + \left.\frac{\partial f(\beta, \mathbf{u}_j)}{\partial \beta}\right|_{(\hat{\beta}, \tilde{\mathbf{u}}_j)} (\beta - \hat{\beta}) + \left.\frac{\partial f(\beta, \mathbf{u}_j)}{\partial \mathbf{u}_j}\right|_{(\hat{\beta}, \tilde{\mathbf{u}}_j)} (\mathbf{u}_j - \tilde{\mathbf{u}}_j) \\ &\approx \hat{\mu}_{ij}^{\mathbf{u}} + h^{'}(\mathbf{x}_{ij}^T\hat{\beta} + \mathbf{z}_{ij}^T\tilde{\mathbf{u}}_j)\ \mathbf{x}_{ij}^T\ (\beta - \hat{\beta}) + h^{'}(\mathbf{x}_{ij}^T\hat{\beta} + \mathbf{z}_{ij}^T\tilde{\mathbf{u}}_j)\ \mathbf{z}_{ij}^T(\mathbf{u}_j - \tilde{\mathbf{u}}_j) \\ &\approx \hat{\mu}_{ij}^{\mathbf{u}} + h^{'}(\hat{\eta}_{ij}^{\mathbf{u}})\,[\mathbf{x}_{ij}^T(\beta - \hat{\beta}) + \mathbf{z}_{ij}^T(\mathbf{u}_j - \tilde{\mathbf{u}}_j)], \end{aligned} \tag{3.6}$$

substituindo em (3.4) obtemos

$$y_{ij} \approx \hat{\mu}_{ij}^{\mathbf{u}} + h^{'}(\hat{\eta}_{ij}^{\mathbf{u}})\,[\mathbf{x}_{ij}^T(\beta - \hat{\beta}) + \mathbf{z}_{ij}^T(\mathbf{u}_j - \tilde{\mathbf{u}}_j)] + e_{ij}. \tag{3.7}$$

Segundo passo: Aplicar a função g nos dados y_{ij}, e expandi-la em série de Taylor de 1ª ordem em torno do valor estimado $\hat{\mu}_{ij}^{\mathbf{u}}$.

A aproximação da série de Taylor de 1ª ordem da função g em torno do valor estimado é construída da seguinte maneira

$$g(y_{ij}) \approx g(\hat{\mu}_{ij}^{\mathbf{u}}) + \left.\frac{\partial g(y_{ij})}{\partial y_{ij}}\right|_{(\hat{\mu}_{ij}^{\mathbf{u}})} (y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})$$

$$\begin{aligned} &\approx g(\hat{\mu}_{ij}^{\mathbf{u}}) + g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}) \\ &\approx \hat{\eta}_{ij}^{\mathbf{u}} + g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}), \end{aligned} \tag{3.8}$$

então, podemos escrever (3.8) da seguinte forma

$$\begin{aligned} y_{ij}^{*} &= \hat{\eta}_{ij}^{\mathbf{u}} + g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}) \\ &= \mathbf{x}_{ij}^{T}\hat{\beta} + \mathbf{z}_{ij}^{T}\tilde{\mathbf{u}}_{j} + g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}). \end{aligned}$$

Terceiro passo: Associar o modelo linearizado ao modelo teórico com distribuição normal. Denotando $g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})$ de e_{ij}^{*} e assumindo que y_{ij}^{*} é distribuido normalmente, podemos escrever o modelo misto linear para y_{ij}^{*} da seguinte maneira

$$y_{ij}^{*} = \mathbf{x}_{ij}^{T}\beta + \mathbf{z}_{ij}^{T}\mathbf{u}_{j} + e_{ij}^{*}, \tag{3.9}$$

em que

$$\mathbf{u}_{j} \sim N(0, \mathbf{D}(\theta)) \text{ e } e_{ij}^{*} \sim N(0, [g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})]^{2} a_{\mathbf{j}}(\phi) v(\hat{\mu}_{ij}^{\mathbf{u}})), \tag{3.10}$$

onde

$$\mathbf{D}(\theta) = \begin{bmatrix} \sigma_{u0}^{2} & \sigma_{u01} \cdots\cdots & \sigma_{u0q} \\ \sigma_{u01} & \sigma_{u1}^{2} \cdots\cdots & \sigma_{u1q} \\ \vdots & \ddots & \vdots \\ \sigma_{u0q} & \sigma_{u1q} \quad \ldots & \sigma_{uq}^{2} \end{bmatrix}.$$

As estimativas atualizadas de $(\beta, \mathbf{u}_{j}, \theta)$ são obtidas resolvendo o modelo misto linear definido por (3.9) e (3.10) simultaneamente, isto é, utilizando o algoritmo QVP1. Segundo Breslow (2003), este algoritmo produz estimativas aproximadas das soluções das equações de máxima verossimilhança restrita (EMVR) para os componentes da variância e dos coeficientes de regressão no caso linear Gaussiano, para os outros modelos este algoritmo produz estimativas aproximadas das soluções das equações de máxima verossimilhança (EMV).

- **Algoritmo da QVP de segunda ordem (QVP2)**

Este algoritmo foi proposto por Goldstein e Rasbash (1996) e envolve uma aproximação da série de Taylor de 2ª ordem em termos unicamente do efeito aleatório $\mathbf{u}_{j}$.

Primeiro passo: Expandir h sobre as estimativas atuais de $(\hat{\beta}, \tilde{\mathbf{u}}_j)$. A partir da equação (3.5) pode-se escrever a aproximação da série de Taylor de $2^{\underline{a}}$ ordem da função em torno dos parâmetros estimados da seguinte forma

$$\begin{aligned} f(\beta, \mathbf{u}_j) &\approx f(\hat{\beta}, \tilde{\mathbf{u}}_j) + \left.\frac{\partial f(\beta, \mathbf{u}_j)}{\partial \beta}\right|_{(\hat{\beta}, \tilde{\mathbf{u}}_j)} (\beta - \hat{\beta}) + \left.\frac{\partial f(\beta, \mathbf{u}_j)}{\partial \mathbf{u}_j}\right|_{(\hat{\beta}, \tilde{\mathbf{u}}_j)} (\mathbf{u}_j - \tilde{\mathbf{u}}_j) \\ &\quad + \frac{1}{2} \left.\frac{\partial^2 f(\beta, \mathbf{u}_j)}{\partial \mathbf{u}_j \partial \mathbf{u}_j^T}\right|_{(\hat{\beta}, \tilde{\mathbf{u}}_j)} (\mathbf{u}_j - \tilde{\mathbf{u}}_j)(\mathbf{u}_j - \tilde{\mathbf{u}}_j)^T \\ &\approx \hat{\mu}_{ij}^{\mathbf{u}} + h^{'}(\mathbf{x}_{ij}^T \hat{\beta} + \mathbf{z}_{ij}^T \tilde{\mathbf{u}}_j)\, \mathbf{x}_{ij}^T (\beta - \hat{\beta}) + h^{'}(\mathbf{x}_{ij}^T \hat{\beta} + \mathbf{z}_{ij}^T \tilde{\mathbf{u}}_j)\, \mathbf{z}_{ij}^T (\mathbf{u}_j - \tilde{\mathbf{u}}_j) \\ &\quad + \frac{1}{2} h^{''}(\mathbf{x}_{ij}^T \hat{\beta} + \mathbf{z}_{ij}^T \tilde{\mathbf{u}}_j)\, \mathbf{z}_{ij}^T (\mathbf{u}_j - \tilde{\mathbf{u}}_j)^T (\mathbf{u}_j - \tilde{\mathbf{u}}_j)\, \mathbf{z}_{ij} \\ &\approx \hat{\mu}_{ij}^{\mathbf{u}} + h^{'}(\hat{\eta}_{ij}^{\mathbf{u}}) \left[\mathbf{x}_{ij}^T (\beta - \hat{\beta}) + \mathbf{z}_{ij}^T (\mathbf{u}_j - \tilde{\mathbf{u}}_j)\right] + \frac{1}{2} h^{''}(\mathbf{x}_{ij}^T \hat{\beta} + \mathbf{z}_{ij}^T \tilde{\mathbf{u}}_j)\, \mathbf{z}_{ij}^2 (\mathbf{u}_j - \tilde{\mathbf{u}}_j)^2. \end{aligned} \quad (3.11)$$

Goldstein e Rasbash (1996), sugerem ignorar o produto cruzado envolvendo diferentes componentes de $\mathbf{u_j}$, adicionando o termo quadrático como um *offsets*, isto é, um termo conhecido que está fora do processo de estimação.

Substituindo (3.11) em (3.4) obtemos

$$y_{ij} \approx \hat{\mu}_{ij}^{\mathbf{u}} + h^{'}(\hat{\eta}_{ij}^{\mathbf{u}}) \left[\mathbf{x}_{ij}^T (\beta - \hat{\beta}) + \mathbf{z}_{ij}^T (\mathbf{u}_j - \tilde{\mathbf{u}}_j)\right] + \frac{1}{2} h^{''}(\mathbf{x}_{ij}^T \beta + \mathbf{z}_{ij}^T \mathbf{u}_j)\, \mathbf{z}_{ij}^2 (\mathbf{u}_j - \tilde{\mathbf{u}}_j)^2 + e_{ij}. \quad (3.12)$$

Segundo passo: Aplicar a função g nos dados y_{ij}, e expandi-la em série de Taylor de $2^{\underline{a}}$ ordem em volta do valor estimado $\hat{\mu}_{ij}^{\mathbf{u}}$.

A aproximação da série de Taylor de $2^{\underline{a}}$ ordem da função em torno do valor estimado é construída da seguinte maneira

$$\begin{aligned} g(y_{ij}) &\approx g(\hat{\mu}_{ij}^{\mathbf{u}}) + \left.\frac{\partial g(y_{ij})}{\partial y_{ij}}\right|_{(\hat{\mu}_{ij}^{\mathbf{u}})} (y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}) + \frac{1}{2} \left.\frac{\partial^2 f(y_{ij})}{\partial y_{ij} \partial y_{ij}^T}\right|_{(\hat{\mu}_{ij}^{\mathbf{u}})} (y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})^T \\ &\approx g(\hat{\mu}_{ij}^{\mathbf{u}}) + g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}) + \frac{1}{2} g^{''}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})^2 \\ &\approx \hat{\eta}_{ij}^{\mathbf{u}} + g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}) + \frac{1}{2} g^{''}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})^2, \end{aligned} \quad (3.13)$$

então, podemos escrever (3.13) da seguinte forma

$$y_{ij}^* = \hat{\eta}_{ij}^{\mathbf{u}} + g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}) - \frac{1}{2} g^{''}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})^2.$$

Terceiro passo: Associar o modelo linearizado ao modelo teórico com distribuição normal. Denotando $g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}}) - \frac{1}{2} g^{''}(\hat{\mu}_{ij}^{\mathbf{u}})(y_{ij} - \hat{\mu}_{ij}^{\mathbf{u}})^2$ de e_{ij}^* e assumindo que y_{ij}^* é distribuído normalmente, o modelo misto linear para y_{ij}^* tem a mesma forma da equação (3.9) em que

$$\mathbf{u}_j \sim N(0, \mathbf{D}(\theta)) \text{ e } e_{ij}^* \sim N(0, [g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})]^2 a_j(\phi) v(\hat{\mu}_{ij}^{\mathbf{u}}) - \frac{1}{2} [g^{''}(\hat{\mu}_{ij}^{\mathbf{u}})]^2 a_j^2(\phi) v^2(\hat{\mu}_{ij}^{\mathbf{u}})), \quad (3.14)$$

como apresentado no Apêndice A.1.

As estimativas atualizadas de $(\beta, \mathbf{u}_j, \theta)$ são obtidas resolvendo as equações (3.9) e (3.14) simultaneamente, que agora leva em consideração a aproximação da série de Taylor de $2^{\underline{a}}$ ordem, isto é, utilizando o algoritmo QVP2. Este algoritmo também melhora as estimativas dos componentes da variância, mas Goldstein (1995) adverte que quando se trabalha com modelos complexos e/ou com conjuntos de dados pequenos surgem problemas de convergência. Em função disso, geralmente, utiliza-se o QVP1.

3.3 Modelos multiníveis para variável resposta binária

Os principais modelos para variáveis binárias são os modelos de regressão logística e o de regressão probit. Estes modelos são casos particulares dos MLGs como já foi mencionado.

Geralmente chama-se de "sucesso" o resultado mais importante da resposta ou aquele resultado que se pretende relacionar com as outras variáveis de interesse. A distribuição de Bernoulli para uma variável aleatória binária, por exemplo, Y_i de parâmetro μ_i especifica as probabilidades de tal forma que

$$\mathrm{P}(Y_i = 0) = 1 - \mu_i \quad \text{e} \quad \mathrm{P}(Y_i = 1) = \mu_i, \tag{3.15}$$

onde μ_i é a proporção de respostas em que $Y_i = 1$. A média e a variância de Y_i são dadas, respectivamente, por

$$\mathrm{E}(Y_i) = 1(\mu_i) + 0(1 - \mu_i) = \mu_i,$$

$$\mathrm{Var}(Y_i) = \mathrm{E}(Y_i^2) - \mathrm{E}^2(Y_i) = [1^2(\mu_i) + 0(1 - \mu_i)] - \mu_i^2 = \mu_i - \mu_i^2 = \mu_i(1 - \mu_i).$$

Seja y_{ij} o valor da variável resposta do i-ésimo indivíduo do j-ésimo grupo e, $\mathbf{u}_j \sim N(\mathbf{0}, \mathbf{D}(\theta))$, onde $g(.)$ é a função de ligação que pode ser logit, probit ou complemento log-log. O MMLG de 2 níveis para resposta binária é definido pela distribuição de probabilidades (3.15) e pelo preditor linear

$$g(\mu_{ij}) = \mathbf{x}_{ij}^T \beta + \mathbf{z}_{ij}^T \mathbf{u}_j, \tag{3.16}$$

com $\mu_{ij} = \mathrm{P}(y_{ij} = 1 \mid \mathbf{u}_j), i = 1, \ldots, n_j, j = 1, \ldots, J$. A especificação do modelo é completa assumindo que dado o vetor condicional de efeitos aleatórios, $\mathbf{u}_j$ as respostas binárias, y_{ij} são independentes. Asssim, a função de verossimilhança condicional tem a seguinte forma

$$L(\beta \mid \mathbf{u}_j) = \prod_{j=1}^{J} \prod_{i=1}^{n_j} f(y_{ij} \mid \mathbf{x}_{ij}, \mathbf{z}_{ij}, \beta, \mathbf{u}_j), \tag{3.17}$$

onde f é a função densidade de y_{ij}. A estratégia padrão para estimar os parâmetros do modelo (β, θ) é integrar a verossimilhança condicional em relação ao efeito aleatório $\mathbf{u}_j$,

$$L(\beta, \mathbf{D}(\theta)) = \int_{\mathbf{u}_j} L(\beta \mid \mathbf{u}_j)\, \phi(\mathbf{u}_j)\, d\mathbf{u}_j, \tag{3.18}$$

onde $\phi(.)$ representa a densidade da normal multivariada. A função de verossimilhança (3.18) pode ser resolvida de várias formas aproximadas, por exemplo, como as soluções propostas por Goldstein (1991), Breslow e Clayton (1993), Goldstein e Rasbash (1996) e Longford (1987, 1994), que utilizam expansões de série de Taylor de $1^{\underline{a}}$ ou $2^{\underline{a}}$ ordem. Será considerada a proposta de Breslow (2003), a QVP1, apresentada na seção anterior.

3.3.1 Modelo logístico multinível

Esta seção apresenta um caso particular do MMLG de 2 níveis, apresentado na seção 3.3, considerando a função de ligação logit, uma única variável explicativa associada ao efeito fixo β e uma variável explicativa associada ao efeito aleatório u_j.

Seja y_{ij} as respostas binárias para o i-ésimo indivíduo do nível 1 do j-ésimo grupo do nível 2, x_{ij} uma variável explicativa do nível 1 associada ao efeito fixo β_1 e $z_{ij} = 1$ uma variável explicativa associada ao efeito aleatório u_{0j}. Assumindo que os elementos y_{ij} são variáveis aleatórias independentes com distribuição Bernoulli (μ_{ij}), sendo (μ_{ij}) modelado pela função de ligação logit, ou seja,

$$\text{logit}(\mu_{ij}) = \beta_0 + \beta_1 x_{ij} + u_{0j}, \tag{3.19}$$

em que

$$\mu_{ij} = \text{P}(y_{ij} = 1) = \frac{\exp(\beta_0 + \beta_1 x_{ij} + u_{0j})}{1 + \exp(\beta_0 + \beta_1 x_{ij} + u_{0j})},$$

onde u_{0j} é o efeito aleatório do nível 2. Note que o componente e_{ij} não está na equação do modelo, porque ele faz parte da especificação do MLG. Note também que os coeficientes β não dependem do grupo j, pois este é um modelo combinado. A equação (3.19) é o preditor linear (3.16) com $\mathbf{x}_{ij}^T = (1, x_{ij})$, $\beta = (\beta_0, \beta_1)^T$, $\mathbf{z}_{ij}^T = 1$ e $g(.) = \text{logit}(.)$.

O modelo (3.19) pode ser escrito, alternativamente como

$$\text{logit}(\mu_{ij}) = \beta_{0j} + \beta_{1j} x_{ij}, \qquad \text{(nível 1)} \tag{3.20}$$

onde

$$\beta_{0j} = \gamma_{00} + u_{0j} \ \text{ e } \ \beta_{1j} = \gamma_{10}\ , \qquad \text{(nível 2)} \tag{3.21}$$

em que $u_{0j} \sim N(0, \sigma_{u0}^2)$. Substituindo a equação (3.21) em (3.20), tem-se (3.19).

Assumindo que dado o vetor condicional de efeitos aleatórios u_{0j} as respostas y_{ij} são independentes. Assim, a função de verossimilhança condicional para o modelo logístico (3.19), é dada por

$$L(\beta \mid u_{0j}) = \prod_{j=1}^{J} \prod_{i=1}^{n_j} \mu_{ij}^{y_{ij}} (1-\mu_{ij})^{1-y_{ij}}, \tag{3.22}$$

A verossimilhança para ajustar o modelo logístico (3.19) para J grupos independentes cada um com n_j subunidades é dada por

$$L(\beta, u_{0j}) = \prod_{j=1}^{J} \int \prod_{i=1}^{n_j} \mu_{ij}^{y_{ij}} (1-\mu_{ij})^{1-y_{ij}} \, \phi(u_{0j}) \, du_{0j}, \tag{3.23}$$

onde $\phi(.)$ representa a função densidade da normal.

O modelo multinível para resposta binária também pode ser obtido de um modelo linear multinível utilizando a chamada conceitualização de uma variável oculta (latente). Suponha que existe uma variável contínua oculta, $\tilde{y}_{ij}$ subjacente a y_{ij}. Supondo que observamos somente a variável resposta binária y_{ij} em lugar de $\tilde{y}_{ij}$, temos que a resposta binária, y_{ij}, é obtida pela dicotomização da variável contínua, $\tilde{y}_{ij}$, sendo $(y_{ij} = 1)$ se $\tilde{y}_{ij} > 0$ e $(y_{ij} = 0)$ se $\tilde{y}_{ij} \leq 0$. Assumindo que o vetor oculto $\tilde{y}_{ij}$ é distribuído normalmente, o modelo de efeito misto para a variável oculta pode ser escrito com a mesma estrutura do modelo (2.9)

$$\tilde{y}_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + \tilde{e}_{ij}, \tag{3.24}$$

onde, $\tilde{e}_{ij} \sim N(0, \sigma^2)$. Segundo Rodríguez e Goldman (1995), com esta suposição a distribuição condicional de y_{ij} é o produto de Bernoulli com probabilidade μ_{ij} satisfazendo a equação (3.19), isto é, o modelo obtido para a variável resposta y_{ij} é exatamente (3.19). Se assumirmos que o vetor oculto $\tilde{y}_{ij}$ não tem distribuição normal, o modelo de efeito misto para a variável oculta pode ser escrito como a estrutura do modelo (3.4)

$$\tilde{y}_{ij} = \text{logit}^{-1}(\beta_0 + \beta_1 x_{ij} + u_{0j}) + \tilde{e}_{ij},$$

onde $\text{Var}(\tilde{e}_{ij}) = a_j(\phi) v(\mu_{ij})$. Portanto, para obter as estimativas dos parâmetros (β_0, β_1 e σ^2_{u0}) e avaliar a integral em (3.23) pode-se utilizar o método QVP1, que é dado pela linearização em (3.9) e (3.10), ou o método QVP2 que é dado pela linearização em (3.9) e (3.14).

3.3.2 Modelo probit multinível

Esta seção apresenta um caso particular do MMLG de 2 níveis, apresentado na seção 3.3, considerando a função de ligação probit, uma única variável explicativa associada ao efeito fixo β e uma variável explicativa associada ao efeito aleatório u_j.

Seja y_{ij} as respostas binárias para o i-ésimo indivíduo do nível 1 do j-ésimo grupo do nível 2, x_{ij} uma variável explicativa do nível 1 associada ao efeito fixo β_1 e $z_{ij} = 1$ uma variável explicativa associada ao efeito aleatório u_{0j}. Assim, como o modelo logístico os elementos de y_{ij} são variáveis aleatórias independentes com distribuição Bernoulli (μ_{ij}), sendo que (μ_{ij}) agora é modelado pela função de ligação probit, ou seja,

$$\text{probit}(\mu_{ij}) = \beta_0 + \beta_1 x_{ij} + u_{0j}. \tag{3.25}$$

em que

$$\mu_{ij} = \text{P}(y_{ij} = 1 \mid u_{0j}) = \Phi(\beta_0 + \beta_1 x_{ij} + u_{0j}),$$

onde $\Phi(.)$ representa a densidade normal padrão acumulada e μ_{0j} é o efeito aleatório do nível 2. Note que o componente e_{ij}, assim como no modelo logístico, não está na equação do modelo, porque ele faz parte da especificação do MLG. Note também que os coeficientes β não dependem do grupo j, pois este é um modelo combinado.
A equação (3.25) é o preditor linear (3.16) com $\mathbf{x}_{ij}^T = (1, x_{ij})$, $\beta = (\beta_0, \beta_1)^T$, $\mathbf{z}_{ij}^T = 1$ e $g(.) = \text{probit}(.)$.

O modelo (3.25) pode ser escrito alternativamente como

$$\text{probit}(\mu_{ij}) = \beta_{0j} + \beta_{1j} x_{ij}, \tag{3.26}$$

onde β_{0j} e β_{1j} é dado pela equação (3.21) em que $u_{0j} \sim N(0, \sigma_{u0}^2)$. Substituindo a equação (3.21) em (3.26), tem-se (3.25).

As funções de verossimilhança condicional e não condicional para o modelo probit geral que também vale para o caso particular (3.25) têm as mesmas formas das equações (3.22) e (3.23), respectivamente.

O modelo (3.25) também pode ser obtido de maneira similar ao modelo logístico. Suponha que existe uma variável contínua oculta, $\tilde{y}_{ij}$ subjacente a y_{ij}. Supondo que observamos somente a variável resposta binária y_{ij} em lugar de $\tilde{y}_{ij}$, temos que a resposta binária, y_{ij}, é obtida pela dicotomização da variável contínua, $\tilde{y}_{ij}$, sendo $(y_{ij} = 1)$ se $\tilde{y}_{ij} > 0$ e $(y_{ij} = 0)$ se $\tilde{y}_{ij} \leq 0$. Assumindo que o vetor oculto $\tilde{y}_{ij}$ é distribuído normalmente o modelo de efeito misto para a variável oculta tem a mesma forma do modelo (3.24) satisfazendo a equação (3.25), portanto é similar ao modelo logístico anterior. Se assumirmos que o vetor oculto $\tilde{y}_{ij}$

não tem distribuição normal o modelo de efeito misto para a variável oculta pode ser escrito como a estrutura do modelo (3.4)

$$\tilde{y}_{ij} = \text{probit}^{-1}(\beta_0 + \beta_1 x_{ij} + u_{0j}) + \tilde{e}_{ij}$$

onde $\text{Var}(\tilde{e}_{ij}) = a_j(\phi)v(\mu_{ij})$. Portanto, para obter as estimativas dos parâmetros (β_0, β_1 e σ^2_{u0}) e avaliar a integral em (3.23) pode-se utilizar o método QVP1 ou QVP2.

4 Modelos para resposta binária má classificada

4.1 Considerações iniciais

Os métodos para análise de agrupamento e dados binários assumem que a resposta binária é medida sem erro, mas na prática isto, geralmente, não ocorre. Por exemplo, quando a resposta indica a presença ou ausência de uma condição médica, ela pode ser identificada por um teste de diagnóstico que em geral é imperfeito. A resposta binária má classificada pode surgir também, por exemplo, em estudos em que os investigadores (assistentes treinados por médicos) fazem exames físicos de cada indivíduo estudado e apresentam o diagnóstico de uma certa condição médica, que pode resultar em um diagnóstico impreciso ou em doenças com difícil diagnóstico.

Neuhaus (1999), investigou o efeito do erro para resposta em regressão binária com uma única observação por sujeito e mostrou que ignorar os erros na resposta pode levar a perda substancial de informações sobre os efeitos dos parâmetros de interesse e que a própria análise da resposta com o erro corrigido pode levar ainda a perdas de eficiência significativa com relação a verdadeira resposta sem erro. Além disso, o referido autor também mostrou que a classe de modelos lineares generalizados apresenta uma propriedade para a resposta má classificada que permite desenvolver um método simples para obter estimativas consistentes dos parâmetros de interesse com uma modificação direta no método do MLG. A idéia deste autor, resume-se em considerar o caso onde a variável resposta verdadeira (resposta sujeita a erros) segue um MLG com função de ligação logit, por exemplo, e ajustar o modelo com a mesma função de ligação para a resposta observada (resposta com o erro corrigido), através de uma expressão geral que avalia a magnitude do viés quando ignoramos os erros na variável resposta.

Neuhaus (2002), fez uma extensão de seu trabalho de 1999, examinando o efeito da má classificação da resposta para dados binários longitudinais utilizando dois procedimentos para estimar os parâmetros de interesse consistentemente: o procedimento de média populacional (segue um MLG) e o procedimento de grupo-específico (segue um MMLG). Este autor, considera que a resposta num tempo particular ou momento particular não depende das respostas anteriores do sujeito, isto significa assumir um modelo de independência entre as observações do mesmo sujeito logo, na abordagem de grupo-específico em que as observações do indivíduo constituem as observações do grupo tem-se que as observações do grupo não

são correlacionadas. Isto deixa claro que o modelo multinível para dados binários não foi tratado em Neuhaus (2002), no qual o interesse principal é considerar as correlações entre as observações do mesmo grupo.

No presente trabalho o interesse está centrado no procedimento de grupo-específico o qual nos inspirou a considerar um modelo hierárquico. Para tanto, se propõem procedimentos para estimar os parâmetros de interesse consistentemente e, examina-se o efeito da má classificação da resposta nos parâmetros do MMLG com resposta binária com funções de ligação logit e probit. Na literatura disponível não encontramos nenhum trabalho com esta abordagem, sendo portanto nossa contribuição original.

4.2 Resposta má classificada e o procedimento grupo-específico

- **Uma variável explicativa**

Sejam t_{ij} as respostas binárias *verdadeiras* para o i-ésimo indivíduo do nível 1 no j-ésimo grupo do nível 2 e x_{ij} uma variável explicativa do nível 1 onde $i = 1, \ldots, n_j$ e $j = 1, \ldots J$. Se ajustará o modelo de regressão binária para avaliar a associação de x_{ij} com $\mathrm{E}(t_{ij})$ bem como se estimará a dependência dentro do grupo da resposta t_{ij}.

Seja u_{0j} o efeito aleatório do nível 2, a probabilidade de $t_{ij} = 1$ (ocorrer uma resposta verdadeira) é especificada como

$$\mu_{ij} = \mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j}) = g^{-1}(\beta_0 + \beta_1 x_{ij} + u_{0j}) = \mathrm{E}(t_{ij} = 1 \mid x_{ij}, u_{0j}), \tag{4.1}$$

em que g^{-1} é a inversa da função de ligação . As seguintes funções de ligação são consideradas:

$$g(\mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j})) = \operatorname{logit} \mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j}) = \beta_0 + \beta_1 x_{ij} + u_{0j}, \tag{4.2}$$

e

$$g(\mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j})) = \operatorname{probit} \mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j}) = \beta_0 + \beta_1 x_{ij} + u_{0j}. \tag{4.3}$$

A função de verossimilhança condicional têm a seguinte estrutura

$$L(\beta \mid u_{0j}) = \prod_{j=1}^{J} \prod_{i=1}^{n_j} \mu_{ij}^{t_{ij}} (1 - \mu_{ij})^{1 - t_{ij}}. \tag{4.4}$$

A especificação do modelo é completa com a suposição de que $u_{0j} \sim N(0, \sigma^2_{u0})$. A verossimilhança não condicional da equação (4.4), tem a mesma estrutura da equação (3.22) e pode ser resolvida de várias formas aproximadas como comentado anteriormente, mas utilizaremos o procedimento da QVP1.

O parâmetro β_1 mede a mudança na esperança condicional dentro do j-ésimo grupo correspondendo a um aumento de uma unidade na covariável x_{ij} do primeiro nível, isto é,

$$\beta_1 = g[\mathrm{P}(t_{ij} = 1 \mid x_{ij} + 1, u_{0j})] - g[\mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j})]. \tag{4.5}$$

Nota-se que esta mudança é suposta igual em todos os grupos. No trabalho de Neuhaus (2002), os índices têm interpretações diferentes: o índice i representa o grupo (indivíduo) e o índice j as unidades dentro do grupo, em que, para cada grupo i têm-se informações de um mesmo indivíduo em diferentes tempos k. Portanto, este autor trabalha com medida repetida no tempo. Neste estudo, temos uma única observação por sujeito e os índices i e j são interpretados como indivíduo e grupo, respectivamente, como citado anteriormente.

No procedimento grupo-específico para resposta binária má classificada não se observa t_{ij}, mas a sua versão com erro y_{ij}. Na teoria, a probabilidade da má classificação pode depender de todas as respostas verdadeiras e das covariáveis, bem como todas as outras respostas observadas. Neuhaus (2002), assume que a probabilidade da má classificação y_{ij} depende somente da resposta verdadeira e possivelmente da covariável, pois segundo este autor, esta suposição é razoável quando a resposta surge de um teste de diagnóstico ou procedimento em que o resultado independe da classificação dos demais sujeitos. Portanto, este autor assume que a classificação de um sujeito num determinado momento não depende da classificação em outro momento, como comentado anteriormente. No presente estudo, tem-se que as classificações dentro de um mesmo grupo estão relacionadas.

Sejam $\tau_1(x_{ij})$ e $\tau_0(x_{ij})$ denotando as probabilidades da resposta ser má classificada em que

$$\tau_1(x_{ij}) = \mathrm{P}(y_{ij} = 0 \mid t_{ij} = 1, x_{ij}), \qquad \text{(falso negativo)} \tag{4.6}$$

e

$$\tau_0(x_{ij}) = \mathrm{P}(y_{ij} = 1 \mid t_{ij} = 0, x_{ij}). \qquad \text{(falso positivo)} \tag{4.7}$$

Associando estas definições a linguagem utilizada para testes diagnósticos, tem-se que a probabilidade $1 - \tau_1(x_{ij})$ é a chamada sensibilidade da medida y_{ij}, isto é, a probabilidade de diagnosticar corretamente a pessoa doente. Enquanto que $1 - \tau_0(x_{ij})$ é a chamada especificidade, isto é, a probabilidade de diagnosticar corretamente os sadios. O modelo *verdadeiro* para a resposta observada y_{ij} tem a seguinte forma

$$\begin{aligned}
\mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, u_{0j}) &= \mathrm{P}_V(y_{ij} = 1 \mid t_{ij} = 0, x_{ij}, u_{0j})\mathrm{P}(t_{ij} = 0 \mid x_{ij}, u_{0j}) + \\
&\quad \mathrm{P}_V(y_{ij} = 1 \mid t_{ij} = 1, x_{ij}, u_{0j})\mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j}) \\
&= \tau_0(x_{ij})\mathrm{P}(t_{ij} = 0 \mid x_{ij}, u_{0j}) + [1 - \tau_1(x_{ij})]\mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j}) \\
&= \tau_0(x_{ij})[1 - \mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j})] + [1 - \tau_1(x_{ij})]\mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j})
\end{aligned}$$

$$= \{1 - \tau_1(x_{ij}) - \tau_0(x_{ij})\}\, \mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j}) + \tau_0(x_{ij}), \tag{4.8}$$

onde P_V denota a probabilidade sob o modelo *verdadeiro* para y_{ij}.

Assume-se que a probabilidade da má classificação não depende da covariável e nem do efeito aleatório, pois esta suposição é razoável quando a resposta surge de um teste de diagnóstico, como comentado em Neuhaus (2002). Suponha que t_{ij} segue um MMLG com função de ligação g. Considerando a equação (4.1) pode-se reescrever a equação (4.8) da seguinte forma

$$\mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, u_{0j}) = \{1 - \tau_1 - \tau_0\} g^{-1}(\beta_0 + \beta_1 x_{ij} + u_{0j}) + \tau_0, \tag{4.9}$$

assume-se que $\tau_1 + \tau_0 < 1$, pois valores de τ_1 e τ_0 maiores do que 0.5 indicam que o procedimento que produz a classificação observada y_{ij} apresenta um desempenho ruim.

Quando t_{ij} segue um MMLG com função de ligação g e com distribuição dos efeitos aleatórios normal tem-se que y_{ij} segue um MMLG com função de ligação modificada, g^*, em que

$$\eta = g\left\{\frac{\mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, u_{0j}) - \tau_0}{1 - \tau_1 - \tau_0}\right\} = g^*\{\mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, u_{0j})\} = g^*(\mu_{ij}), \tag{4.10}$$

onde $\mu_{ij} = \mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, u_{0j})$. Assim, a verossimilhança para a resposta observada y_{ij} terá a mesma forma da integral da verossimilhança para a resposta verdadeira t_{ij}, mas irá depender da função de ligação g^* ao invés de g.

A maximização da verossimilhança para a resposta observada pode ser estimada pela derivada da função de ligação g^* fazendo algumas modificações simples na derivada de g ou aplicando o algoritmo QVP1 ou QVP2 que utiliza a função de ligação especificada. Neste estudo utilizaremos o algoritmo QVP1 que estima consistentemente todos os parâmetros do modelo analisando a resposta observada y_{ij}.

4.3 O viés devido a resposta má classificada ignorada

- **Uma variável explicativa**

Para estudar o viés devido aos erros ignorados na resposta binária admitimos, assim como em Neuhaus (1999), que ajusta-se o modelo para os dados observados, y_{ij}, com função de ligação g e não g^*. Em particular, assume-se que o verdadeiro modelo que relaciona a variável explicativa x_{ij} e o efeito aleatório u_{0j} com a verdadeira resposta t_{ij} tem a seguinte forma

$$g\{\mathrm{P}(t_{ij} = 1 \mid x_{ij}, u_{0j})\} = \beta_0 + \beta_1 x_{ij} + u_{0j}, \tag{4.11}$$

mas, quando se ajusta o modelo para y_{ij} considerando função de ligação g, neste caso, está ajustando-se um modelo mal especificado dado por

$$g\{\mathrm{P_F}(y_{ij}=1 \mid x_{ij}, u_{0j})\} = \beta_0^* + \beta_1^* x_{ij} + u_{0j}^*, \tag{4.12}$$

onde $u_{0j}^* \sim N(0,\, \sigma_{u0}^{2*})$ e $\mathrm{P_F}$ denota a probabilidade sob o modelo *falso* para y_{ij}. Assim, o objetivo da análise é comparar os β^*'s com os β's e σ_{u0}^{2*} com σ_{u0}^2.

Os trabalhos de Huber (1967), Akaike (1973) e White (1982) para modelos mal especificados, mostraram que as estimativas $(\hat{\beta}_0^*, \hat{\beta}_1^*, \hat{\sigma}_{u0}^*)$ do modelo *falso* ($\mathrm{P_F}$) convergem para valores $(\beta_0^*, \beta_1^*, \sigma_{u0}^*)$ que minimizam a divergência de Kullback Leibler (Kullback, 1959) entre o modelo *verdadeiro* (4.9) e o modelo *falso* (4.12). Neuhaus (1999), utilizou os resultados destes trabalhos e provou que a minimização desta divergência leva a um sistema de equações de difícil solução, e então encontrou uma solução aproximada para a minimização do critério de divergência de Kullback Leibler. Seguindo a proposta deste autor, obteve-se a seguinte solução aproximada para o caso dos MMLG's

$$\mathrm{P_V}(y_{ij}=1 \mid x_{ij}) \approx \mathrm{P_F}(y_{ij}=1 \mid x_{ij}) = g^{-1}(\beta_0^* + \beta_1^* x_{ij} + u_{0j}^*), \tag{4.13}$$

$\forall\, x_{ij}$. Esta aproximação é razoável uma vez que

$$\mathrm{P_V}(u_{0j}) = \int \mathrm{P_V}(y_{ij}=1 \mid x_{ij}, u_{0j})\, \phi(u_{0j})\, du_{0j},$$

e

$$\mathrm{P_F}(u_{0j}) = \int \mathrm{P_F}(y_{ij}=1 \mid x_{ij}, u_{0j})\, \phi(u_{0j})\, du_{0j},$$

onde $\phi(.)$ representa a função densidade da normal, $\mathrm{P_V}(u_{0j})$ e $\mathrm{P_F}(u_{0j})$ são as integrais indefinidas de $\mathrm{P_V}(y_{ij}=1 \mid x_{ij})$ e $\mathrm{P_F}(y_{ij}=1 \mid x_{ij})$, respectivamente, logo $\mathrm{P_V}(u_{0j}) \simeq \mathrm{P_F}(u_{0j})$, $\forall u_{0j} \in \{x_{ij}\} \forall i \forall j$ se e somente se $\mathrm{P_V}(y_{ij}=1 \mid x_{ij}) \simeq \mathrm{P_F}(y_{ij}=1 \mid x_{ij})$, $\forall x_{ij}$. Assim, as estimativas obtidas da probabilidade do modelo *falso* (4.12) são aproximadamente iguais as estimativas obtidas a partir da probabilidade *verdadeira* (4.9). Se $x_{ij}=0$, pode-se obter β_0^* em termos da $\mathrm{P_V}(y_{ij}=1 \mid x_{ij}, u_{0j})$ a partir de (4.13), neste caso tem-se

$$\beta_0^* = g\{\mathrm{P_F}(y_{ij}=1 \mid x_{ij}=0, u_{0j})\} = g\{\mathrm{P_V}(y_{ij}=1 \mid x_{ij}=0, u_{0j})\}.$$

Para todos os outros valores de x_{ij} temos que a medida do efeito da variável explicativa pelo modelo *falso* é

$$\beta_1^* = g\{\mathrm{P_F}(y_{ij}=1 \mid x_{ij}+1, u_{0j})\} - g\{\mathrm{P_F}(y_{ij}=1 \mid x_{ij}, u_{0j})\}. \tag{4.14}$$

Substituindo a probabilidade do modelo *falso* ($\mathrm{P_F}$) em (4.14) pela probabilidade *verdadeira* ($\mathrm{P_V}$) de (4.8), podemos relacionar os parâmetros estimados do modelo *falso*, β_1^*, com os

valores analógos ao do modelo *verdadeiro*, β_1, levando em consideração as probabilidades de má classificação τ_1 e τ_0. Assim,

$$\beta_1^* = g\{\mathrm{P_V}(y_{ij} = 1 \mid x_{ij} + 1, u_{0j})\} - g\{\mathrm{P_V}(y_{ij} = 1 \mid x_{ij}, u_{0j})\} = H(\beta_1). \tag{4.15}$$

Observe que $H(0) = 0$. Seguindo o procedimento proposto por Neuhaus (1999), pode-se aproximar esta relação expandindo (4.15) em série de Taylor sob $\beta_1 = 0$ e obter, como apresentado no Apêndice A.2, a seguinte expressão

$$\beta_1^* \approx \beta_1 H'(0) = \beta_1 \frac{(1 - \tau_1 - \tau_0) g'[(1 - \tau_1 - \tau_0) g^{-1}\{\beta_0 + u_{0j}\} + \tau_0]}{g'[g^{-1}\{\beta_0 + u_{0j}\}]}. \tag{4.16}$$

O fator $H'(0)$ corresponde a um fator de viés. A estimação desta constante é discutida por Li e Duan (1989) quando $u_{0j} = 0$. Quando não há erro na resposta, isto é, quando $(\tau_0 = \tau_1 = 0), H'(0) = 1$ e quando há erro na resposta ($\tau_0 \neq 0$ ou $\tau_1 \neq 0$), temos que $H'(0) \neq 1$. Para função de ligação com $1/g'$ côncava, como por exemplo, a logit, a probit e o complemento log-log, Neuhaus (1999), mostrou que $0 \leq H'(0) \leq 1$ quando $u_{0j} = 0$, ou seja, ignorar os erros na resposta conduzem a estimativas "atenuadas", e portanto menos eficientes do efeito da variável explicativa. Além disso, o referido autor, mostrou ainda que utilizar a derivada da função de ligação identidade também conduz a $0 \leq H'(0) \leq 1$. A extensão da prova é direta quando se considera o efeito u_{0j}.

A equação (4.16) indica que a magnitude do viés devido a má classificação da resposta depende de $(\beta_0 + u_{0j})$ e das duas probabilidades de má classificação τ_1 e τ_0 e, não do tamanho da amostra.

Ao se considerar a função de ligação logit, como demonstrado no Apêndice A.3, a equação (4.16) pode ser escrita como

$$\begin{aligned} \beta_1^* &\approx \beta_1 H'(0) \\ &\approx \beta_1 \frac{(1 - \tau_1 - \tau_0)\ g^{-1}\{\beta_0 + u_{0j}\}\ [1 - g^{-1}(\beta_0 + u_{0j})]}{[(1 - \tau_1 - \tau_0)\, g^{-1}\{\beta_0 + u_{0j}\} + \tau_0]\, [(1 - \tau_1 - \tau_0)\, (1 - g^{-1}\{\beta_0 + u_{0j}\}) + \tau_1]}. \end{aligned} \tag{4.17}$$

Se a função de ligação a ser considerada é a probit que a partir da equação (4.16), como provado no Apêndice A.4, pode ser escrita como

$$\begin{aligned} \beta_1^* &\approx \beta_1 H'(0) \\ &\approx \beta_1 \frac{(1 - \tau_1 - \tau_0)\ \phi\{\beta_0 + u_{0j}\}}{\phi[\Phi^{-1}((1 - \tau_1 - \tau_0)\, g^{-1}\{\beta_0 + u_{0j}\} + \tau_0)]}, \end{aligned} \tag{4.18}$$

onde $\phi[.]$ é a densidade Normal padronizada e $\Phi(.)$ é a distribuição Normal padrão acumulada.

- **Mais de uma variável explicativa**

A extensão de MMLG para as respostas binárias má classificadas com múltiplos preditores é simples. Esta seção apresenta a extensão das principais expressões das equações apresentadas nas seções 4.2 e 4.3. Considerando t_{ij} as respostas binárias verdadeiras para o i-ésimo indivíduo do nível 1 no j-ésimo grupo do nível 2, $\mathbf{x}_{ij}$ o vetor de observações das variáveis regressoras $(p+1)$-dimensional associado ao efeito fixo, β e $\mathbf{z}_{ij}$ o vetor de observações das variáveis regressoras $(q+1)$-dimensional associado ao efeito aleatório, $\mathbf{u}_j$, em que $\mathbf{u}_j \sim N(\mathbf{0}, \mathbf{D}(\theta))$, o MMLG para a variável resposta t_{ij} pode ser escrito como

$$g(\mathrm{P}(t_{ij}=1 \mid x_{1ij},\ldots,x_{pij};u_{0j},\ldots,u_{qj})) = \beta_0+\beta_1 x_{1ij}+\ldots+\beta_p x_{pij} \\ +u_{0j}+u_{1j}z_{1ij}+\ldots+u_{qj}z_{qij}.$$

A extensão das principais expressões são apresentadas abaixo:

- Extensão da expressão (4.1)

$$\mathrm{P}(t_{ij}=1 \mid x_{1ij},\ldots,x_{pij};u_{0j},\ldots,u_{qj}) = g^{-1}(\beta_0+\beta_1 x_{1ij}+\ldots+\beta_p x_{pij} \\ +u_{0j}+u_{1j}z_{1ij}+\ldots+u_{qj}z_{qij}), \qquad (4.19)$$

onde $\mathrm{P}(t_{ij}=1 \mid x_{1ij},\ldots,x_{pij};u_{0j},\ldots,u_{qj})$ é a probabilidade de ocorrer a resposta verdadeira, como comentado anteriormente.

- Extensão da expressão (4.2)

$$\beta_k = g[\mathrm{P}(t_{ij}=1 \mid x_{1ij}\ldots,x_{(k-1)ij},x_{kij}+1,x_{(k+1)ij},\ldots,x_{pij};\, u_{0j},\ldots,u_{qj})] \\ -g[\mathrm{P}(t_{ij}=1 \mid x_{1ij}\ldots,x_{(k-1)ij},x_{kij},x_{(k+1)ij},\ldots,x_{pij};\, u_{0j},\ldots,u_{qj})],$$

onde os x_{kij}'s são independentes e $k=1,\ldots,p$. Assim, a extensão da expressão (4.2) pode ser escrita como:

$$\beta_k = g[\mathrm{P}(t_{ij}=1 \mid (x_{kij}+1);\, u_{0j},\ldots,u_{qj})] \\ -g[\mathrm{P}(t_{ij}=1 \mid x_{kij};\, u_{0j},\ldots,u_{qj})], \qquad (4.20)$$

em que β_k mede a mudança na esperança condicional no j-ésimo grupo correspondendo a um aumento de uma unidade na k-ésima covariável.

- Extensão da expressão (4.9)

$$\mathrm{P}_V(y_{ij}=1 \mid x_{1ij},\ldots,x_{pij};u_{0j},\ldots,u_{qj}) = \{1-\tau_1-\tau_0\}g^{-1}(\beta_0+\beta_1x_{1ij}+\ldots+\beta_px_{pij} \\ +u_{0j}+u_{1j}z_{1ij}+\ldots+u_{qj}z_{qij})+\tau_0, \qquad (4.21)$$

onde P_V denota a probabilidade sob o modelo *verdadeiro* para y_{ij} com a suposição de que o erro da má classificação não depende das covariáveis.

- Extensão da expressão (4.10)

$$\eta = g\left\{\frac{\mathrm{P}_V(y_{ij}=1 \mid x_{1ij},\ldots,x_{pij};u_{0j},\ldots,u_{qj})-\tau_0}{1-\tau_1-\tau_0}\right\} \\ = g^*\{\mathrm{P}_V(y_{ij}=1 \mid x_{1ij},\ldots,x_{pij};u_{0j},\ldots,u_{qj})\} = g^*(\mu_{ij}), \qquad (4.22)$$

em que g^* é a função de ligação modificada para a variável observada, y_{ij}.

- Extensão da expressão (4.12)

$$g\{\mathrm{P}_F(y_{ij}=1 \mid x_{1ij},\ldots,x_{pij};u_{0j},\ldots,u_{qj})\} = \beta_0^*+\beta_1^*x_{1ij}+\ldots+\beta_p^*x_{pij} \\ +u_{0j}^*+u_{1j}^*z_{1ij}+\ldots+u_{qj}^*z_{qij}, \qquad (4.23)$$

onde P_F denota a probabilidade sob o modelo *falso* para y_{ij}.

- Considerando a equação (4.13)

$$P_V(y_{ij}=1 \mid x_{kij}) \approx P_F(y_{ij}=1 \mid x_{kij}),$$

temos

$$P_V(y_{ij}=1 \mid x_{kij},u_{0j},\ldots,u_{qj}) \simeq P_F(y_{ij}=1 \mid x_{kij},u_{0j},\ldots,u_{qj}). \qquad (4.24)$$

- Extensão da expressão (4.15)

$$\beta_k^* = g\{P_V(y_{ij}=1 \mid (x_{kij}+1);\, u_{0j},\ldots,u_{qj})\} \\ -g\{P_V(y_{ij}=1 \mid x_{kij};\, u_{0j},\ldots,u_{qj})\} = H(\beta_k). \qquad (4.25)$$

- Extensão da expressão (4.16)

$$\begin{aligned}\beta_k^* &\approx \beta_k H'(0) \\ &\approx \beta_k \frac{(1-\tau_1-\tau_0)g'[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}+\ldots+u_{qj}z_{qij}\}+\tau_0]}{g'[g^{-1}\{\beta_0+u_{0j}+\ldots+u_{qj}z_{qij}\}]},\end{aligned} \tag{4.26}$$

denota-se $\beta_0+u_{0j}+\ldots+u_{qj}z_{qij}$ por w para simplificar a notação das expressões apresentadas a seguir.

- Extensão da expressão (4.17)

$$\begin{aligned}\beta_k^* &\approx \beta_k H'(0) \\ &\approx \beta_k \frac{(1-\tau_1-\tau_0)\ \ g^{-1}\{w\}\ \ [1-g^{-1}(w)]}{[(1-\tau_1-\tau_0)\,g^{-1}\{w\}+\tau_0]\,[(1-\tau_1-\tau_0)\,(1-g^{-1}\{w\})+\tau_1]}.\end{aligned} \tag{4.27}$$

- Extensão da expressão (4.18)

$$\begin{aligned}\beta_k^* &\approx \beta_k H'(0) \\ &\approx \beta_k \frac{(1-\tau_1-\tau_0)\ \ \phi\{w\}}{\phi[\Phi^{-1}((1-\tau_1-\tau_0)\,g^{-1}\{w\}+\tau_0)]},\end{aligned} \tag{4.28}$$

onde $\phi[.]$ é a densidade Normal padronizada e $\Phi(.)$ é a densidade Normal padrão acumulada.

5 Simulações

5.1 Considerações iniciais

Neste capítulo são investigados os comportamentos dos estimadores baseados na metodologia apresentada no capítulo 4 de modelos para resposta binária má classificada. Considerou-se duas funções de ligação pertencentes a família binomial; logit e probit. Para cada função de ligação a investigação foi realizada considerando quatro estudos de simulações. O primeiro estudo investiga o comportamento dos vieses das estimativas dos parâmetros a partir das equações (4.2) e (4.3) apresentadas no capítulo 4 sob a má classificação da variável resposta binária, e a qualidade da aproximação das equações (4.17) e (4.18) nos modelos logístico e probit, respectivamente. O segundo estudo avalia o viés devido a má classificação da variável resposta binária com mais de uma variável explicativa. O terceiro estudo investiga o comportamento das estimativas das probabilidades de má classificação consideradas como parâmetros desconhecidos e das estimativas dos parâmetros a partir das equações (4.2) e (4.3) apresentadas no capítulo 4. O quarto estudo avalia o comportamento das estimativas das probabilidades de má classificação consideradas como parâmetros desconhecidos num modelo com mais de uma variável explicativa, conjuntamente com o desempenho das estimativas dos parâmetros regressores.

As investigações foram realizadas através de simulações de Monte Carlo. O número de réplicas de Monte Carlo utilizado em cada experimento foi de 1.000 e, cada amostra de Monte Carlo envolveu 200 grupos de tamanhos balanceados: 5, 10, 20, 50 e 100, respectivamente. Em cada estudo de simulação, 16 diferentes valores para o par (τ_0, τ_1) e cinco diferentes valores de tamanhos de amostra por grupo foram considerados, totalizando oitenta diferentes situações.

5.2 Primeiro estudo de simulação

Neste primeiro estudo de simulação são investigados os comportamentos dos vieses das estimativas dos parâmetros β_0, β_1 e σ_{u0} a partir das equações (4.2) e (4.3) discutidas no capítulo 4 sob a má classificação da variável resposta binária, e a qualidade das aproximações propostas pelas equações (4.17) e (4.18) para $\beta_1 \neq 0$ nos modelos logístico e probit, respectivamente.

O desempenho das estimativas de β_0, β_1 e σ_{u0} foi avaliado pelos valores da média e desvio padrão amostral, além do viés. As estimativas de máxima verossimilhança para os

parâmetros de interesse foram obtidas utilizando o algoritmo de Newton-Rapson e Quase-verossimilhança penalizada para avaliar a integral em (3.20) pelo procedimento `glmmPQL` do `R` que utiliza a chamada conceitualização de uma variável oculta.

Para ilustrar o procedimento dessa investigação, segue-se os seguintes passos:

1. Gera-se uma amostra de observações para a variável dependente binária t_{ij} conforme o modelo logístico hierárquico (4.2) com um intercepto aleatório u_{0j} que segue uma distribuição normal padrão, e uma variável explicativa, x_{ij} que também segue uma distribuição normal padrão. A variável independente x_{ij} é gerada independentemente dentro de cada grupo. Os valores verdadeiros dos parâmetros do modelo logístico hierárquico considerados são: $\beta_0 = -1$, $\beta_1 = 1$ e $\theta = \sigma^2_{u0} = 1$. Para efeito de comparação, ressalta-se que estes valores são iguais aos considerados no trabalho de Neuhaus (2002).

2. As estimativas dos parâmetros β_0, β_1 e σ_{u0} são obtidas resolvendo as equações (3.9) e (3.10) pelo procedimento `glmmPQL` do `R`.

3. Gera-se a resposta observada y_{ij} a partir da resposta verdadeira t_{ij}. Em que $y_{ij} = 0$ quando $t_{ij} = 1$ com probabilidade $1 - \tau_1$ (sensibilidade) e $y_{ij} = 1$ quando $t_{ij} = 0$ com probabilidade $1 - \tau_0$ (especificidade).

4. Ajusta-se o modelo logístico de efeito misto para a resposta observada y_{ij} resolvendo as equações (3.9) e (3.10) pelo procedimento `glmmPQL` do `R`.

O mesmo procedimento é feito para o modelo probit, sendo que a variável dependente binária t_{ij} é gerada conforme o modelo probit hierárquico (4.3).

Os Quadros 1a e 1b (págs. 60 e 61) apresentam os valores médios dos parâmetros estimados do modelo logístico hierárquico de efeito misto (4.2) obtidos com erros na resposta, o desvio padrão dessas estimativas, os valores preditos utilizando a aproximação (4.17) com base no tamanho de amostra $n = 5$, e a quantidade de convergência pelo procedimento `glmmPQL` (ver Apêndice B.1). Nos quadros referidos observa-se que quando se fixa τ_0 o viés, em módulo, da estimativa de β_1 aumenta quando τ_1 aumenta. Da mesma forma, quando se fixa τ_1 o viés, em módulo, da estimativa β_1 aumenta quando τ_0 aumenta. O mesmo comportamento é observado para a estimativa corrigida segundo a expressão (4.17). Isto indica que quanto maior for a probabilidade de má classifição menor será as estimativas obtidas dos parâmetros de interesse. Ao comparar os vieses em módulo, observa-se que o viés da estimativa corrigida é relativamente menor, para diferentes combinações de τ_0 e τ_1.

Estes resultados, indicam que a expressão (4.17) fornece uma estimativa mais próxima do verdadeiro valor para os parâmetros de interesse quando a resposta binária é má classificada, como era esperado. Analisando o comportamento da estimativa σ^2_{u0} observa-se que a estimativa do parâmetro σ^2_{u0} é próxima de zero, para diferentes valores de τ_0 e τ_1, indicando

que este parâmetro está sendo sub-estimado, pois o procedimento `glmmPQL` no `R` não ajusta a função de variância. Comparando os Quadros 1a e 1b com os valores apresentados no artigo de Neuhaus (2002), para o modelo logístico de grupo-específico para dados longitudinais pode-se observar que a qualidade das aproximações são similares.

Os Quadros 2a a 5b (págs. 62 a 69) apresentam os resultados para os tamanhos de amostras $n = 10, 20, 50$ e 100, respectivamente, nos quais observa-se o mesmo comportamento para as estimativas β_1 e σ^2_{u0}, para diferentes valores de τ_0 e τ_1 (ver Apêndice B.1), e como era esperado, a magnitude destes vieses não depende do tamanho da amostra.

Nos Quadros 6a a 10b (págs. 70 a 79), apresentados nos Apêndice B.1, encontram-se os resultados das simulações para as análises das estimativas dos parâmetros do modelo probit hierárquico (4.3) obtidos com erros na resposta, os desvios padrão dessas estimativas, os valores preditos utilizando a aproximação (4.18), para $n = 5, 10, 20, 50$ e 100, e a quantidade de convergência pelo procedimento `glmmPQL`. As estimativas obtidas para os parâmetros deste modelo tendem a estar mais afastadas dos verdadeiros valores, em relação aos parâmetros do modelo logístico (4.2). Estes resultados podem sugerir que a especificação da função de ligação pode ser uma importante consideração quando se ajusta modelos multiníveis para dados binários utilizando o algoritmo da Quase-verossimilhança Penalizada, como comentado em Renard (2002). De forma análoga ao que acontece com as estimativas dos parâmetros do modelo logístico hierárquico, observa-se que a magnitude dos vieses das estimativas de β_0, β_1 e σ^2_{u0} não depende do tamanho da amostra, como era esperado. Quando comparam-se os vieses, em módulo, das estimativa β_1, observa-se que o viés da estimativa corrigida é relativamente menor, indicando que esta expressão apresenta estimativa consistente para o parâmetro β_1. Analisando o comportamento da estimativa σ^2_{u0} observa-se que a média da estimativa do parâmetro σ^2_{u0} é muito próxima de zero, para diferentes valores de τ_0 e τ_1, indicando que este parâmetro está sendo sub-estimado, como era esperado.

5.3 Segundo estudo de simulação

Este segundo estudo de simulação tem por objetivo avaliar o viés devido a má classificação da variável resposta binária num modelo com mais de uma variável explicativa.

Os resultados numéricos apresentados a seguir são obtidos a partir de um modelo misto linear generalizado com a seguinte estrutura:

$$g(\mathrm{P}(t_{ij} = 1 \mid x_{1ij}, x_{2ij}, u_{0j}, u_{1j})) = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_{0j} + z_{1ij} u_{1j},$$

$i = 1, \ldots, n_j; j = 1, \ldots, J$, onde g é a função de ligação; x_{1ij} e x_{2ij} são as variáveis explicativas associadas aos efeitos fixos β_1 e β_2, repectivamente, z_{ij} é a variável explicativa associada ao efeito aleatório u_{1j}; u_{0j} e u_{1j} são variáveis aleatórias independentes com as seguintes

suposições : $u_{0j} \sim N(0, \sigma^2_{u0})$, $u_{1j} \sim N(0, \sigma^2_{u1})$ e $cov(u_{0j}, u_{1j}) = 0$. As seguintes funções de ligação são consideradas:

$$\text{logit}\, \mathrm{P}(t_{ij} = 1 \mid x_{1ij}, x_{2ij}, u_{0j}, u_{1j}) = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_{0j} + z_{1ij} u_{1j}, \tag{5.1}$$

e

$$\text{probit}\, \mathrm{P}(t_{ij} = 1 \mid x_{1ij}, x_{2ij}, u_{0j}, u_{1j}) = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_{0j} + z_{1ij} u_{1j}, \tag{5.2}$$

em que $z_{1ij} = x_{1ij}$, com suas respectivas equações aproximadas

$$\beta^*_k \approx$$

$$\approx \beta_k \frac{(1 - \tau_1 - \tau_0)\ g^{-1}\{\beta_0 + u_{0j} + u_{1j} z_{1ij}\}\ [1 - g^{-1}(\beta_0 + u_{0j} + u_{1j} z_{1ij})]}{[(1 - \tau_1 - \tau_0)\, g^{-1}\{\beta_0 + u_{0j} + u_{1j} z_{1ij}\} + \tau_0]\, [(1 - \tau_1 - \tau_0)\, (1 - g^{-1}\{\beta_0 + u_{0j} + u_{1j} z_{1ij}\}) + \tau_1]}, \tag{5.3}$$

para $\beta_k \neq 0$, $k = 1, 2$ e

$$\beta^*_k \approx \beta_k \frac{(1 - \tau_1 - \tau_0)\ \phi\{\beta_0 + u_{0j} + u_{1j} z_{1ij}\}}{\phi[\Phi^{-1}((1 - \tau_1 - \tau_0)\, g^{-1}\{\beta_0 + u_{0j} + u_{1j} z_{1ij}\} + \tau_0)]}, \; k = 1, 2, \tag{5.4}$$

em que $\beta_k \neq 0$, $\phi[.]$ é a distribuição Normal padronizada e $\Phi(.)$ é a densidade Normal padrão acumulada. Note que o fator de correção não depende da variável x_2

A avaliação do comportamento dos vieses das estimativas dos parâmetros $\beta_0, \beta_1, \beta_2$, σ_{u0} e σ_{u1} das equações (5.1) e (5.2), e a avaliação da qualidade da aproximação das equações (5.3) e (5.4) para os seus respectivos parâmetros β_1 e β_2 foi realizada através de simulações de Monte Carlo, como citado anteriormente.

Para ilustrar o procedimento dessa investigação, segue-se os seguintes passos:

1. Gera-se uma amostra de observações para a variável dependente binária t_{ij} conforme o modelo logístico hierárquico (5.1) com duas variáveis explicativas, x_{1ij} e x_{2ij} ambas com distribuição normal padrão, com um intercepto aleatório, u_{0j}, que segue uma distribuição normal padrão e, com uma variável explicativa $z_{1ij} = x_{1ij}$, relacionada ao efeito aleatório u_{1j} no qual considera-se $cov(u_{0j}, u_{1j}) = 0$. As variáveis independentes x_{1ij} e x_{2ij} são geradas independentemente dentro de cada grupo. Considera-se valores verdadeiros dos parâmetros $\beta_0 = -1$, $\beta_1 = 1$, $\beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$.

2. As estimativas dos parâmetros $\beta_0, \beta_1, \beta_2, \sigma_{u0}$ e σ_{u1} são obtidas resolvendo as equações (3.9) e (3.10) em que $\mathbf{D}(\theta) = \begin{bmatrix} \sigma^2_{u0} & 0 \\ 0 & \sigma^2_{u1} \end{bmatrix}$, pelo procedimento `glmmPQL` do `R`.

3. Gera-se a resposta observada y_{ij} a partir da resposta verdadeira t_{ij}. Em que $y_{ij} = 0$ quando $t_{ij} = 1$ com probabilidade $1 - \tau_1$ (sensibilidade) e $y_{ij} = 1$ quando $t_{ij} = 0$ com probabilidade $1 - \tau_0$ (especificidade).

4. Ajusta-se o modelo logístico hierárquico para a resposta observada y_{ij} resolvendo as equações (3.9) e (3.10) pelo procedimento `glmmPQL` do `R`, como mencionado anteriormente.

O mesmo procedimento é feito para o modelo probit hierárquico (5.2), sendo que a variável dependente binária t_{ij} é gerada conforme o modelo probit hierárquico (5.2).

Nos Quadros 11a a 15b (págs. 80 a 89), apresentados no Apêndice B.2, encontram-se os resultados de simulação dos parâmetros estimados do modelo logístico hierárquico (5.1) obtidos com erros na resposta, apresentando os desvios padrão, os valores preditos utilizando a aproximação (5.3), para $n = 5, 10, 20, 50$ e 100, e a quantidade de convergência pelo procedimento `glmmPQL`. Através dos resultados pode-se observar que com a consideração do intercepto e a inclinação aleatórios, as estimativas dos parâmetros deste modelo tendem a estar mais afastadas em relação ao verdadeiro valor dos parâmetros do modelo (5.1). Nota-se também que, em todas as situações, o viés, em módulo, associado ao parâmetro β_2 é menor que o associado ao parâmetro β_1, talvez pelo fato de que a expressão do erro corrigido para o parâmetro β_2 não dependa da variável x_2. Os vieses, em módulo, das estimativas dos parâmetros β_1 e β_2 aumentam quando fixa-se τ_0 e, τ_1 aumenta. Este mesmo comportamento é observado quando fixa-se τ_1 e τ_0 aumenta para todos os tamanhos de amostra considerados, mostrando assim que a magnitude destes vieses não depende do tamanho da amostra.

Em adição, ao comparar os vieses, em módulo, das estimativas dos parâmetros β_1 e β_2 com os vieses, em módulo, das estimativas corrigidas obtidas pela expressão (5.3), observa-se que esta expressão apresenta estimativas aproximadas para os parâmetros de interesse com resposta binária má classificada. Analisando o comportamento das estimativas dos parâmetros σ^2_{u0} e σ^2_{u1} observa-se que suas estimativas são próximas de zero, para diferentes valores de τ_0 e τ_1, indicando que estes parâmetros estão sendo sub-estimados.

Quando se compara os vieses, em módulo, das estimativas do parâmetro β_1 do modelo logístico hierárquico (4.2) com os vieses, em módulo, das estimativas do parâmetro β_1 do modelo logístico hierárquico (5.1), observa-se que para todas as combinações de τ_0 e τ_1 e, para diferentes tamanhos de amostras os vieses, em módulo, do modelo (5.1) são relativamente maiores. Este mesmo comportamento é observado quando se compara os vieses, em módulo, das estimativas corrigidas pela expressão (5.3) para o parâmetro β_1 com os vieses, em módulo, das estimativas corrigidas obtidas pela expressão (4.17) (ver Apêndices B.1 e B.2). Estes mesmos resultados são encontrados para o modelo probit hierárquico. Ainda, tem-se que as estimativas dos parâmetros do modelo probit (5.2) são muito distantes dos valores verdadeiros dos parâmetros β_0, β_1 e β_2 para diferentes tamanhos de amostra (ver Apêndice B.2).

Estes resultados sugerem que a especificação da função de ligação probit não desempenha muito bem no modelo com *inclinação e intercepto aleatórios*, quando utiliza-se o algoritmo da quase-verossimilhança penalizada, como comentado em Renard (2002).

5.4 Terceiro estudo de simulação

Neste terceiro estudo de simulação são avaliados os comportamento das estimativas das probabilidades de má classificação τ_0 e τ_1, consideradas como parâmetros desconhecidos e, das estimativas dos parâmetros β_0, β_1 e σ_{u0} obtidos a partir das equações (4.2) e (4.3) discutidas no capítulo 4 sob a má classificação da variável resposta binária nos modelos logístico e probit, respectivamente.

Para cada experimento de Monte Carlo, calculou-se a mediana e o intervalo interquartílico em lugar da média e do desvio padrão. As estimativas de máxima verossimilhança dos parâmetros de interesse foram obtidas utilizando o algoritmo de Newton-Rapson e Quadratura Gauss-Hermite para avaliar a integral em (3.20) via procedimento `gnlmm` do `R`. Esta função, ajusta equações de regressão não lineares especificadas para um e dois parâmetros das distribuições, onde o intercepto da regressão tem um efeito aleatório distribuído normalmente. Para maiores detalhes ver (http://finzi.psych.upenn.edu/R/library/repeated/html /gnlmm.html). Estas simulações geraram respostas binárias agrupadas verdadeiras e com erros, mas considerando a probabilidade de má classificação τ_0 e τ_1 como parâmetros desconhecidos para construir a verossimilhança em termos da forma (4.9) com função de ligação g dos tipos logit e probit.

Para investigar o modelo logístico (4.2), parametrizou-se a verossimilhança em termos de $\lambda_0 = \log(\sigma_{u0})$, $\omega_0 = \log(\frac{\tau_0}{1-\tau_0})$ e $\omega_1 = \log(\frac{\tau_1}{1-\tau_1})$ para considerar a limitação em $\sigma_{u0},\ \tau_0$ e τ_1, sendo este procedimento igual ao considerado no trabalho de Neuhaus (2002).

Para ilustrar o procedimento dessa investigação, segue-se os seguintes passos:

1. Gera-se uma amostra de observações para a variável dependente binária t_{ij} conforme o modelo logístico hierárquico (4.2) com um intercepto aleatório u_{0j} que segue uma distribuição normal padrão, e uma variável explicativa, x_{ij} que também segue uma distribuição normal padrão. A variável independente x_{ij} é gerada, independentemente, dentro de cada grupo. Os valores verdadeiros dos parâmetros do modelo logístico hierárquico considerados foram: $\beta_0 = -1$, $\beta_1 = 1$ e $\theta = \sigma^2_{u0} = 1$.

2. Gera-se a resposta observada y_{ij} a partir da resposta verdadeira t_{ij}. Em que $y_{ij} = 0$ quando $t_{ij} = 1$ com probabilidade $1 - \tau_1$ (sensibilidade) e $y_{ij} = 1$ quando $t_{ij} = 0$ com probabilidade $1 - \tau_0$ (especificidade).

3. Estima-se os parâmetros do modelo logístico hierárquico verdadeiro para a resposta observada y_{ij} e, as probabilidades de má classificação (τ_0 e τ_1) simultaneamente com o procedimento `gnlmm` do `R`, utilizando a parametrização alternativa, acima citada.

Vale ressaltar que considerando a verossimilhança em sua forma original, isto é, sem utilizar esta parametrização alternativa, consegue-se obter estimativas próximas aos valores

verdadeiros dos parâmetros de interesse, quando utiliza-se o procedimento `gnlmm`.

Para o modelo probit a verossimilhança é considerada em sua forma original. Vale a pena destacar que esta função de ligação não foi trabalhada em Neuhaus (2002). Para este modelo a variável dependente binária t_{ij} é gerada conforme o modelo probit hierárquico (4.3) e, a estimação dos parâmetros de interesse é feita utilizando o procedimento `gnlmm`.

O procedimento `gnlmm` do `R` não estima os componentes de variância quando se trabalha com variável resposta binomial. Isto ocorre devido ao fato de que as variáveis discretas frequentemente tem uma relação natural entre a média e variância da distribuição. Em termos de modelos multiníveis, isto conduz a uma relação entre os parâmetros da parte fixa e os parâmetros da parte aleatória, como citado na seção (2.4). Portanto para estimar os parâmetros de interesse simultaneamente com a razão de erros, padronizou-se a distribuição do efeito aleatório, $u_{0j} \sim N(0, \sigma^2_{u0})$, em que

$$b_{0j} = \frac{u_{0j}}{\sigma_{u0}} \;\Rightarrow\; u_{0j} = \sigma_{u0} b_{0j},$$

onde $b_{0j} \sim N(0,1)$. Assim, o modelo *verdadeiro*, correspondente a expressão (4.9), para a resposta observada y_{ij}, é escrito da seguinte forma

$$\mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, b_{0j}) = \{1 - \tau_1 - \tau_0\} g^{-1}(\beta_0 + \beta_1 x_{ij} + \sigma_{u0}\, b_{0j}) + \tau_0. \tag{5.5}$$

Vale ressaltar que quando se trabalha com função de ligação logit a expressão (5.5) é escrita utilizando a forma da parametrização alternativa

$$\begin{aligned}\mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, b_j) &= \{1 - \tau_1 - \tau_0\} \mathrm{logit}^{-1}(\beta_0 + \beta_1 x_{ij} + \sigma_{u0}\, b_{0j}) + \tau_0 \\ &= \{1 - \tau_1 - \tau_0\} \frac{\exp(\beta_0 + \beta_1 x_{ij} + \sigma_{u0}\, b_{0j})}{(1 + \exp(\beta_0 + \beta_1 x_{ij} + \sigma_{u0}\, b_{0j}))} + \tau_0, \end{aligned} \tag{5.6}$$

onde $\tau_0 = \frac{\exp(\omega_0)}{(1+\exp(\omega_0))}$, $\tau_1 = \frac{\exp(\omega_1)}{(1+\exp(\omega_1))}$ e $\sigma_{u0} = \exp(\lambda_0)$.

Quando se trabalha com a função de ligação probit tem-se

$$\mathrm{P}_V(y_{ij} = 1 \mid x_{ij}, b_{0j}) = \{1 - \tau_1 - \tau_0\} \mathrm{probit}^{-1}(\beta_0 + \beta_1 x_{ij} + \sigma_{u0}\, b_{0j}) + \tau_0. \tag{5.7}$$

Os Quadros 21a e 21b (págs. 101 e 102), apresentados no Apêndice B.3, apresentam a mediana e o intervalo interquartílico dos parâmetros estimados do modelo logístico hierárquico de efeito misto do tipo (5.6), e as razões de erros, além da quantidade de convergência pelo procedimento `gnlmm` para o tamanho de amostra $n = 5$. Através dos resultados apresentados nos quadros pode-se notar que a mediana observada dos parâmetros do modelo logístico hierárquico são maiores em relação aos valores verdadeiros dos parâmetros, indicando que estes parâmetros estão sendo levemente superestimados, enquanto que as estimativas das

probabilidades da má classificação são menores, indicando que os parâmetros τ_0 e τ_1 estão sendo subestimados. Nota-se também que há casos em que as estimativas dos parâmetros estão fora da diferença máxima de 10% dos seus valores verdadeiros, por exemplo, fixando $\tau_0 = 0,05$ e variando τ_1, as estimativas do parâmetro β_1 estão todas fora da diferença máxima de 10% do seu valor verdadeiro. Este mesmo comportamento é observado para a estimativa da probabilidade do erro τ_0, por exemplo, quando fixa-se $\tau_0 = 0,10$ e varia τ_1 as estimativas de τ_0 estão todas fora da diferença máxima de 10% do seu valor verdadeiro. Isto pode ser uma indicação de que o procedimento `gnlmm` não ajusta bem os parâmetros de interesse para tamanho de amostra muito pequeno. Comparando os Quadros 21a e 21b com os valores apresentados na Tabela 4 do artigo de Neuhaus (2002) pode-se observar que as conclusões são similares, apesar de se tratar de modelos diferentes.

Os Quadros 22a a 25b (págs. 102 a 109), apresentam os resultados para os tamanhos de amostras $n = 10, 20, 50$ e 100, nos quais observa-se que as estimativas medianas dos parâmetros do modelo logístico hierárquico (5.6) e dos parâmetros de probabilidade de má classificação estão todas dentro da diferença máxima de 10% dos seus valores verdadeiros.

Os Quadros 26a a 30b (págs. 110 a 119), apresentados também no Apêndice B.3, contem os resultados das simulações para o modelo probit hierárquico (5.7), para $n = 5, 10, 20, 50$ e 100, respectivamente. De forma similar ao que acontece com algumas estimativas dos parâmetros do modelo logit para tamanho de amostra 5, pode-se notar que há casos em que as estimativas dos parâmetros de interesse estão fora da diferença máxima de 10% dos seus valores verdadeiro, por exemplo, fixando $\tau_0 = 0,05$ e considerando $\tau_1 = 0,05$ ou $\tau_1 = 0,15$, nota-se que as estimativas do parâmetro β_1 estão fora da diferença máxima de 10% do valor verdadeiro do parâmetro (ver Quadros 26a a 26b, no Apêndice B.3). Para os demais tamanhos de amostra, observa-se que as estimativas medianas dos parâmetros do modelo probit (5.7) e dos parâmetros das probabilidades de má classificação estão todas dentro da diferença máxima de 10% dos seus valores. Estes resultados sugerem que quando se especifica corretamente o modelo subjacente, obtem-se estimativas próximas dos verdadeiros valores dos parâmetros dos modelos considerados, logístico hierárquico (5.6) e probit hierárquico (5.7), respectivamente, e das probabilidades de má classificação dos respectivos modelos (ver Apêndice B.3).

5.5 Quarto estudo de simulação

Este quarto estudo de simulação tem por objetivo avaliar os comportamentos das estimativas das probabilidades de má classificação τ_0 e τ_1, consideradas como parâmetros desconhecidos num modelo com mais de uma variável explicativa.

Os resultados são obtidos a partir do modelo misto linear generalizado das equações

(5.1) e (5.2) sob a má classificação da variável resposta binária nos modelos logístico e probit, respectivamente.

Para cada experimento de Monte Carlo calculou-se a mediana e o intervalo interquartílico em lugar da média e do desvio padrão, como citado no terceiro estudo de simulação. As estimativas de máxima verossimilhança para os parâmetros de interesse foram obtidas utilizando o algoritmo de Newton-Rapson e Quadratura Gauss-Hermite para avaliar a integral em (3.20) pelo procedimento `gnlmm` do `R`. Estas simulações geraram respostas binárias agrupadas verdadeiras e com erros, em que as probabilidades de má classificação τ_0 e τ_1 são consideradas parâmetros desconhecidos para construir a verossimilhança em termos da forma (4.9) com função de ligação g, logit ou probit.

Para investigar o modelo logístico (5.1), parametrizou-se a verossimilhança em termos de $\lambda_0 = \log(\sigma_{u0})$, $\lambda_1 = \log(\sigma_{u1})$, $\omega_0 = \log(\frac{\tau_0}{1-\tau_0})$ e $\omega_1 = \log(\frac{\tau_1}{1-\tau_1})$ para considerar a limitação em $\sigma_{u0},\ \sigma_{u1},\ \tau_0$ e τ_1, como comentado anteriormente.

Para ilustrar o procedimento dessa investigação, segue-se os seguintes passos:

1. Gera-se uma amostra de observações para a variável dependente binária t_{ij} conforme o modelo logístico hierárquico (5.1) com duas variáveis explicativas, x_{1ij} e x_{2ij} ambas com uma distribuição normal padrão, com um intercepto aleatório, u_{0j}, que segue uma distribuição normal padrão e, com uma variável explicativa $z_{1ij} = x_{1ij}$, relacionada ao efeito aleatório u_{1j} no qual considera-se $cov(u_{0j}, u_{1j}) = 0$. As variáveis independentes x_{1ij} e x_{2ij} são geradas independentemente dentro de cada grupo. Os valores verdadeiros dos parâmetros do modelo logístico hierárquico considerados são : $\beta_0 = -1$, $\beta_1 = 1$, $\beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$.

2. Gera-se a resposta observada y_{ij} a partir da resposta verdadeira t_{ij}. Em que $y_{ij} = 0$ quando $t_{ij} = 1$ com probabilidade $1 - \tau_1$ (sensibilidade) e $y_{ij} = 1$ quando $t_{ij} = 0$ com probabilidade $1 - \tau_0$ (especifidade).

3. Estima-se os parâmetros do modelo logístico hierárquico verdadeiro para a resposta y_{ij} e, as probabilidades de má classificação (τ_0 e τ_1) simultaneamente com o procedimento `gnlmm` do `R`, utilizando a parametrização alternativa, acima citada.

Para o modelo probit (5.2) a verossimilhança é considerada em sua forma original. A variável dependente binária t_{ij} é gerada conforme o modelo probit hierárquico (5.2) e, a estimação dos parâmetros de interesse é feita utilizando o procedimento `gnlmm`. Como o procedimento `gnlmm` do `R` não estima os componentes de variância quando se trabalha com variável resposta binomial, como comentado anteriormente, padronizou-se as distribuições dos efeitos aleatórios, $u_{0j} \sim N(0, \sigma^2_{u0})$ e $u_{1j} \sim N(0, \sigma^2_{u1})$, em que

$$b_{0j} = \frac{u_{0j}}{\sigma_{u0}} \ \Rightarrow\ u_{0j} = \sigma_{u0} b_{0j} \ \text{ e } \ b_{1j} = \frac{u_{1j}}{\sigma_{u1}} \ \Rightarrow\ u_{1j} = \sigma_{u1} b_{1j},$$

onde $b_{0j} \sim N(0, 1)$ e $b_{1j} \sim N(0, 1)$. Assim, supondo que a probabilidade de má classificação

não depende da covariável e nem do efeito aleatório, o modelo *verdadeiro* para a resposta observada y_{ij} é escrito da seguinte forma

$$\begin{aligned} \mathrm{P}_V(y_{ij}=1 \mid x_{1ij}, x_{2ij}, b_{0j}, b_{1j}) = \{1-\tau_1-\tau_0\} g^{-1}(\beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} \\ + \sigma_{u0}\, b_{0j} + \sigma_{u1}\, b_{1j}\, z_{1ij}) + \tau_0, \end{aligned} \tag{5.8}$$

em que a equação (5.8) é uma extensão da expressão (5.5). Denota-se $\beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \sigma_{u0}\, b_{0j} + \sigma_{u1}\, b_{1j}\, z_{1ij}$ por W para simplificar a notação da expressão apresentada a seguir. Vale ressaltar que quando se trabalha com função de ligação logit a expressão (5.8) é escrita utilizando a forma da parametrização alternativa

$$\begin{aligned} \mathrm{P}_V(y_{ij}=1 \mid x_{1ij}, x_{2ij}, b_{0j}, b_{1j}) &= \{1-\tau_1-\tau_0\} \mathrm{logit}^{-1}(W) + \tau_0 \\ &= \{1-\tau_1-\tau_0\} \frac{\exp(W)}{(1+\exp(W)} + \tau_0, \end{aligned} \tag{5.9}$$

considerando a parametrização $\tau_0 = \frac{\exp(\omega_0)}{(1+\exp(\omega_0))}$, $\tau_1 = \frac{\exp(\omega_1)}{(1+\exp(\omega_1))}$, $\sigma_{u0} = \exp(\lambda_0)$ e $\sigma_{u1} = \exp(\lambda_1)$.

Quando se trabalha com a função de ligação probit tem-se

$$\mathrm{P}_V(y_{ij}=1 \mid x_{1ij}, x_{2ij}, b_{0j}, b_{1j}) = \{1-\tau_1-\tau_0\} \mathrm{probit}^{-1}(W) + \tau_0 \tag{5.10}$$

Nos Quadros 31a a 35b (págs. 120 a 129), apresentados no Apêndice B.4, encontram-se os resultados de simulação dos parâmetros estimados do modelo logístico hierárquico (5.9) com probabilidade de má classificação desconhecida, apresentando as medianas e os intervalos interquartílicos das estimativas dos parâmetros deste modelo e das probabilidades de má classificação para $n = 5, 10, 20, 50$ e 100, além da quantidade de convergência pelo procedimento `gnlmm`. Através dos resultados pode-se observar que para o tamanho de amostra $n = 5$, há casos em que as estimativas medianas dos parâmetros de interesse estão todas fora dos 10% dos seus valores verdadeiros, por exemplo, quando fixa-se $\tau_0 = 0,15$ e $\tau_1 = 0,15$ as estimativas dos parâmetros β_0, β_1, β_2 e σ_{u0} estão todas fora da diferença máxima de 10% dos valores verdadeiros de seus respectivos parâmetros (ver Quadro 31a e 31b, no Apêndice B.4). Observa-se também que, para os demais tamanhos de amostra $n = 10, 20, 50$ e 100, as estimativas medianas dos parâmetros de interesse estão todas dentro da diferença máxima de 10% dos valores verdadeiros dos seus respectivos parâmetros, (ver Quadros 31a a 31b, no Apêndice B.4).

Comportamentos semelhantes foram observados para o modelo probit hierárquico (5.10), em que se considera as probabilidades de má classificação desconhecida (ver Quadros 36a a 40b, no Apêndice B.4).

Quando se compara a estimativa mediana do parâmetro β_1 do modelo logístico hierárquico (5.6), do terceiro estudo de simulação, em que se considera as probabilidades de má classificação desconhecidas, com a estimativa mediana do parâmetro β_1 do modelo logístico hierárquico (5.9), do quarto estudo, que também considera as probabilidades de má classificação desconhecidas, observa-se que para diferentes tamanhos de amostras as estimativas medianas dos parâmetros do modelo (5.9) são levemente inferiores, situação inversa ocorre quando se compara as estimativas das probabilidades de má classificação. Estes mesmos comportamentos também foram observados quando considera-se o modelo probit hierárquico (5.10) com probabilidades de má classificação desconhecidas (ver Apêndices B.3 e B.4).

6 Aplicação

Este capítulo tem por objetivo principal a ilustração com dados reais da proposta de considerar erros de classificação em modelos multiníveis, não tendo por intuito a escolha do melhor modelo para este conjunto de dados. Esta escolha requeriria um processo mais cuidadoso que envolveria uma discussão ampla sobre a seleção dos componentes de variância envolvidos nestes tipos de modelos.

Os dados desta aplicação estão relacioanados ao indicador de exposição HIV/AIDS através do ato sexual e, estão disponíveis na página http://www.aids.gov.br, dentro do Relatório de pesquisa: Projeto "Comportamento Sexual da População Brasileira e Percepções do HIV/AIDS" de setembro do ano 2000 do Programa Nacional de DST e AIDS do Ministério da Saúde.

A referida pesquisa foi realizada de dezembro de 1997 a dezembro de 1998 tendo como universo da pesquisa de estudo indivíduos de ambos os sexos, com idade entre 16 e 65 anos (em três categorias) moradores nas áreas urbanas de 169 micro-regiões do Brasil.

A pesquisa tinha por objetivo geral, identificar representações, comportamento, atitudes e práticas sexuais da população brasileira, e conhecimento sobre HIV/Aids, com vistas a estabelecer estratégias de intervenções preventivas das DST e HIV (relatório da pesquisa, pág. 2).

A amostra estudada foi de 3315 pessoas. O plano amostral foi do tipo estratificado em múltiplos estágios, com probabilidades desiguais, sorteando-se, no primeiro estágio, micro-regiões, no segundo, setores censitários, no terceiro, domicílios particulares e, finalmente no quarto, uma pessoa de 16 a 65 anos. Os estratos correspondem às regiões Norte-Nordeste, Sul expandido e Centro-Oeste expandido (relatório da pesquisa, pág. 3).

Na pesquisa considerou-se como "indicador de fator de exposição, na sua forma dicotômica restringe-se ao uso ou não do preservativo, pois foram considerados como não expostos apenas os indivíduos que declararam usar camisinhas nas suas relações sexuais" (relatório da pesquisa pág 106). Assim, este indicador está sujeito a erro de classificação da resposta, pois a pessoa pode se sentir constrangida ao responder a pergunta, ou ter algum tipo de esquecimento.

Nesta aplicação de modelos multiníveis para resposta binária, pretende-se apresentar modelos logísticos hierárquicos de 2 níveis admitindo erro na variável resposta indicadora de exposição. Para exemplificar a técnica, com base nos dados disponíveis, considerou-se como primeiro nível as pessoas e como segundo nível o estado em que elas moravam. As informações sobre os estados foram obtidas do banco eletrônico do Atlas de Desenvolvimento Humano (http://www.ipea.gov.br/TemasEspeciais/especiais.php).

As variáveis do primeiro nível consideradas foram: idade, estado conjugal, classe socio econômica, segundo a divisão do IBGE, nível de instrução, sexo e religião. As variáveis do segundo nível consideradas foram: porcentagem de analfabetismo, Índice de Desenvolvimento Humano de Renda, Índice de Desenvolvimento de Educação e Índice de Desenvolvimento Humano Municipal.

Foram considerados modelos com todas as variáveis fixas do tipo:

$$\eta_{ij} = \beta_{0j} + \sum_{l\in\{1,\ldots,7\}} \beta_{lj} x_{lij}, \qquad \text{(nível 1)}$$

$$\beta_{0j} = \gamma_{00} + \sum_{r\in\{1,\ldots,4\}} \gamma_{0r} Z_{rj} + u_{0j}; \qquad \text{(nível 2)}$$

$$\beta_{lj} = \gamma_{l0},$$

em que $u_{0j} \sim N\left(0, \sigma^2_{u0}\right)$, u_{0j} independentes, onde

x_{lij} l-ésima característica do i-ésimo indivíduo do estado j (preditor do nível 1);

β_{0j} intercepto do j-ésimo estado;

β_{lj} coeficiente que representa o efeito da variável x_{lij} do indivíduo no j-ésimo estado, $l = 1, \ldots, 7$;

γ_{00} é o intercepto na modelagem de β_{0j};

Z_{rj} característica do j-ésimo estado (preditor do nível 2), $r = 1, \ldots 4$;

γ_{0r} coeficiente que representa o efeito da variável de estado Z_{rj} em β_{0j};

u_{0j} efeito aleatório que representa o desvio do coeficiente β_{0j} do estado j em relação ao modelo ajustado no nível 2;

γ_{l0} é o intercepto no modelo do nível 2 para o coeficiente β_{lj}, $l = 1, \ldots, 7$.

As variáveis no nível do indivíduo foram:

- x_{1ij} : 1 = idade do indivíduo entre 26 e 34 anos, 0 = caso contrário,
- x_{2ij} : 1 = idade do indivíduo superior a 35 anos, 0 = caso contrário ;
- x_{3ij} : 1 = indivíduo casado ou em união consensual, 0 = em outra condição;
- x_{4ij} : 1 = classes econômicas A ou B, 0 = classes econômicas C, D ou E;
- x_{5ij} : 1 = analfabeto ou com fundamental incompleto, 0 = em outra condição;
- x_{6ij} : 1 = católica, 0 = outra religião ou nenhuma;
- x_{7ij} : 1 = sexo masculino, 0 = sexo feminino.

As variáveis no nível do estado foram:

- Z_{1j} : porcentagem de analfabetismo; Z_{2j} índice de Desenvolvimento de Renda; Z_{3j} índice de Desenvolvimento de Educação e Z_{4j} índice de Desenvolvimento Humano Municipal.

Obtendo modelos combinados da forma:

$$\eta_{ij} = \gamma_{00} + \sum_{r\in\{1,...,4\}} \gamma_{0r} Z_{rj} + u_{0j} + \sum_{l\in\{1,...,7\}} \gamma_{l0} x_{lij}.$$

Utilizando as parametrizações descritas nos capítulos anteriores e considerando as probabilidades de erro de classificação τ_0 e τ_1, tem-se que

$$P_V\left(y_{ij} = 1 \mid x_{1ij}, ..., x_{lij}, b_{oj}\right) =$$

$$(1 - w_0 - w_1)\,\text{logit}^{-1}\left(\gamma_{00} + \sum_{r\in\{1,...,4\}} \gamma_{0r} Z_{rj} + \sigma_{u0}\, b_{0j} + \sum_{l\in\{1,...,7\}} \gamma_{l0} x_{lij}\right) + w_0,$$

em que $w_0 = \log\left(\frac{\tau_0}{1-\tau_0}\right)$, $w_1 = \log\left(\frac{\tau_1}{1-\tau_1}\right)$ e $b_{0j} = \frac{u_{0j}}{\sigma_{u0}} \sim N(0,1)$. Note que esta parametrização tem a vantagem que permite construir testes de hipóteses para o parâmetro σ_{u0}, que não possui restrições em relação ao espaço paramétrico.

Foram considerados alguns modelos em que algumas das variáveis foram consideradas como efeito aleatório, assim as probabilidades ajustadas foram do tipo:

$$P_V\left(y_{ij} = 1 \mid x_{1ij}, ..., x_{lij}, b_{oj}, b_{1j}, ..., b_{lj}\right) = (1 - w_0 - w_1)$$

$$\text{logit}^{-1}\left(\gamma_{00} + \sum_{r\in\{1,...,4\}} \gamma_{0r} Z_{rj} + \sigma_{u0} b_{0j} + \sum_{l\in\{1,...,7\}} \gamma_{l0} x_{lij} + \sum_{l\in\{1,...,7\}} \sigma_{ul}\, b_{lj} x_{lij}\right) + w_0.$$

No apêndice D (pág. 164) apresenta-se um dos programas de ajuste utilizados, e no apêndice E (pág. 166) as saídas de todos os modelos considerados.

Dos Quadros 1 e 2 podem-se fazer os seguintes comentários: ao considerar o modelo unicamente com as variáveis do primeiro nível como fixas (Modelo 1 - M1) observa-se, por meio da estatística de Wald, que os coeficientes das variáveis nível de intrução e classe social não são significativos. Ao ajustar o modelo em que no lugar destas variáveis são considerados a porcentagem de analfabetismo e ID de renda (Modelo 2 - M2), os coeficientes relacionados a estas variáveis também resultam não significativos. Destacando também que as probabilidades de erro de classificação destes modelos são similares, $\tau_0 = 0,164$ e $\tau_1 = 0,123$, referente ao modelo (M1) e $\tau_0 = 0,168$ e $\tau_1 = 0,124$, referente ao modelo (M2).

Quadro 1. Estimativas dos parâmetros do modelo M1, utilizando $\tau_0 = \tau_1 = 0,269$.

Parâmetros	Estimativa	Erro padrão	Estatística de Wald	p-valor
γ_{00} : Intercepto	−0,72520	0,89100	0,66238	0,41572
γ_{10} : Idade(26-34)	2,34310	0,96610	5,88210	0,01530
γ_{20} : Idade($\geq$35)	2,08080	0,94350	4,86432	0,02742
γ_{30} : Casado	7,16100	1,57220	20,74662	<0,00001
γ_{40} : Classe econômicas A/B	−0,51180	0,45410	1,27021	0,25973
γ_{50} : Analfabeto	0,24110	0,38670	0,38871	0,53298
γ_{60} : Católica	−1,12240	0,46830	5,74468	0,01654
γ_{70} : Masculino	−1,51150	0,51610	8,57829	0,00340
σ_{u0}	0,11874	0,80066	2,07834	0,14940
τ_0	0,16412	0,53937	106,47448	<0,00001
τ_1	0,12315	0,52697	330,21392	<0,00001

Quadro 2. Estimativas dos parâmetros do modelo M2, utilizando $\tau_0 = \tau_1 = 0,269$.

Parâmetros	Estimativa	Erro padrão	Estatística de Wald	p-valor
γ_{00} : Intercepto	−6,30023	7,54060	0,69807	0,40343
γ_{10} : Idade(26-34)	2,42622	1,18789	4,17163	0,04111
γ_{20} : Idade($\geq$35)	2,24103	1,14031	3,86233	0,04938
γ_{30} : Casado	7,39040	1,90757	15,00985	<0,00011
γ_{40} : ID de Renda	6,53898	8,74297	0,55937	0,454513
γ_{50} : (%) Analfabetismo	0,05636	0,09285	0,36842	0,543867
γ_{60} : Católica	−1,15832	0,49512	5,47322	0,01931
γ_{70} : Masculino	−1,68704	0,59048	8,16276	0,00428
σ_{u0}	0,01535	1,00000	0,09462	0,75838
τ_0	0,16773	0,53802	110,52120	<0,00001
τ_1	0,12377	0,53061	254,89399	<0,00001

A nível de ilustração, considera-se como modelo escolhido aquele que contém só as variáveis do primeiro nível como fixas sem considerar nível de instrução e classe social (Modelo 3 - M3). Do Quadro 3, pode-se concluir que: as probabilidades de estar exposto aumentam com a idade, para os individuos de sexo feminino, indivíduos casados ou em união consensual, e para outras religiões ou sem religião em relação à católica.

Quadro 3. Estimativas dos parâmetros do modelo M3, utilizando $\tau_0 = \tau_1 = 0,269$.

Parâmetros	Estimativa	Erro padrão	Estatística de Wald	p-valor
γ_{00} : Intercepto	−0,87040	0,87550	0,99158	0,03190
γ_{10} : Idade(26-34)	2,29620	0,98160	5,44721	0,01960
γ_{20} : Idade($\geq$35)	2,08900	0,97500	4,59064	0,03210
γ_{30} : Casado	7,1990	1,58080	20,74044	<0,00001
γ_{60} : Católica	−1,14240	0,47360	5,81780	0,01590
γ_{70} : Masculino	−1,60010	0,54690	8,56080	0,00340
σ_{u0}	0,00003	1,1e+26	0,03400	0,85370
τ_0	0,16520	0,54110	337,41525	<0,00001
τ_1	0,12346	0,52570	96,36183	<0,00001

Nota-se que os valores das probabilidades de classificação de todos os modelos considerados são superiores a 0,10 que pode-se considerar um valor alto, o qual motiva a realização de uma pesquisa para uma melhor definição do indicador de exposição e de uma escolha cuidadosa das possíveis variáveis explicativas deste fator. As conclusões obtidas no relatório da pesquisa são similares as que foram encontradas no presente estudo. Vale a pena ressaltar que o modelo ilustrado com certeza não é o melhor modelo que se ajusta aos dados de pesquisa usando o método dos modelos multiníveis para resposta binária má classificada.

7 Conclusões e sugestões

Os modelos multiníveis para resposta binária são uma ferramenta muito útil para realizar análises de dados binários. Estes modelos assumem que a variável resposta binária é medida sem erro, isto na prática não acontece devido a uma pergunta de difícil resposta ou a problemas que requerem um diagnóstico médico. Na presente dissertação apresentou-se a expressão geral para obter estimativas consistentes dos parâmetros de interesse no modelo multinível de 2 níveis para resposta binária má classificada, para o caso de duas funções de ligação, logit e probit. Na literatura disponível não encontramos nenhum trabalho com esta abordagem, sendo portanto nossa contribuição original. Considerou-se duas situações: a primeira em que as probabilidades de má classificação eram de forma conhecida e na segunda de forma desconhecida.

A investigação do efeito de má classificação da resposta para dados binários agrupados foi realizada no presente trabalho por simulações de Monte Carlo, para avaliar o desempenho dos estimadores de máxima verossimilhança em uma classe particular de modelos mistos lineares generalizados para resposta binária. O resultado mostra que ignorar os erros na resposta pode levar a perda substancial de informações sobre os efeitos dos parâmetros de interesse e que a própria análise da resposta com o erro corrigido pode ainda levar a uma perda substancial na precisão das estimativas. Além disso, quando as probabilidades de má classificação são desconhecidas, pode-se estimá-las simultaneamente com os parâmetros do modelo misto e obter estimativas muito próximas dos valores verdadeiros. Isto ocorre sempre que o modelo subjacente é especificado corretamente.

Outro aspecto importante observado foi com relação aos dois métodos aproximados utilizados, o algoritmo da Quase-verossimilhança Penalizada, em que se utilizou o procedimento `glmmPQL` do `R` e o algoritmo da Quadratura de Gauss-Hermite, em que se utilizou o procedimento `glnmm` do `R`, para obter as estimativas dos parâmetros dos modelos estudados. Em relação aos modelos ajustados pelo procedimento `glmmPQL`, pode-se concluir que a função de ligação probit não desempenha bem no modelo com *inclinação e interceptos aleatórios* quando se utiliza o algoritmo da Quase-verossimilhança Penalizada, confirmando os resultados obtidos por Renard (2002) e, ainda tem-se que o procedimento `glmmPQL` não estima a função de variância, logo as estimativas dos efeitos aleatórios são altamente viesadas. Já utilizando o procedimento `gnlmm` é possível estimar os parâmetros relacionados aos efeitos fixos e aleatórios do modelo sob investigação e, obter estimativas consistentes para os parâmetros de interesse quando se padronizam as distribuições dos efeitos aleatórios, pois este procedimento não estima os componentes de variância quando se trabalha com variável resposta binomial. Ainda constatou-se que este procedimento não tem bom desempenho quando o

tamanho de amostra é muito pequeno.

O resultado da aplicação que objetivou estimar a probabilidade de um indivíduo estar exposto à contaminação pelo HIV por meio de modelos multiníveis considerando a existência de erros da resposta, revelou que estes erros podem ser consideráveis, o que motiva o uso da metodologia apresentado neste trabalho.

Em resumo pode-se concluir que incorporar as probabilidades de má classificação no método dos modelos multiníveis permitem tanto uma mensuração dos erros quanto uma melhor inferência dos parâmetros associados.

Embora o trabalho aqui assuma que a probabilidade de má classificação não depende da covariável, segundo Neuhaus (2002), pode-se assumir que a probabilidade de má classificação depende das covariáveis conhecidas e estimar os parâmetros de interesse. Contudo, quando a teoria considera a dependência do erro na covariável, isto pode dificultar a modificação específica na rotina do software para considerar estes erros.

Sugestões para a continuidade deste trabalho:

• Realizar um estudo mais detalhado da questão da probabilidade de má classificação depender da covariável, considerando a classe de modelos mistos lineares generalizados.

• Estudar o modelo logístico multinível com má classificação na resposta e com erros nas covariáveis.

• Fazer um estudo dos testes dos efeitos aleatórios para a classe de modelos mistos lineares generalizados.

Apêndice

A. Demonstrações

A.1 - Prova da variância de e^*_{ij}, representada na expressão (3.14), que é utilizada no algoritmo QVP2.

Na secção 3.2.1 denotamos $g^{'}(\hat{\mu}^{\mathbf{u}}_{ij})(y_{ij} - \hat{\mu}^{\mathbf{u}}_{ij}) - \frac{1}{2}\, g^{''}(\hat{\mu}^{\mathbf{u}}_{ij})(y_{ij} - \hat{\mu}^{\mathbf{u}}_{ij})^2$ de e^*_{ij} e assumindo que y^*_{ij} é distribuído normalmente. Por definição tem-se que $\hat{e}_{ij} = (y_{ij} - \hat{\mu}^{\mathbf{u}}_{ij})$ e $\mathrm{Var}(\hat{e}_{ij}) = a_j(\phi)v(\hat{\mu}^{\mathbf{u}}_{ij})$. Portanto, basta calcular $\mathrm{Var}(e^*_{ij})$,

$$\begin{aligned}
\mathrm{Var}(e^*_{ij}) &= \mathrm{Var}\left(g^{'}(\hat{\mu}^{\mathbf{u}}_{ij})\, \hat{e}_{ij} - \frac{1}{2}\, g^{''}(\hat{\mu}^{\mathbf{u}}_{ij})\, \hat{e}_{ij} \right) \\
&= [g^{'}(\hat{\mu}^{\mathbf{u}}_{ij})]^2 \,\mathrm{Var}(\hat{e}_{ij}) - \frac{1}{4}[g^{''}(\hat{\mu}^{\mathbf{u}}_{ij})]^2 \,\mathrm{Var}(\hat{e}^2_{ij}) \\
&= [g^{'}(\hat{\mu}^{\mathbf{u}}_{ij})]^2 \, a_j(\phi)v(\hat{\mu}^{\mathbf{u}}_{ij}) - \frac{1}{4}[g^{''}(\hat{\mu}^{\mathbf{u}}_{ij})]^2 \,\mathrm{Var}(\hat{e}^2_{ij}),
\end{aligned}$$

Precisa-se calcular $\mathrm{Var}(\hat{e}^2_{ij})$. Seja $\hat{e}^2_{ij} = e^2$, assim, $\mathrm{Var}(\hat{e}^2_{ij})$ pode ser escrita como $\mathrm{Var}(e^2) = \mathrm{E}(e^4) - \mathrm{E}(e^2)$. Sabe-se que $e \sim N(0, a_j(\phi)v(\mu^{\mathbf{u}}_{ij}))$, então sua função denidade é da forma:

$$f(e) = \frac{1}{\sqrt{2\pi a_j(\phi)v(\mu^{\mathbf{u}}_{ij})}} \exp\left\{ -\frac{1}{2a_j(\phi)v(\mu^{\mathbf{u}}_{ij})}\, e^2 \right\},\ e \in I\!R \text{ e } a_j(\phi)v(\mu^{\mathbf{u}}_{ij}) > 0,$$

$$\mathrm{E}(e^4) = \int_{-\infty}^{\infty} e^4 \frac{1}{\sqrt{2\pi a_j(\phi)v(\mu^{\mathbf{u}}_{ij})}} \exp\left\{ -\frac{1}{2a_j(\phi)v(\mu^{\mathbf{u}}_{ij})}\, e^2 \right\} de,$$

Seja $z = \frac{e}{\sqrt{a_j(\phi)v(\mu^{\mathbf{u}}_{ij})}}$ e $dz = \frac{1}{\sqrt{a_j(\phi)v(\mu^{\mathbf{u}}_{ij})}} de$. Daí, obtem-se

$$\begin{aligned}
\mathrm{E}(e^4) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [a_j(\phi)v(\mu^{\mathbf{u}}_{ij})]^2 z^4 \exp\left\{ -\frac{z^2}{2} \right\} dz \\
&= \frac{[a_j(\phi)v(\mu^{\mathbf{u}}_{ij})]^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^4 \exp\left\{ -\frac{z^2}{2} \right\} dz,
\end{aligned}$$

a integração é feita por partes. Considerando:

$$u = z^3 \ \Rightarrow\ du = 3z^2 dz$$

$$dv = z\exp\left\{ -\frac{z^2}{2} \right\} dz \ \Rightarrow\ v = -\exp\left\{ -\frac{z^2}{2} \right\}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} z^4 \exp\left\{ -\frac{z^2}{2} \right\} dz &= -z^3 \exp\left\{ -\frac{z^2}{2} \right\} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} 3z^2 \exp\left\{ -\frac{z^2}{2} \right\} dz \\
&= 0 + \int_{-\infty}^{\infty} 3z^2 \exp\left\{ -\frac{z^2}{2} \right\} dz = \int_{-\infty}^{\infty} 3z^2 \exp\left\{ -\frac{z^2}{2} \right\} dz,
\end{aligned}$$

a integração é feita por partes. Considerando:

$$a = 3z \Rightarrow da = 3dz$$

$$db = z\text{exp}\Big\{ -\frac{z^2}{2} \Big\} \Rightarrow b = -\text{exp}\Big\{ -\frac{z^2}{2} \Big\}$$

$$\int_{-\infty}^{\infty} 3z^2\text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz = -3z\text{exp}\Big\{ -\frac{z^2}{2} \Big\}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} 3\text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz$$

$$= 0 + \int_{-\infty}^{\infty} 3\text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz = 3\int_{-\infty}^{\infty} \text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz,$$

então

$$\int_{-\infty}^{\infty} z^4\text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz = 3\int_{-\infty}^{\infty} \text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz.$$

Portanto,

$$\begin{aligned}
\text{E}(e^4) &= \frac{[a_j(\phi)v(\mu_{ij}^{\mathbf{u}})]^2}{\sqrt{2\pi}}\, 3\int_{-\infty}^{\infty} \text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz \\
&= [a_j(\phi)v(\mu_{ij}^{\mathbf{u}})]^2 3\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\text{exp}\Big\{ -\frac{z^2}{2} \Big\}dz \\
&= 3a_j^2(\phi)v^2(\mu_{ij}^{\mathbf{u}}).
\end{aligned}$$

Como $\text{var}(e) = \text{E}(e^2) - \text{E}^2(e^2) \Rightarrow \text{E}^2(e^2) = \text{var}(e) = a_j(\phi)v(\mu_{ij}^{\mathbf{u}})$. Então,

$$\begin{aligned}
\text{var}(e^2) &= \text{E}(e^4) - \text{E}^2(e^2) \\
&= 3a_j^2(\phi)v^2(\mu_{ij}^{\mathbf{u}}) - [a_j(\phi)v(\mu_{ij}^{\mathbf{u}})]^2 \\
&= 3a_j^2(\phi)v^2(\mu_{ij}^{\mathbf{u}}) - a_j^2(\phi)v^2(\mu_{ij}^{\mathbf{u}}) \\
&= 2a_j^2(\phi)v^2(\mu_{ij}^{\mathbf{u}}).
\end{aligned}$$

Logo, voltando para $\text{var}(e_{ij}^*)$, tem-se

$$\begin{aligned}
\text{var}(e_{ij}^*) &= [g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})]^2\, a_j(\phi)v(\hat{\mu}_{ij}^{\mathbf{u}}) - \frac{1}{4}[g^{''}(\hat{\mu}_{ij}^{\mathbf{u}})]^2\, 2a_j^2(\phi)v^2(\mu_{ij}^{\mathbf{u}}) \\
&= [g^{'}(\hat{\mu}_{ij}^{\mathbf{u}})]^2\, a_j(\phi)v(\hat{\mu}_{ij}^{\mathbf{u}}) - \frac{1}{2}[g^{''}(\hat{\mu}_{ij}^{\mathbf{u}})]^2\, a_j^2(\phi)v^2(\mu_{ij}^{\mathbf{u}}),
\end{aligned}$$

fica provada a variância de e_{ij}^*, representada na expressão (3.14).

A.2 - Prova da expressão (4.16) que avalia aproximadamente a relação entre β_1^* e β_1. Para avaliar a relação aproximada entre β_1^* e β_1 expande-se a função $H(\beta_1)$ (4.15) em série de Taylor sobre o valor $\beta_1 = 0$.

A função $H(\beta_1)$ considerada neste trabalho é da forma

$$H(\beta_1) = g\{P_V(y_{ij} = 1 \mid x_{ij} + 1, u_{0j})\} - g\{P_V(y_{ij} = 1 \mid x_{ij}, u_{0j})\},$$

onde

$$P_V(y_{ij} = 1 \mid x_{ij} + 1, u_{0j}) = (1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + \beta_1(x_{ij} + 1) + u_{0j}\} + \tau_0$$

e

$$P_V(y_{ij} = 1 \mid x_{ij}, u_{0j}) = (1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + \beta_1 x_{ij} + u_{0j}\} + \tau_0$$

em que P_V denota a probabilidade sob o verdadeiro modelo y_{ij}. Assim, $H(\beta_1)$ pode ser escrito como

$$H(\beta_1) = g[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + \beta_1(x_{ij} + 1) + u_{0j}\} + \tau_0] - g[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + \beta_1 x_{ij} + u_{0j}\} + \tau_0].$$

Agora, expandimos $H(\beta_1)$ em série de Taylor sobre o valor $\beta_1 = 0$, $\beta_1^* = H(\beta_1) \approx \beta_1 H'(0)$,

$$H'(\beta_1) = g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + \beta_1(x_{ij} + 1) + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + \beta_1(x_{ij} + 1) + u_{0j}\}(x_{ij} + 1)$$

$$-g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + \beta_1 x_{ij} + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + \beta_1 x_{ij} + u_{0j}\}x_{ij}.$$

Substituindo pelo valor $\beta_1 = 0$, tem-se que

$$\begin{aligned} H'(0) = {} & g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + u_{0j}\}(x_{ij} + 1) \\ & - g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + u_{0j}\}x_{ij} \\ = {} & g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + u_{0j}\}x_{ij} \\ & + g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + u_{0j}\} \\ & - g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + u_{0j}\}x_{ij} \\ = {} & g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + u_{0j}\}. \end{aligned}$$

Assim,

$$H'(0) = g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0](1 - \tau_1 - \tau_0)g^{-1'}\{\beta_0 + u_{0j}\}.$$

Utilizando a relação $g^{-1'}\{\beta_0 + u_{0j}\} = \frac{1}{g'[g^{-1}\{\beta_0 + u_{0j}\}]}$ tem-se

$$H'(0) = \frac{(1 - \tau_1 - \tau_0)g'[(1 - \tau_1 - \tau_0)g^{-1}\{\beta_0 + u_{0j}\} + \tau_0]}{g'[g^{-1}\{\beta_0 + u_{0j}\}]},$$

logo a expressão geral para avaliar a aproximação da relação entre β_1^* e β_1 é dada por

$$\beta_1^* \approx \beta_1 H'(0) = \beta_1 \frac{(1-\tau_1-\tau_0)g'[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0]}{g'[g^{-1}\{\beta_0+u_{0j}\}]},$$

fica provada a expressão (4.16).

A.3 - Prova da expressão (4.17) que avalia aproximadamente a relação entre β_1^* e β_1 utilizando a função de ligação logit. Para avaliar essa relação considera-se a função logit na equação (4.16). A equação (4.16) é dada da seguinte forma

$$\beta_1^* \approx \beta_1 H'(0) = \beta_1 \frac{(1-\tau_1-\tau_0)g'[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0]}{g'[g^{-1}\{\beta_0+u_{0j}\}]}.$$

Considerando a função de ligação logit tem-se a seguinte relação : $g(\mu) = \log\{\frac{\mu}{1-\mu}\}$ e $g'(\mu) = \frac{1}{\mu(1-\mu)}$. Asssim, temos que

$$\begin{aligned}
g'[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0] &= \frac{1}{[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0][1-(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}-\tau_0]} \\
&= \frac{1}{[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0][1+\tau_1-\tau_1-(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}-\tau_0]} \\
&= \frac{1}{[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0][(1-\tau_1-\tau_0)-(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_1]} \\
&= \frac{1}{[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0][(1-\tau_1-\tau_0)(1-g^{-1}\{\beta_0+u_{0j}\})+\tau_1]}
\end{aligned}$$

e

$$g'[g^{-1}\{\beta_0+u_{0j}\}] = \frac{1}{g^{-1}\{\beta_0+u_{0j}\}[1-g^{-1}\{\beta_0+u_{0j}\}]},$$

logo a expressão que avalia aproximadamente a relação entre β_1^* e β_1 utilizando a função de ligação logit é dada por

$$\beta_1^* \approx \beta_1 H'(0) = \beta_1 \frac{(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}[1-g^{-1}\{\beta_0+u_{0j}\}]}{[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0][(1-\tau_1-\tau_0)(1-g^{-1}\{\beta_0+u_{0j}\})+\tau_1]},$$

fica provada a expressão (4.17).

A.4 - Prova da expressão (4.18) que avalia aproximadamente a relação entre β_1^* e β_1 utilizando a função de ligação probit. Para avaliar essa relação considera-se a função probit na equação (4.16).

Utilizando a relação $g^{-1'}\{\beta_0 + u_{0j}\} = \frac{1}{g'[g^{-1}\{\beta_0+u_{0j}\}]}$ e a função de ligação probit tem-se que

$g(\mu) = \Phi^{-1}(\mu)$ e $g'(\mu) = \Phi^{-1'}(\mu) = \frac{1}{\phi\{\Phi^{-1}(\mu)\}}$, onde $\Phi(.)$ é a densidade Normal padrão acumulada e $\phi\{.\}$ é a densidade Normal padronizada. Asssim, temos que

$$g'[(1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0] = \frac{1}{\phi[\Phi^{-1}((1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0)]}$$

e

$$g'[g^{-1}\{\beta_0+u_{0j}\}] = \frac{1}{\phi[\Phi^{-1}(g^{-1}\{\beta_0+u_{0j}\})]} = \frac{1}{\phi[\beta_0+u_{0j}]},$$

onde $(\Phi^{-1}g^{-1})$ é a identidade, logo a expressão que avalia aproximadamente a relação entre β_1^* e β_1 utilizando a função de ligação probit é dada por

$$\beta_1^* \approx \beta_1 H'(0) = \beta_1 \frac{(1-\tau_1-\tau_0)\phi[\beta_0+u_{0j}]}{\phi[\Phi^{-1}((1-\tau_1-\tau_0)g^{-1}\{\beta_0+u_{0j}\}+\tau_0)]},$$

fica provada a expressão (4.18).

B. Quadros dos estudos de simulações

B.1 - Quadros do primeiro estudo de simulação.

Quadro 1a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1			σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Est.	
5	0,05	0,05	−0,722 (0,079)	0,278	0,719 (0,087)	−0,281	0,739	0,053 (0,079)	827
		0,10	−0,805 (0,074)	0,195	0,696 (0,077)	−0,304	0,722	0,055 (0,081)	1000
		0,15	−0,879 (0,074)	0,121	0,665 (0,078)	−0,335	0,696	0,050 (0,077)	1000
		0,20	−0,958 (0,077)	0,042	0,642 (0,083)	−0,358	0,673	0,057 (0,089)	1000
	0,10	0,05	−0,553 (0,067)	0,447	0,648 (0,078)	−0,352	0,673	0,042 (0,069)	1000
		0,10	−0,624 (0,073)	0,376	0,608 (0,076)	−0,392	0,639	0,044 (0,067)	886
		0,15	−0,698 (0,071)	0,302	0,580 (0,073)	−0,420	0,614	0,048 (0,076)	800
		0,20	−0,777 (0,106)	0,223	0,567 (0,100)	−0,434	0,586	0,045 (0,060)	**699**

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de convergência pelo procedimento `glmmPQL`
(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 1b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
5	0,15	0,05	$-0,394$	0,606	0,583	$-0,417$	0,611	$-0,389$	0,039	1000
			(0,065)		(0,075)				(0,062)	
		0,10	$-0,461$	0,539	0,548	$-0,452$	0,580	$-0,420$	0,043	1000
			(0,070)		(0,070)				(0,064)	
		0,15	$-0,534$	0,466	0,515	$-0,485$	0,551	$-0,449$	0,043	1000
			(0,068)		(0,073)				(0,063)	
		0,20	$-0,602$	0,398	0,481	$-0,519$	0,521	$-0,480$	0,044	825
			(0,073)		(0,075)				(0,067)	
	0,20	0,05	$-0,244$	0,756	0,525	$-0,475$	0,560	$-0,440$	0,040	1000
			(0,064)		(0,072)				(0,061)	
		0,10	$-0,309$	0,691	0,493	$-0,507$	0,526	$-0,475$	0,040	1000
			(0,068)		(0,069)				(0,064)	
		0,15	$-0,375$	0,625	0,456	$-0,544$	0,495	$-0,505$	0,041	1000
			(0,067)		(0,071)				(0,062)	
		0,20	$-0,447$	0,553	0,426	$-0,575$	0,464	$-0,536$	0,039	860
			(0,065)		(0,071)				(0,061)	

(*) Prob. erro = probabilidade do erro; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 2a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
10	0,05	0,05	$-0,726$	0,274	0,719	$-0,281$	0,737	$-0,263$	0,018	868
			(0,063)		(0,066)				(0,030)	
		0,10	$-0,803$	0,197	0,692	$-0,308$	0,718	$-0,282$	0,022	1000
			(0,050)		(0,056)				(0,033)	
		0,15	$-0,869$	0,131	0,661	$-0,339$	0,692	$-0,308$	0,022	891
			(0,080)		(0,072)				(0,032)	
		0,20	$-0,954$	0,046	0,634	$-0,366$	0,671	$-0,329$	0,024	844
			(0,064)		(0,059)				(0,036)	
	0,10	0,05	$-0,555$	0,445	0,641	$-0,359$	0,664	$-0,336$	0,016	886
			(0,051)		(0,059)				(0,025)	
		0,10	$-0,623$	0,377	0,610	$-0,390$	0,636	$-0,364$	0,020	1000
			(0,049)		(0,052)				(0,030)	
		0,15	$-0,697$	0,303	0,579	$-0,421$	0,609	$-0,391$	0,021	895
			(0,058)		(0,060)				(0,030)	
		0,20	$-0,770$	0,230	0,547	$-0,453$	0,583	$-0,417$	0,020	1000
			(0,048)		(0,051)				(0,030)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 2b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
10	0,15	0,05	$-0,391$	0,609	0,580	$-0,421$	0,603	$-0,397$	0,018	1000
			(0,048)		(0,051)				(0,028)	
		0,10	$-0,460$	0,540	0,544	$-0,456$	0,572	$-0,428$	0,019	1000
			(0,047)		(0,051)				(0,030)	
		0,15	$-0,529$	0,471	0,512	$-0,489$	0,546	$-0,454$	0,018	1000
			(0,047)		(0,051)				(0,028)	
		0,20	$-0,602$	0,398	0,483	$-0,518$	0,515	$-0,487$	0,019	1000
			(0,048)		(0,049)				(0,029)	
	0,20	0,05	$-0,243$	0,757	0,529	$-0,471$	0,550	$-0,450$	0,017	1000
			(0,046)		(0,051)				(0,027)	
		0,10	$-0,307$	0,693	0,492	$-0,508$	0,519	$-0,481$	0,018	888
			(0,051)		(0,056)				(0,027)	
		0,15	$-0,529$	0,471	0,462	$-0,538$	0,507	$-0,493$	0,018	1000
			(0,047)		(0,050)				(0,028)	
		0,20	$-0,445$	0,556	0,423	$-0,577$	0,457	$-0,543$	0,018	888
			(0,049)		(0,052)				(0,029)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 3a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
20	0,05	0,05	$-0,728$	0,272	0,721	$-0,279$	0,737	$-0,263$	0,010	1000
			(0,036)		(0,040)				(0,015)	
		0,10	$-0,800$	0,200	0,690	$-0,310$	0,714	$-0,286$	0,010	1000
			(0,037)		(0,039)				(0,016)	
		0,15	$-0,877$	0,123	0,646	$-0,340$	0,692	$-0,308$	0,010	1000
			(0,037)		(0,039)				(0,015)	
		0,20	$-0,956$	0,044	0,634	$-0,366$	0,669	$-0,331$	0,012	1000
			(0,037)		(0,039)				(0,017)	
	0,10	0,05	$-0,553$	0,447	0,641	$-0,359$	0,663	$-0,337$	0,010	1000
			(0,035)		(0,038)				(0,015)	
		0,10	$-0,624$	0,376	0,611	$-0,389$	0,635	$-0,365$	0,009	1000
			(0,034)		(0,037)				(0,014)	
		0,15	$-0,696$	0,304	0,581	$-0,420$	0,608	$-0,392$	0,010	1000
			(0,035)		(0,037)				(0,015)	
		0,20	$-0,769$	0,231	0,550	$-0,450$	0,582	$-0,418$	0,010	1000
			(0,036)		(0,038)				(0,014)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 3b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
20	0,15	0,05	$-0,393$	0,607	0,578	$-0,422$	0,601	$-0,399$	0,009	1000
			(0,034)		(0,036)				(0,013)	
		0,10	$-0,462$	0,538	0,545	$-0,455$	0,573	$-0,427$	0,009	1000
			(0,034)		(0,035)				(0,013)	
		0,15	$-0,533$	0,469	0,514	$-0,487$	0,544	$-0,457$	0,008	1000
			(0,034)		(0,036)				(0,013)	
		0,20	$-0,604$	0,396	0,482	$-0,518$	0,514	$-0,487$	0,009	1000
			(0,034)		(0,036)				(0,014)	
	0,20	0,05	$-0,242$	0,758	0,528	$-0,473$	0,549	$-0,451$	0,009	1000
			(0,034)		(0,036)				(0,013)	
		0,10	$-0,310$	0,691	0,492	$-0,509$	0,516	$-0,484$	0,008	1000
			(0,033)		(0,035)				(0,012)	
		0,15	$-0,376$	0,624	0,460	$-0,540$	0,486	$-0,514$	0,009	1000
			(0,033)		(0,036)				(0,013)	
		0,20	$-0,447$	0,554	0,425	$-0,576$	0,456	$-0,544$	0,009	1000
			(0,033)		(0,036)				(0,013)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 4a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
50	0,05	0,05	$-0,726$	0,274	0,719	$-0,281$	0,735	$-0,265$	0,004	1000
			(0,023)		(0,025)				(0,006)	
		0,10	$-0,803$	0,197	0,691	$-0,309$	0,714	$-0,287$	0,004	1000
			(0,023)		(0,025)				(0,006)	
		0,15	$-0,877$	0,123	0,663	$-0,337$	0,691	$-0,309$	0,004	1000
			(0,023)		(0,024)				(0,006)	
		0,20	$-0,954$	0,046	0,634	$-0,366$	0,667	$-0,333$	0,004	1000
			(0,023)		(0,024)				(0,006)	
	0,10	0,05	$-0,553$	0,447	0,640	$-0,360$	0,661	$-0,339$	0,003	1000
			(0,022)		(0,023)				(0,005)	
		0,10	$-0,625$	0,375	0,609	$-0,391$	0,635	$-0,365$	0,003	1000
			(0,022)		(0,024)				(0,005)	
		0,15	$-0,695$	0,305	0,579	$-0,421$	0,608	$-0,392$	0,004	1000
			(0,022)		(0,023)				(0,006)	
		0,20	$-0,769$	0,231	0,549	$-0,451$	0,583	$-0,417$	0,004	1000
			(0,023)		(0,023)				(0,006)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 4b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
50	0,15	0,05	$-0,392$	0,608	0,578	$-0,422$	0,600	$-0,400$	0,003	1000
			(0,021)		(0,022)				(0,006)	
		0,10	$-0,461$	0,539	0,544	$-0,456$	0,571	$-0,429$	0,004	1000
			(0,021)		(0,023)				(0,005)	
		0,15	$-0,530$	0,469	0,512	$-0,488$	0,541	$-0,459$	0,004	1000
			(0,021)		(0,023)				(0,005)	
		0,20	$-0,601$	0,400	0,481	$-0,519$	0,512	$-0,488$	0,004	1000
			(0,021)		(0,023)				(0,005)	
	0,20	0,05	$-0,242$	0,758	0,526	$-0,474$	0,548	$-0,452$	0,003	1000
			(0,022)		(0,022)				(0,005)	
		0,10	$-0,308$	0,693	0,490	$-0,510$	0,516	$-0,484$	0,003	1000
			(0,021)		(0,021)				(0,005)	
		0,15	$-0,376$	0,624	0,456	$-0,544$	0,484	$-0,516$	0,003	1000
			(0,021)		(0,021)				(0,005)	
		0,20	$-0,445$	0,555	0,424	$-0,576$	0,454	$-0,546$	0,003	1000
			(0,021)		(0,022)				(0,005)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 5a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
100	0,05	0,05	$-0,726$	0,274	0,719	$-0,281$	0,735	$-0,265$	0,002	1000
			(0,016)		(0,017)				(0,003)	
		0,10	$-0,801$	0,199	0,690	$-0,310$	0,714	$-0,286$	0,002	1000
			(0,016)		(0,017)				(0,003)	
		0,15	$-0,878$	0,122	0,662	$-0,338$	0,691	$-0,309$	0,002	1000
			(0,017)		(0,018)				(0,003)	
		0,20	$-0,955$	0,045	0,634	$-0,366$	0,668	$-0,332$	0,002	1000
			(0,017)		(0,018)				(0,003)	
	0,10	0,05	$-0,552$	0,448	0,640	$-0,360$	0,660	$-0,340$	0,002	1000
			(0,015)		(0,017)				(0,003)	
		0,10	$-0,623$	0,377	0,610	$-0,390$	0,634	$-0,366$	0,002	1000
			(0,016)		(0,017)				(0,003)	
		0,15	$-0,696$	0,304	0,580	$-0,421$	0,608	$-0,392$	0,002	1000
			(0,016)		(0,017)				(0,003)	
		0,20	$-0,769$	0,231	0,549	$-0,451$	0,582	$-0,418$	0,002	1000
			(0,016)		(0,016)				(0,003)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 5b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
100	0,15	0,05	$-0,392$	0,608	0,579	$-0,422$	0,599	$-0,401$	0,002	1000
			(0,016)		(0,016)				(0,003)	
		0,10	$-0,462$	0,539	0,545	$-0,455$	0,569	$-0,431$	0,002	1000
			(0,015)		(0,016)				(0,003)	
		0,15	$-0,530$	0,470	0,513	$-0,487$	0,541	$-0,458$	0,002	1000
			(0,016)		(0,016)				(0,003)	
		0,20	$-0,601$	0,399	0,481	$-0,519$	0,512	$-0,488$	0,002	1000
			(0,016)		(0,016)				(0,002)	
	0,20	0,05	$-0,240$	0,760	0,525	$-0,475$	0,547	$-0,453$	0,002	1000
			(0,015)		(0,016)				(0,002)	
		0,10	$-0,309$	0,692	0,491	$-0,509$	0,516	$-0,485$	0,002	1000
			(0,015)		(0,016)				(0,002)	
		0,15	$-0,377$	0,623	0,458	$-0,542$	0,484	$-0,516$	0,002	1000
			(0,015)		(0,016)				(0,003)	
		0,20	$-0,445$	0,555	0,425	$-0,575$	0,454	$-0,547$	0,002	1000
			(0,014)		(0,016)				(0,003)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.17).

Quadro 6a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
5	0,05	0,05	$-0,591$	0,409	0,578	$-0,423$	0,617	$-0,383$	0,017	856
			(0,044)		(0,054)				(0,029)	
		0,10	$-0,633$	0,367	0,552	$-0,448$	0,598	$-0,402$	0,017	825
			(0,044)		(0,052)				(0,028)	
		0,15	$-0,676$	0,324	0,530	$-0,471$	0,573	$-0,427$	0,019	853
			(0,051)		(0,055)				(0,035)	
		0,20	$-0,719$	0,281	0,507	$-0,493$	0,550	$-0,450$	0,021	1000
			(0,048)		(0,053)				(0,035)	
	0,10	0,05	$-0,462$	0,538	0,501	$-0,499$	0,545	$-0,455$	0,015	886
			(0,049)		(0,053)				(0,027)	
		0,10	$-0,498$	0,502	0,477	$-0,523$	0,526	$-0,474$	0,015	818
			(0,052)		(0,058)				(0,026)	
		0,15	$-0,544$	0,456	0,451	$-0,549$	0,505	$-0,495$	0,021	856
			(0,054)		(0,057)				(0,033)	
		0,20	$-0,580$	0,420	0,432	$-0,568$	0,474	$-0,526$	0,018	840
			(0,040)		(0,050)				(0,032)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 6b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
5	0,15	0,05	−0,345	0,655	0,448	−0,552	0,486	−0,514	0,013	888
			(0,049)		(0,053)				(0,025)	
		0,10	−0,387	0,614	0,424	−0,576	0,464	−0,536	0,012	1000
			(0,043)		(0,047)				(0,022)	
		0,15	−0,413	0,587	0,399	−0,602	0,442	−0,558	0,013	857
			(0,068)		(0,070)				(0,022)	
		0,20	−0,462	0,538	0,369	−0,631	0,418	−0,582	0,014	821
			(0,050)		(0,049)				(0,025)	
	0,20	0,05	−0,242	0,758	0,405	−0,595	0,432	−0,568	0,014	1000
			(0,042)		(0,044)				(0,023)	
		0,10	−0,275	0,725	0,364	−0,636	0,403	−0,597	0,020	728
			(0,070)		(0,082)				(0,030)	
		0,15	−0,315	0,685	0,351	−0,649	0,386	−0,614	0,014	806
			(0,044)		(0,046)				(0,024)	
		0,20	−0,331	0,669	0,291	−0,709	0,368	−0,632	0,018	**613**
			(0,106)		(0,092)				(0,034)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 7a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
10	0,05	0,05	$-0,567$	0,433	0,558	$-0,442$	0,604	$-0,396$	0,004	**631**
			(0,108)		(0,109)				(0,008)	
		0,10	$-0,633$	0,367	0,549	$-0,451$	0,589	$-0,411$	0,006	1000
			(0,035)		(0,036)				(0,012)	
		0,15	$-0,673$	0,327	0,525	$-0,476$	0,566	$-0,434$	0,008	1000
			(0,033)		(0,038)				(0,014)	
		0,20	$-0,717$	0,283	0,502	$-0,498$	0,542	$-0,458$	0,007	1000
			(0,034)		(0,035)				(0,013)	
	0,10	0,05	$-0,461$	0,539	0,502	$-0,498$	0,542	$-0,458$	0,005	1000
			(0,030)		(0,035)				(0,0115)	
		0,10	$-0,501$	0,500	0,475	$-0,525$	0,521	$-0,480$	0,006	1000
			(0,030)		(0,033)				(0,011)	
		0,15	$-0,537$	0,463	0,452	$-0,548$	0,496	$-0,504$	0,006	1000
			(0,031)		(0,034)				(0,011)	
		0,20	$-0,576$	0,424	0,424	$-0,576$	0,474	$-0,526$	0,006	1000
			(0,031)		(0,032)				(0,011)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 7b.Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
10	0,15	0,05	$-0,346$	0,654	0,448	$-0,552$	0,481	$-0,520$	0,006	1000
			(0,029)		(0,032)				(0,011)	
		0,10	$-0,384$	0,616	0,420	$-0,580$	0,459	$-0,541$	0,006	1000
			(0,031)		(0,032)				(0,011)	
		0,15	$-0,422$	0,578	0,393	$-0,607$	0,436	$-0,564$	0,006	1000
			(0,030)		(0,033)				(0,011)	
		0,20	$-0,460$	0,541	0,369	$-0,631$	0,413	$-0,587$	0,007	1000
			(0,030)		(0,033)				(0,012)	
	0,20	0,05	$-0,241$	0,759	0,402	$-0,598$	0,428	$-0,572$	0,006	1000
			(0,028)		(0,032)				(0,011)	
		0,10	$-0,275$	0,725	0,375	$-0,625$	0,406	$-0,594$	0,006	1000
			(0,028)		(0,032)				(0,011)	
		0,15	$-0,314$	0,686	0,349	$-0,651$	0,382	$-0,618$	0,006	1000
			(0,030)		(0,031)				(0,011)	
		0,20	$-0,352$	0,648	0,325	$-0,675$	0,361	$-0,640$	0,006	1000
			(0,029)		(0,032)				(0,011)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`
(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 8a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
20	0,05	0,05	$-0,592$	0,408	0,576	$-0,424$	0,609	$-0,391$	0,002	1000
			(0,024)		(0,026)				(0,005)	
		0,10	$-0,632$	0,368	0,550	$-0,450$	0,588	$-0,412$	0,003	1000
			(0,025)		(0,026)				(0,006)	
		0,15	$-0,673$	0,327	0,525	$-0,475$	0,534	$-0,466$	0,003	1000
			(0,024)		(0,024)				(0,005)	
		0,20	$-0,716$	0,285	0,499	$-0,501$	0,502	$-0,498$	0,003	1000
			(0,024)		(0,026)				(0,005)	
	0,10	0,05	$-0,460$	0,540	0,501	$-0,499$	0,539	$-0,461$	0,002	1000
			(0,022)		(0,024)				(0,005)	
		0,10	$-0,500$	0,500	0,476	$-0,524$	0,518	$-0,482$	0,002	1000
			(0,023)		(0,024)				(0,005)	
		0,15	$-0,538$	0,462	0,449	$-0,551$	0,495	$-0,505$	0,003	1000
			(0,023)		(0,024)				(0,005)	
		0,20	$-0,578$	0,422	0,423	$-0,577$	0,472	$-0,528$	0,003	1000
			(0,022)		(0,025)				(0,005)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 8b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
20	0,15	0,05	−0.346	0.654	0.446	−0.554	0.580	−0,420	0.003	1000
			(0.022)		(0.023)				(0.005)	
		0,10	−0.384	0.616	0.420	−0.580	0.457	−0,543	0.003	1000
			(0.021)		(0.024)				(0.005)	
		0,15	−0.422	0.578	0.394	−0.606	0.435	−0,565	0.003	1000
			(0.021)		(0.024)				(0.005)	
		0,20	−0.458	0.542	0.368	−0.632	0.411	−0,589	0.003	1000
			(0.020)		(0.022)				(0.005)	
	0,20	0,05	−0,239	0,761	0,402	−0,598	0,425	−0,575	0,003	1000
			(0,021)		(0,022)				(0,005)	
		0,10	−0,276	0,724	0,376	−0,624	0,404	−0,596	0,003	1000
			(0,021)		(0,022)				(0,005)	
		0,15	−0,313	0,687	0,348	−0,652	0,381	−0,619	0,003	1000
			(0,021)		(0,022)				(0,005)	
		0,20	−0,350	0,650	0,323	−0,677	0,358	−0,642	0,003	1000
			(0,021)		(0,022)				(0,005)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 9a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
50	0,05	0,05	−0,591	0,409	0,575	−0,425	0,608	−0,392	0,001	1000
			(0,015)		(0,016)				(0,002)	
		0,10	−0,632	0,368	0,548	−0,452	0,586	−0,414	0,001	1000
			(0,015)		(0,016)				(0,002)	
		0,15	−0,673	0,327	0,524	−0,476	0,564	−0,436	0,001	1000
			(0,014)		(0,016)				(0,002)	
		0,20	−0,714	0,286	0,499	−0,501	0,541	−0,459	0,001	1000
			(0,015)		(0,016)				(0,002)	
	0,10	0,05	−0,461	0,539	0,501	−0,499	0,538	−0,462	0,001	1000
			(0,014)		(0,016)				(0,002)	
		0,10	−0,500	0,501	0,474	−0,526	0,518	−0,482	0,001	1000
			(0,014)		(0,015)				(0,002)	
		0,15	−0,539	0,462	0,449	−0,551	0,495	−0,505	0,001	1000
			(0,014)		(0,016)				(0,002)	
		0,20	−0,577	0,423	0,424	−0,576	0,472	−0,528	0,001	1000
			(0,014)		(0,015)				(0,002)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 9b.Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
50	0,15	0,05	$-0,345$	0,655	0,446	$-0,554$	0,479	$-0,521$	0,001	1000
			(0,014)		(0,015)				(0,002)	
		0,10	$-0,383$	0,617	0,419	$-0,581$	0,457	$-0,543$	0,001	1000
			(0,013)		(0,015)				(0,002)	
		0,15	-0.421	0.579	0.394	-0.606	0.434	$-0,566$	0.001	1000
			(0.013)		(0.014)				(0.002)	
		0,20	$-0,458$	0,542	0,367	$-0,633$	0,411	$-0,589$	0,001	1000
			(0,014)		(0,014)				(0,002)	
	0,20	0,05	$-0,239$	0,761	0,403	$-0,598$	0,425	$-0,575$	0,001	1000
			(0,013)		(0,014)				(0,002)	
		0,10	$-0,277$	0,723	0,375	$-0,625$	0,403	$-0,597$	0,001	1000
			(0,013)		(0,014)				(0,002)	
		0,15	$-0,314$	0,686	0,349	$-0,651$	0,381	$-0,619$	0,001	1000
			(0,013)		(0,014)				(0,002)	
		0,20	$-0,351$	0,649	0,323	$-0,677$	0,358	$-0,642$	0,001	1000
			(0,013)		(0,014)				(0,002)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 10a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
100	0,05	0,05	$-0,591$	0,409	0,574	$-0,426$	0,608	$-0,392$	0,000	1000
			(0,010)		(0,011)				(0,001)	
		0,10	$-0,632$	0,368	0,548	$-0,452$	0,586	$-0,414$	0,000	1000
			(0,010)		(0,011)				(0,001)	
		0,15	$-0,672$	0,328	0,523	$-0,477$	0,549	$-0,451$	0,001	1000
			(0,011)		(0,012)				(0,001)	
		0,20	$-0,713$	0,287	0,499	$-0,501$	0,541	$-0,459$	0,001	1000
			(0,010)		(0,011)				(0,001)	
	0,10	0,05	$-0,460$	0,540	0,501	$-0,499$	0,539	$-0,461$	0,000	1000
			(0,010)		(0,011)				(0,001)	
		0,10	$-0,498$	0,502	0,474	$-0,526$	0,517	$-0,483$	0,001	1000
			(0,010)		(0,010)				(0,001)	
		0,15	$-0,538$	0,462	0,449	$-0,551$	0,494	$-0,506$	0,001	1000
			(0,010)		(0,011)				(0,001)	
		0,20	$-0,578$	0,422	0,424	$-0,576$	0,471	$-0,529$	0,001	1000
			(0,010)		(0,011)				(0,001)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 10b.Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0		β_1				σ^2_{u0}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	
100	0,15	0,05	$-0,346$	0,654	0,446	$-0,554$	0,478	$-0,522$	0,001	1000
			(0,010)		(0,010)				(0,001)	
		0,10	$-0,383$	0,617	0,419	$-0,581$	0,456	$-0,544$	0,000	1000
			(0,009)		(0,010)				(0,001)	
		0,15	$-0,420$	0,580	0,393	$-0,607$	0,434	$-0,566$	0,001	1000
			(0,009)		(0,010)				(0,001)	
		0,20	$-0,459$	0,541	0,368	$-0,632$	0,411	$-0,589$	0,001	1000
			(0,009)		(0,011)				(0,001)	
	0,20	0,05	$-0,240$	0,760	0,402	$-0,598$	0,424	$-0,575$	0,001	1000
			(0,009)		(0,010)				(0,001)	
		0,10	$-0,276$	0,724	0,375	$-0,625$	0,403	$-0,597$	0,001	1000
			(0,009)		(0,010)				(0,001)	
		0,15	$-0,314$	0,686	0,349	$-0,651$	0,380	$-0,620$	0,001	1000
			(0,009)		(0,010)				(0,001)	
		0,20	$-0,350$	0,650	0,324	$-0,676$	0,357	$-0,643$	0,001	1000
			(0,010)		(0,010)				(0,001)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (4.18).

Quadro 11a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
5	0,05	0,05	−0,678	0,322	0,676	−0,324	0,641	−0,359	0,705	−0,294	0,710	−0,290	0,052	0,092	700
			(0,066)		(0,134)				(0,095)				(0,056)	(0,152)	
		0,10	−0,691	0,309	0,524	−0,476	0,585	−0,415	0,636	−0,364	0,660	−0,340	0,058	0,072	730
			(0,216)		(0,180)				(0,190)				(0,108)	(0,111)	
		0,15	−0,824	0,176	0,514	−0,486	0,565	−0,435	0,606	−0,394	0,631	−0,369	0,131	0,116	740
			(0,075)		(0,088)				(0,061)				(0,124)	(0,084)	
		0,20	−0,933	0,067	0,509	−0,491	0,508	−0,492	0,571	−0,429	0,606	−0,394	0,083	0,092	700
			(0,085)		(0,065)				(0,042)				(0,139)	(0,115)	
	0,10	0,05	−0,549	0,451	0,550	−0,450	0,603	−0,397	0,594	−0,406	0,623	−0,377	0,064	0,072	715
			(0,303)		(0,303)				(0,357)				(0,063)	(0,067)	
		0,10	−0,577	0,423	0,512	−0,488	0,552	−0,448	0,560	−0,440	0,604	−0,396	0,065	0,104	747
			(0,074)		(0,077)				(0,079)				(0,066)	(0,104)	
		0,15	−0,589	0,411	0,436	−0,564	0,524	−0,476	0,505	−0,495	0,586	−0,414	0,080	0,094	717
			(0,167)		(0,143)				(0,154)				(0,078)	(0,069)	
		0,20	−0,687	0,313	0,422	−0,578	0,468	−0,532	0,493	−0,507	0,535	−0,465	0,095	0,075	714
			(0,210)		(0,144)				(0,153)				(0,137)	(0,076)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 11b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
5	0,15	0,05	−0, 325	0,699	0,439	−0, 561	0,506	−0, 494	0,505	−0, 495	0,569	−0, 431	0,012	0,046	704
			(0,038)		(0,050)				(0,048)				(0,013)	(0,035)	
		0,10	−0, 330	0,670	0,363	−0, 637	0,510	−0, 490	0,465	−0, 535	0,579	−0, 421	0,097	0,076	706
			(0,167)		(0,189)				(0,231)				(0,113)	(0,062)	
		0,15	−0, 502	0,498	0,355	−0, 645	0,467	−0, 533	0,384	−0, 616	0,545	−0, 455	0,035	0,119	708
			(0,060)		(0,058)				(0,085)				(0,041)	(0,113)	
		0,20	−0, 622	0,378	0,353	−0, 647	0,400	−0, 600	0,374	−0, 626	0,470	−0, 530	0,017	0,072	704
			(0,057)		(0,090)				(0,046)				(0,031)	(0,105)	
	0,20	0,05	−0, 173	0,827	0,437	−0, 564	0,501	−0, 499	0,500	−0, 500	0,541	−0, 459	0,032	0,093	723
			(0,060)		(0,097)				(0,063)				(0,048)	(0,086)	
		0,10	−0, 226	0,774	0,396	−0, 604	0,521	−0, 479	0,406	−0, 594	0,538	−0, 462	0,023	0,040	702
			(0,107)		(0,036)				(0,006)				(0,032)	(0,054)	
		0,15	−0, 299	0,701	0,393	−0, 607	0,435	−0, 565	0,385	−0, 615	0,460	−0, 540	0,051	0,095	711
			(0,135)		(0,144)				(0,119)				(0,069)	(0,103)	
		0,20	−0, 440	0,560	0,341	−0, 659	0,423	−0, 577	0,376	−0, 624	0,472	−0, 528	0,101	0,077	702
			(0,017)		(0,048)				(0,088)				(0,150)	(0,096)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 12a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
10	0,05	0,05	−0,587	0,413	0,558	−0,442	0,606	−0,394	0,608	−0,392	0,687	−0,313	0,033	0,040	712
			(0,189)		(0,189)				(0,197)				(0,051)	(0,046)	
		0,10	−0,747	0,253	0,557	−0,443	0,597	−0,403	0,605	−0,395	0,666	−0,334	0,034	0,055	739
			(0,051)		(0,042)				(0,050)				(0,046)	(0,052)	
		0,15	−0,806	0,194	0,523	−0,477	0,579	−0,421	0,580	−0,4196	0,604	−0,396	0,020	0,030	703
			(0,031)		(0,047)				(0,050)				(0,021)	(0,033)	
		0,20	−0,894	0,106	0,516	−0,484	0,543	−0,457	0,579	−0,421	0,613	−0,387	0,046	0,032	711
			(0,061)		(0,044)				(0,039)				(0,048)	(0,025)	
	0,10	0,05	−0,476	0,524	0,526	−0,474	0,552	−0,448	0,577	−0,423	0,609	−0,391	0,012	0,036	715
			(0,043)		(0,045)				(0,055)				(0,017)	(0,036)	
		0,10	−0,565	0,435	0,491	−0,509	0,535	−0,465	0,543	−0,457	0,582	−0,418	0,022	0,043	753
			(0,054)		(0,057)				(0,047)				(0,034)	(0,049)	
		0,15	−0,559	0,410	0,445	−0,555	0,509	−0,491	0,509	−0,491	0,579	−0,421	0,033	0,039	709
			(0,213)		(0,173)				(0,170)				(0,046)	(0,030)	
		0,20	−0,704	0,296	0,431	−0,569	0,489	−0,511	0,491	−0,509	0,550	−0,450	0,034	0,028	718
			(0,061)		(0,047)				(0,062)				(0,034)	(0,034)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 12b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
10	0,15	0,05	−0,301	0,699	0,439	−0,561	0,506	−0,494	0,505	−0,495	0,569	−0,431	0,012	0,046	710
			(0,038)		(0,050)				(0,048)				(0,013)	(0,035)	
		0,10	−0,415	0,585	0,431	−0,569	0,477	−0,523	0,489	−0,511	0,534	−0,466	0,038	0,052	709
			(0,041)		(0,051)				(0,054)				(0,037)	(0,050)	
		0,15	−0,463	0,537	0,390	−0,610	0,461	−0,539	0,442	−0,558	0,511	−0,489	0,012	0,031	731
			(0,100)		(0,084)				(0,094)				(0,017)	(0,037)	
		0,20	−0,527	0,473	0,366	−0,634	0,439	−0,561	0,407	−0,593	0,497	−0,503	0,023	0,046	710
			(0,193)		(0,137)				(0,148)				(0,033)	(0,047)	
	0,20	0,05	−0,188	0,812	0,446	−0,554	0,475	−0,525	0,489	−0,511	0,518	−0,482	0,013	0,024	734
			(0,054)		(0,044)				(0,040)				(0,023)	(0,036)	
		0,10	−0,199	0,801	0,386	−0,614	0,459	−0,541	0,480	−0,520	0,496	−0,504	0,016	0,035	705
			(0,041)		(0,020)				(0,036)				(0,026)	(0,045)	
		0,15	−0,336	0,664	0,384	−0,616	0,451	−0,549	0,387	−0,613	0,479	−0,521	0,018	0,025	705
			(0,029)		(0,073)				(0,049)				(0,025)	(0,028)	
		0,20	−0,387	0,613	0,350	−0,650	0,389	−0,611	0,367	−0,633	0,434	−0,566	0,026	0,028	745
			(0,043)		(0,046)				(0,045)				(0,037)	(0,032)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 13a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				$\hat{\sigma}_{u0}$	$\hat{\sigma}_{u1}$	con.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (5.3)	Viés (5.3)	Est.	Viés	Pred. (5.3)	Viés (5.3)	Pred.	Pred.	
15	0,05	0,05	−0,658	0,342	0,591	−0,409	0,622	−0,378	0,655	−0,345	0,673	−0,327	0,010	0,022	750
			(0,032)		(0,033)				(0,035)				(0,013)	(0,027)	
		0,10	−0,730	0,270	0,553	−0,447	0,587	−0,413	0,616	−0,384	0,651	−0,349	0,011	0,019	830
			(0,075)		(0,062)				(0,066)				(0,017)	(0,0200)	
		0,15	−0,811	0,189	0,533	−0,467	0,562	−0,438	0,596	−0,404	0,635	−0,365	0,010	0,017	758
			(0,035)		(0,039)				(0,037)				(0,015)	(0,021)	
		0,20	−0,905	0,095	0,508	−0,492	0,545	−0,455	0,566	−0,434	0,611	−0,389	0,013	0,021	749
			(0,037)		(0,036)				(0,039)				(0,017)	(0,022)	
	0,10	0,05	−0,446	0,554	0,488	−0,512	0,558	−0,442	0,543	−0,457	0,600	−0,400	0,011	0,020	715
			(0,128)		(0,137)				(0,155)				(0,015)	(0,022)	
		0,10	−0,558	0,442	0,487	−0,513	0,525	−0,475	0,521	−0,479	0,586	−0,414	0,011	0,025	715
			(0,041)		(0,036)				(0,035)				(0,017)	(0,018)	
		0,15	−0,620	0,380	0,459	−0,541	0,506	−0,494	0,503	−0,497	0,563	−0,437	0,011	0,025	732
			(0,118)		(0,094)				(0,099)				(0,013)	(0,019)	
		0,20	−0,708	0,292	0,440	−0,560	0,489	−0,511	0,487	−0,513	0,527	−0,473	0,008	0,015	722
			(0,031)		(0,043)				(0,031)				(0,013)	(0,015)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 13b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
20	0,15	0,05	−0,326	0,674	0,469	−0,531	0,506	−0,494	0,530	−0,470	0,559	−0,441	0,011	0,018	776
			(0,036)		(0,038)				(0,034)				(0,015)	(0,019)	
		0,10	−0,404	0,596	0,442	−0,558	0,491	−0,509	0,505	−0,495	0,512	−0,488	0,006	0,019	706
			(0,035)		(0,055)				(0,042)				(0,006)	(0,018)	
		0,15	−0,475	0,525	0,409	−0,591	0,454	−0,546	0,461	−0,539	0,503	−0,497	0,010	0,019	805
			(0,032)		(0,037)				(0,038)				(0,013)	(0,019)	
		0,20	−0,547	0,453	0,387	−0,613	0,426	−0,574	0,427	−0,573	0,476	−0,524	0,009	0,014	815
			(0,032)		(0,034)				(0,034)				(0,014)	(0,017)	
	0,20	0,05	−0,173	0,827	0,423	−0,577	0,457	−0,543	0,471	−0,529	0,506	−0,494	0,013	0,013	714
			(0,036)		(0,034)				(0,041)				(0,013)	(0,015)	
		0,10	−0,250	0,750	0,400	−0,600	0,436	−0,564	0,440	−0,560	0,483	−0,517	0,010	0,016	869
			(0,033)		(0,036)				(0,037)				(0,014)	(0,017)	
		0,15	−0,322	0,678	0,363	−0,637	0,406	−0,594	0,399	−0,601	0,445	−0,555	0,011	0,018	736
			(0,068)		(0,069)				(0,074)				(0,016)	(0,018)	
		0,20	−0,390	0,610	0,332	−0,668	0,382	−0,618	0,387	−0,613	0,428	−0,572	0,011	0,015	741
			(0,029)		(0,031)				(0,035)				(0,014)	(0,015)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 14a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
50	0,05	0,05	−0,654	0,346	0,582	−0,418	0,605	−0,395	0,660	−0,340	0,671	−0,329	0,005	0,013	724
			(0,025)		(0,022)				(0,020)				(0,008)	(0,013)	
		0,10	−0,728	0,272	0,548	−0,452	0,588	−0,412	0,616	−0,384	0,654	−0,346	0,005	0,007	812
			(0,074)		(0,056)				(0,064)				(0,007)	(0,009)	
		0,15	−0,819	0,181	0,532	−0,468	0,565	−0,435	0,590	−0,410	0,631	−0,369	0,005	0,008	922
			(0,024)		(0,026)				(0,024)				(0,006)	(0,009)	
		0,20	−0,867	0,133	0,487	−0,513	0,546	−0,454	0,540	−0,460	0,614	−0,386	0,006	0,009	728
			(0,171)		(0,098)				(0,108)				(0,007)	(0,010)	
	0,10	0,05	−0,472	0,528	0,512	−0,488	0,551	−0,449	0,566	−0,434	0,608	−0,392	0,004	0,008	752
			(0,069)		(0,077)				(0,084)				(0,005)	(0,010)	
		0,10	−0,560	0,440	0,492	−0,508	0,525	−0,475	0,547	−0,453	0,585	−0,415	0,004	0,008	899
			(0,022)		(0,024)				(0,024)				(0,006)	(0,007)	
		0,15	−0,637	0,363	0,463	−0,537	0,505	−0,495	0,520	−0,480	0,560	−0,440	0,004	0,008	756
			(0,024)		(0,024)				(0,023)				(0,006)	(0,009)	
		0,20	−0,712	0,288	0,438	−0,562	0,480	−0,520	0,489	−0,511	0,535	−0,465	0,004	0,008	955
			(0,043)		(0,032)				(0,036)				(0,005)	(0,008)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 14b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
50	0,15	0,05	−0, 322	0,678	0,470	−0, 530	0,499	−0, 501	0,524	−0, 476	0,557	−0, 443	0,004	0,008	936
			(0,027)		(0,035)				(0,036)				(0,006)	(0,008)	
		0,10	−0, 396	0,604	0,437	−0, 563	0,473	−0, 527	0,488	−0, 512	0,529	−0, 471	0,004	0,007	867
			(0,038)		(0,041)				(0,045)				(0,005)	(0,008)	
		0,15	−0, 474	0,526	0,413	−0, 587	0,450	−0, 550	0,458	−0, 542	0,501	−0, 499	0,004	0,007	828
			(0,022)		(0,024)				(0,023)				(0,005)	(0,007)	
		0,20	−0, 550	0,450	0,387	−0, 613	0,428	−0, 572	0,428	−0, 572	0,474	−0, 526	0,005	0,007	793
			(0,019)		(0,022)				(0,020)				(0,006)	(0,008)	
	0,20	0,05	−0, 178	0,822	0,427	−0, 573	0,459	−0, 541	0,482	−0, 518	0,510	−0, 490	0,005	0,007	758
			(0,022)		(0,020)				(0,022)				(0,007)	(0,007)	
		0,10	−0, 245	0,755	0,396	−0, 604	0,431	−0, 569	0,444	−0, 556	0,482	−0, 518	0,004	0,006	777
			(0,021)		(0,019)				(0,022)				(0,005)	(0,006)	
		0,15	−0, 318	0,682	0,370	−0, 630	0,403	−0, 597	0,409	−0, 591	0,446	−0, 554	0,005	0,007	729
			(0,023)		(0,019)				(0,017)				(0,006)	(0,008)	
		0,20	−0, 394	0,606	0,345	−0, 655	0,379	−0, 621	0,377	−0, 623	0,420	−0, 580	0,004	0,005	806
			(0,020)		(0,022)				(0,021)				(0,005)	(0,006)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 15a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
100	0,05	0,05	$-0,657$	0,343	0,584	$-0,416$	0,600	$-0,400$	0,651	$-0,349$	0,673	$-0,327$	0,003	0,004	743
			(0,014)		(0,018)				(0,022)				(0,004)	(0,004)	
		0,10	$-0,715$	0,285	0,539	$-0,461$	0,587	$-0,413$	0,598	$-0,402$	0,650	$-0,350$	0,002	0,003	734
			(0,127)		(0,098)				(0,107)				(0,003)	(0,005)	
		0,15	$-0,812$	0,188	0,528	$-0,472$	0,564	$-0,436$	0,592	$-0,408$	0,630	$-0,370$	0,002	0,004	724
			(0,015)		(0,018)				(0,012)				(0,003)	(0,005)	
		0,20	$-0,851$	0,149	0,476	$-0,524$	0,543	$-0,457$	0,531	$-0,469$	0,606	$-0,394$	0,002	0,005	718
			(0,213)		(0,120)				(0,134)				(0,002)	(0,005)	
	0,15	0,05	$-0,474$	0,526	0,509	$-0,491$	0,530	$-0,470$	0,567	$-0,433$	0,610	$-0,390$	0,002	0,005	751
			(0,069)		(0,074)				(0,083)				(0,003)	(0,005)	
		0,10	$-0,543$	0,457	0,483	$-0,517$	0,524	$-0,476$	0,536	$-0,464$	0,582	$-0,418$	0,003	0,003	737
			(0,093)		(0,084)				(0,092)				(0,004)	(0,003)	
		0,15	$-0,635$	0,365	0,463	$-0,537$	0,500	$-0,500$	0,515	$-0,485$	0,557	$-0,443$	0,002	0,003	795
			(0,014)		(0,014)				(0,016)				(0,002)	(0,004)	
		0,20	$-0,712$	0,288	0,440	$-0,560$	0,481	$-0,519$	0,487	$-0,513$	0,534	$-0,466$	0,002	0,004	927
			(0,017)		(0,016)				(0,017)				(0,003)	(0,004)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 15b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
100	0,15	0,05	−0,325	0,675	0,471	−0,529	0,499	−0,501	0,524	−0,476	0,555	−0,445	0,002	0,003	877
			(0,016)		(0,016)				(0,016)				(0,003)	(0,004)	
		0,10	−0,400	0,600	0,444	−0,556	0,474	−0,526	0,492	−0,508	0,526	−0,474	0,002	0,003	947
			(0,016)		(0,017)				(0,017)				(0,003)	(0,004)	
		0,15	−0,472	0,528	0,410	−0,590	0,450	−0,550	0,458	−0,542	0,500	−0,500	0,002	0,003	902
			(0,037)		(0,034)				(0,037)				(0,002)	(0,004)	
		0,20	−0,535	0,465	0,377	−0,623	0,422	−0,578	0,422	−0,578	0,472	−0,528	0,002	0,003	752
			(0,077)		(0,056)				(0,061)				(0,002)	(0,004)	
	0,20	0,05	−0,173	0,827	0,428	−0,572	0,457	−0,543	0,475	−0,525	0,509	−0,491	0,002	0,004	856
			(0,016)		(0,016)				(0,015)				(0,003)	(0,004)	
		0,10	−0,249	0,751	0,398	−0,602	0,431	−0,569	0,443	−0,557	0,479	−0,521	0,002	0,003	941
			(0,014)		(0,015)				(0,015)				(0,002)	(0,003)	
		0,15	−0,312	0,688	0,363	−0,637	0,403	−0,597	0,406	−0,594	0,448	−0,552	0,001	0,003	744
			(0,050)		(0,059)				(0,064)				(0,002)	(0,004)	
		0,20	−0,393	0,607	0,341	−0,659	0,378	−0,622	0,380	−0,620	0,421	−0,579	0,001	0,004	793
			(0,013)		(0,017)				(0,016)				(0,002)	(0,004)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.3).

Quadro 16a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
5	0,05	0,05	−0, 531	0,469	0,457	−0, 543	0,458	−0, 542	0,523	−0, 477	0,544	−0, 456	0,025	0,041	719
			(0,055)		(0,062)				(0,050)				(0,038)	(0,065)	
		0,10	−0, 579	0,421	0,415	−0, 585	0,456	−0, 544	0,465	−0, 535	0,532	−0, 468	0,015	0,033	720
			(0,051)		(0,045)				(0,054)				(0,026)	(0,046)	
		0,15	−0, 585	0,415	0,407	−0, 593	0,443	−0, 557	0,462	−0, 537	0,512	−0, 488	0,030	0,062	732
			(0,064)		(0,067)				(0,055)				(0,044)	(0,082)	
		0,20	−0, 632	0,368	0,352	−0, 648	0,419	−0, 581	0,431	−0, 569	0,498	−0, 502	0,022	0,029	704
			(0,049)		(0,046)				0,060				(0,026)	(0,035)	
	0,10	0,05	−0, 365	0,635	0,370	−0, 630	0,467	−0, 533	0,416	−0, 584	0,509	−0, 491	0,019	0,038	706
			(0,029)		(0,048)				(0,068)				(0,029)	(0,043)	
		0,10	−0, 350	0,650	0,356	−0, 644	0,456	−0, 544	0,366	−0, 634	0,467	−0, 533	0,000	0,015	705
			(0,199)		(0,169)				(0,206)				(0,000)	(0,022)	
		0,15	−0, 502	0,498	0,346	−0, 654	0,415	−0, 585	0,362	−0, 638	0,458	−0, 542	0,030	0,043	707
			(0,068)		(0,047)				(0,065)				(0,042)	(0,048)	
		0,20	−0, 486	0,514	0,322	−0, 678	0,394	−0, 606	0,357	−0, 643	0,452	−0, 548	0,043	0,028	705
			(0,043)		(0,020)				(0,036				(0,054)	(0,026)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 16b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
5	0,15	0,05	−0,306	0,694	0,349	−0,651	0,419	−0,581	0,408	−0,592	0,419	−0,581	0,025	0,030	706
			(0,039)		(0,063)				(0,017)				(0,033)	(0,051)	
		0,10	−0,321	0,679	0,318	−0,682	0,353	−0,647	0,368	−0,632	0,400	−0,600	0,020	0,020	712
			(0,042)		(0,055)				(0,029)				(0,034)	(0,040)	
		0,15	−0,375	0,625	0,308	−0,692	0,339	−0,661	0,355	−0,645	0,384	−0,616	0,022	0,023	709
			(0,055)		(0,049)				(0,048)				(0,037)	(0,046)	
		0,20	−0,403	0,597	0,296	−0,704	0,297	−0,703	0,267	−0,733	0,347	−0,653	0,007	0,000	704
			(0,053)		(0,047)				(0,048)				(0,015)	(0,000)	
	0,20	0,05	−0,253	0,747	0,311	−0,689	0,354	−0,646	0,343	−0,657	0,380	−0,620	0,004	0,028	702
			(0,018)		(0,052)				(0,068)				(0,005)	(0,039)	
		0,10	−0,184	0,816	0,303	−0,697	0,333	−0,667	0,320	−0,680	0,358	−0,642	0,004	0,024	703
			(0,160)		(0,194)				(0,192)				(0,007)	(0,041)	
		0,15	−0,257	0,743	0,282	−0,718	0,316	−0,684	0,310	−0,689	0,352	−0,648	0,013	0,011	704
			(0,035)		(0,068)				(0,047)				(0,025)	(0,023)	
		0,20	−0,302	0,698	0,242	−0,758	0,298	−0,702	0,296	−0,704	0,344	−0,656	0,015	0,024	711
			(0,026)		(0,040)				(0,059)				(0,020)	(0,031)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 17a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
10	0,05	0,05	−0,498	0,502	0,456	−0,544	0,497	−0,503	0,484	−0,516	0,535	−0,465	0,009	0,007	707
			(0,031)		(0,034)				(0,023)				(0,016)	(0,019)	
		0,10	−0,371	0,629	0,448	−0,552	0,456	−0,544	0,456	−0,544	0,479	−0,521	0,000	0,000	703
			(0,324)		(0,215)				(0,259)				(0,000)	(0,000)	
		0,15	−0,593	0,407	0,411	−0,589	0,440	−0,560	0,424	−0,576	0,469	−0,531	0,013	0,023	705
			(0,043)		(0,024)				(0,054)				(0,025)	(0,037)	
		0,20	−0,672	0,328	0,389	−0,611	0,416	−0,584	0,430	−0,570	0,450	−0,550	0,017	0,002	704
			(0,048)		(0,017)				(0,037)				(0,034)	(0,005)	
	0,10	0,05	−0,375	0,625	0,387	−0,613	0,412	−0,588	0,419	−0,581	0,472	−0,528	0,003	0,015	714
			(0,023)		(0,027)				(0,039)				(0,005)	(0,030)	
		0,10	−0,411	0,589	0,354	−0,646	0,392	−0,608	0,399	−0,601	0,461	−0,539	0,010	0,000	703
			(0,021)		(0,038)				(0,027)				(0,017)	(0,000)	
		0,15	−0,469	0,531	0,341	−0,659	0,352	−0,648	0,398	−0,602	0,455	−0,545	0,013	0,016	704
			(0,014)		(0,023)				(0,030)				(0,024)	(0,024)	
		0,20	−0,497	0,503	0,303	−0,697	0,356	−0,644	0,353	−0,647	0,405	−0,595	0,010	0,007	705
			(0,019)		(0,029)				(0,024)				(0,011)	(0,016)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 17b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
10	0,15	0,05	−0, 272	0,728	0,362	−0, 638	0,365	−0, 635	0,406	−0, 594	0,426	−0, 574	0,000	0,000	703
			(0,041)		(0,029)				(0,019)				(0,000)	(0,000)	
		0,10	−0, 337	0,663	0,317	−0, 683	0,334	−0, 666	0,394	−0, 606	0,387	−0, 613	0,000	0,006	703
			(0,003)		(0,013)				(0,055)				(0,000)	(0,011)	
		0,15	−0, 365	0,635	0,306	−0, 694	0,311	−0, 689	0,386	−0, 614	0,374	−0, 626	0,007	0,000	704
			(0,021)		(0,017)				(0,028)				(0,014)	(0,000)	
		0,20	−0, 389	0,611	0,287	−0, 713	0,303	−0, 697	0,352	−0, 648	0,364	−0, 636	0,005	0,000	704
			(0,012)		(0,017)				(0,025)				(0,009)	(0,000)	
	0,20	0,05	−0, 171	0,829	0,307	−0, 693	0,318	−0, 682	0,344	−0, 656	0,375	−0, 625	0,010	0,013	**609**
			(0,026)		(0,023)				(0,025)				(0,010)	(0,013)	
		0,10	−0, 239	0,761	0,299	−0, 701	0,307	−0, 693	0,333	−0, 667	0,345	−0, 655	0,015	0,012	**607**
			(0,015)		(0,021)				(0,029)				(0,016)	(0,011)	
		0,15	−0, 248	0,752	0,267	−0, 737	0,299	−0, 701	0,309	−0, 691	0,331	−0, 669	0,009	0,012	**608**
			(0,033)		(0,0323)				(0,034)				(0,014)	(0,013)	
		0,20	−0, 333	0,667	0,231	−0, 769	0,244	−0, 756	0,308	−0, 692	0,311	−0, 689	0,013	0,006	**604**
			(0,028)		(0,030)				(0,035)				(0,017)	(0,007)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 18a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
20	0,05	0,05	−0,259	0,741	0,419	−0,581	0,479	−0,521	0,258	−0,462	0,546	−0,454	0,000	0,000	702
			(0,366)		(0,310)				(0,336)				(0,000)	(0,000)	
		0,10	−0,538	0,462	0,386	−0,614	0,439	−0,561	0,481	−0,519	0,514	−0,486	0,000	0,000	706
			(0,026)		(0,009)				(0,024)				(0,000)	(0,000)	
		0,15	−0,589	0,411	0,383	−0,617	0,416	−0,584	0,442	−0,558	0,477	−0,523	0,001	0,014	703
			(0,017)		(0,033)				(0,042)				(0,001)	(0,013)	
		0,20	−0,563	0,437	0,317	−0,683	0,406	−0,594	0,364	−0,636	0,468	−0,532	0,005	0,003	708
			(0,229)		(0,130)				(0,148)				(0,009)	(0,008)	
	0,10	0,05	−0,364	0,636	0,385	−0,615	0,408	−0,592	0,427	−0,573	0,465	−0,535	0,006	0,010	702
			(0,002)		(0,007)				(0,005)				(0,009)	(0,014)	
		0,10	−0,427	0,573	0,374	−0,626	0,386	−0,614	0,405	−0,595	0,450	−0,550	0,001	0,008	704
			(0,032)		(0,021)				(0,027)				(0,001)	(0,016)	
		0,15	−0,447	0,553	0,356	−0,644	0,372	−0,628	0,375	−0,625	0,414	−0,586	0,001	0,007	703
			(0,017)		(0,007)				(0,018)				(0,002)	(0,012)	
		0,20	−0,518	0,482	0,314	−0,686	0,349	−0,651	0,357	−0,643	0,407	−0,593	0,005	0,009	706
			(0,012)		(0,007)				(0,016)				(0,007)	(0,014)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 18b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
20	0,15	0,05	−0, 274	0,726	0,344	−0, 656	0,357	−0, 643	0,394	−0, 606	0,419	−0, 581	0,000	0,000	702
			(0,047)		(0,033)				(0,006)				(0,000)	(0,000)	
		0,10	−0, 316	0,684	0,320	−0, 680	0,339	−0, 661	0,369	−0, 631	0,400	−0, 600	0,006	0,012	705
			(0,025)		(0,030)				(0,009)				(0,007)	(0,011)	
		0,15	−0, 353	0,647	0,300	−0, 700	0,338	−0, 662	0,336	−0, 664	0,382	−0, 618	0,006	0,005	707
			(0,033)		(0,0245)				(0,029)				(0,006)	(0,008)	
		0,20	−0, 430	0,570	0,299	−0, 701	0,306	−0, 694	0,330	−0, 700	0,369	−0, 631	0,000	0,011	702
			(0,007)		(0,021)				(0,018)				(0,000)	(0,016)	
	0,20	0,05	−0, 160	0,840	0,300	−0, 700	0,328	−0, 672	0,349	−0, 651	0,376	−0, 624	0,007	0,010	702
			(0,006)		(0,007)				(0,015)				(0,004)	(0,005)	
		0,10	−0, 215	0,785	0,283	−0, 717	0,303	−0, 697	0,322	−0, 678	0,349	−0, 651	0,008	0,006	703
			(0,008)		(0,015)				(0,017)				(0,012)	(0,007)	
		0,15	−0, 263	0,737	0,254	−0, 746	0,295	−0, 705	0,305	−0, 695	0,348	−0, 652	0,004	0,007	711
			(0,028)		(0,027)				(0,026)				(0,005)	(0,010)	
		0,20	−0, 283	0,717	0,244	−0, 756	0,270	−0, 730	0,275	−0, 725	0,316	−0, 316	0,005	0,009	707
			(0,011)		(0,024)				(0,018)				(0,005)	(0,007)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 19a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
50	0,15	0,05	−0,654	0,346	0,582	−0,418	0,605	−0,395	0,660	−0,340	0,671	−0,329	0,005	0,0128	724
			(0,025)		(0,022)				(0,020)				(0,008)	(0,013)	
		0,10	−0,728	0,272	0,548	−0,452	0,588	−0,412	0,616	−0,384	0,654	−0,346	0,005	0,007	812
			(0,074)		(0,056)				(0,064)				(0,007)	(0,009)	
		0,15	−0,819	0,181	0,532	−0,468	0,565	−0,435	0,590	−0,410	0,631	−0,369	0,005	0,008	922
			(0,024)		(0,026)				(0,024)				(0,006)	(0,009)	
		0,20	−0,867	0,133	0,487	−0,514	0,546	−0,454	0,540	−0,465	0,614	−0,386	0,006	0,009	728
			(0,171)		(0,098)				(0,108)				(0,007)	(0,010)	
	0,20	0,05	−0,472	0,528	0,512	−0,488	0,551	−0,449	0,566	−0,434	0,608	−0,392	0,004	0,008	752
			(0,069)		(0,077)				(0,084)				(0,005)	(0,010)	
		0,10	−0,560	0,440	0,492	−0,508	0,525	−0,475	0,547	−0,453	0,585	−0,415	0,004	0,008	899
			(0,022)		(0,024)				(0,024)				(0,006)	(0,007)	
		0,15	−0,637	0,363	0,463	−0,537	0,505	−0,495	0,520	−0,480	0,560	−0,440	0,004	0,008	756
			(0,024)		(0,024)				(0,023)				(0,006)	(0,009)	
		0,20	−0,712	0,288	0,438	−0,562	0,480	−0,520	0,489	−0,511	0,535	−0,465	0,004	0,008	955
			(0,043)		(0,032)				(0.036)				(0,005)	(0,008)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 19b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
50	0,15	0,05	−0,322	0,678	0,470	−0,530	0,499	−0,501	0,524	−0,476	0,557	−0,443	0,004	0,008	936
			(0,027)		(0,034)				(0,036)				(0,006)	(0,008)	
		0,10	−0,396	0,604	0,437	−0,563	0,473	−0,527	0,488	−0,512	0,529	−0,471	0,004	0,007	867
			(0,038)		(0,041)				(0,045)				(0,005)	(0,008)	
		0,15	−0,474	0,526	0,413	−0,587	0,450	−0,550	0,458	−0,542	0,501	−0,499	0,004	0,007	828
			(0,022)		(0,024)				(0,023)				(0,005)	(0,007)	
		0,20	−0,550	0,450	0,387	−0,613	0,428	−0,572	0,428	−0,572	0,474	−0,526	0,005	0,007	793
			(0,019)		(0,022)				(0,020)				(0,006)	(0,008)	
	0,20	0,05	−0,178	0,822	0,427	−0,573	0,459	−0,541	0,482	−0,518	0,510	−0,490	0,005	0,007	758
			(0,022)		(0,020)				(0,022)				(0,007)	(0,007)	
		0,10	−0,245	0,755	0,396	−0,604	0,431	−0,569	0,444	−0,556	0,482	−0,518	0,004	0,006	777
			(0,021)		(0,019)				(0,022)				(0,005)	(0,006)	
		0,15	−0,318	0,682	0,370	−0,630	0,404	−0,596	0,409	−0,591	0,446	−0,554	0,005	0,007	729
			(0,023)		(0,019)				(0,017)				(0,006)	(0,008)	
		0,20	−0,394	0,606	0,345	−0,655	0,379	−0,621	0,377	−0,623	0,420	−0,580	0,004	0,005	806
			(0,020)		(0,022)				(0,021)				(0,005)	(0,006)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 20a. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
100	0,05	0,05	−0,497	0,503	0,415	−0,585	0,443	−0,557	0,477	−0,523	0,509	−0,491	0,000	0,003	702
			(0,005)		(0,033)				(0,027)				(0,001)	(0,005)	
		0,10	−0,551	0,449	0,393	−0,607	0,435	−0,565	0,474	−0,526	0,496	−0,504	0,000	0,000	702
			(0,000)		(0,009)				(0,003)				(0,000)	(0,000)	
		0,15	−0,583	0,417	0,381	−0,619	0,406	−0,594	0,449	−0,552	0,486	−0,514	0,000	0,000	702
			(0,015)		(0,011)				(0,010)				(0,000)	(0,000)	
		0,20	−0,630	0,370	0,350	−0,650	0,397	−0,603	0,426	−0,574	0,476	−0,524	0,000	0,000	703
			(0,012)		(0,011)				(0,007)				(0,000)	(0,000)	
	0,10	0,05	−0,379	0,621	0,372	−0,628	0,399	−0,601	0,430	−0,570	0,465	−0,535	0,000	0,001	840
			(0,016)		(0,017)				(0,018)				(0,001)	(0,002)	
		0,10	−0,422	0,578	0,352	−0,648	0,384	−0,616	0,405	−0,595	0,446	−0,554	0,000	0,0007	846
			(0,018)		(0,016)				(0,017)				(0,001)	(0,002)	
		0,15	−0,480	0,520	0,338	−0,662	0,375	−0,625	0,383	−0,617	0,432	−0,568	0,000	0,002	800
			(0,009)		(0,033)				(0,016)				(0,000)	(0,003)	
		0,20	−0,518	0,483	0,318	−0,682	0,352	−0,648	0,354	−0,646	0,398	−0,602	0,001	0,003	702
			(0,002)		(0,006)				(0,013)				(0,001)	(0,004)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 20b. Média observada, desvio padrão amostral (entre parênteses) e viés predito dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1.0$, $\beta_1 = 1.0$, $\beta_2 = 1.0$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1.0$.(*)

n	Prob. erro		β_0		β_1				β_2				σ^2_{u0}	σ^2_{u1}	Conv.
	τ_0	τ_1	Est.	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Viés	Pred. (**)	Viés	Est.	Est.	
100	0,15	0,05	−0,264	0,736	0,327	−0,673	0,356	−0,644	0,390	−0,610	0,424	−0,576	0,001	0,001	703
			(0,009)		(0,011)				(0,006)				(0,003)	(0,002)	
		0,10	−0,312	0,688	0,313	−0,687	0,340	−0,660	0,361	−0,639	0,397	−0,603	0,001	0,001	804
			(0,014)		(0,015)				(0,016)				(0,001)	(0,002)	
		0,15	−0,362	0,638	0,288	−0,712	0,329	−0,671	0,346	−0,654	0,381	−0,619	0,000	0,003	702
			(0,020)		(0,008)				(0,010)				(0,000)	(0,004)	
		0,20	−0,392	0,608	0,267	−0,733	0,303	−0,697	0,320	−0,680	0,350	−0,650	0,002	0,002	702
			(0,018)		(0,007)				(0,006)				(0,003)	(0,003)	
	0,20	0,05	−0,169	0,831	0,304	−0,696	0,318	−0,682	0,349	−0,651	0,371	−0,629	0,001	0,001	1000
			(0,010)		(0,010)				(0,010)				(0,001)	(0,002)	
		0,10	−0,211	0,789	0,282	−0,718	0,290	−0,710	0,324	−0,676	0,352	−0,648	0,001	0,001	960
			(0,013)		(0,016)				(0,018)				(0,001)	(0,002)	
		0,15	−0,237	0,763	0,242	−0,758	0,285	−0,715	0,286	−0,714	0,340	−0,660	0,000	0,001	717
			(0,061)		(0,063)				(0,074)				(0,001)	(0,002)	
		0,20	−0,292	0,708	0,239	−0,761	0,268	−0,732	0,274	−0,726	0,313	−0,687	0,001	0,001	840
			(0,026)		(0,023)				(0,025)				(0,001)	(0,002)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

(**) Pred. = predito pela expressão aproximada (5.4).

Quadro 21a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
5	0,05	0,05	$-1,092$	**1,168**	1,029	**0,033**	**0,035**	701
			$(-1,240; -0,919)$	(1,023;1,318)	(0,919;1,204)	(0,012;0,056)	(0,000;0,119)	
		0,10	$\mathbf{-1,114}$	**1,152**	1,048	0,036	0,090	798
			$(-1,353; -1,134)$	(1,022;1,389)	(0,882;1,294)	(0,013;0,066)	(0,001;0,199)	
		0,15	$-1,098$	**1,145**	1,060	0,037	0,135	770
			$(-1,332; -0,895)$	(0,952;1,322)	(0,868;1,278)	(0,011;0,056)	(0,045;0,310)	
		0,20	$\mathbf{-1,107}$	**1,127**	1,023	0,037	0,189	728
			$(-1,327; -0,899)$	(0,945;1,377)	(0,861;1,273)	(0,012;0,060)	(0,104;0,429)	
	0,10	0,05	$-1,045$	**1,171**	1,056	**0,080**	0,043	1000
			$(-1,226; -1,046)$	(1,023;1,344)	(0,908;1,236)	(0,043;0,117)	(0,000;0,129)	
		0,10	$-1,070$	**1,124**	1,020	**0,081**	0,092	1000
			$(-1,269; -1,072)$	(0,968;1,320)	(0,855;1,282)	(0,045;0,117)	(0,000;0,200)	
		0,15	$-1,073$	1,096	1,046	0,086	0,140	1000
			$(-1,304; -0,861)$	(0,914;1,333)	(0,844;1,289)	(0,051;0,119)	(0,035;0,298)	
		0,20	$-1,079$	1,093	1,048	**0,085**	0,179	792
			$(-1,357; -0,811)$	(0,885;1,374)	(0,818;1,354)	(0,047;0,114)	(0,095;0,411)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 21b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
5	0,15	0,05	$-1,012$	**1,151**	1,001	**0,126**	0,044	729
			$(-1,164; -0,751)$	(0,988;1,340)	(0,863;1,256)	(0,077;0,181)	(0,000;0,131)	
		0,10	$-1,013$	**1,135**	1,023	0,135	0,096	700
			$(-1,247; -0,797)$	(0,927;1,365)	(0,851;1,305)	(0,088;0,186)	(0,000;0,194)	
		0,15	$-1,079$	1,097	1,040	0,145	0,141	712
			$(-1,354; -0,832)$	(0,887;1,357)	(0,827;1,338)	(0,092;0,193)	(0,003;0,287)	
		0,20	$-1,090$	1,063	1,016	0,145	0,184	807
			$(-1,379; -0,787)$	(0,832;1,434)	(0,788;1,383)	(0,095;0,193)	(0,043;0,394)	
	0,20	0,05	$-1,052$	**1,186**	1,060	0,220	0,042	1000
			$(-1,229; -0,777)$	(1,001;1,399)	(0,897;1,297)	(0,156;0,281)	(0,000;0,129)	
		0,10	$-1,034$	**1,128**	1,052	**0,226**	0,097	812
			$(-1,288; -0,811)$	(0,923;1,415)	(0,848;1,325)	(0,162;0,281)	(0,000;0,205)	
		0,15	$-1,001$	**1,155**	1,092	**0,226**	0,135	782
			$(-1,332; -0,736)$	(0,879;1,438)	(0,836;1,350)	(0,145;0,280)	(0,000;0,321)	
		0,20	$-1,196$	1,096	1,057	**0,235**	0,185	709
			$(-1,507; -0,797)$	(0,820;1,559)	(0,787;1,642)	(0,159;0,277)	(0,000;0,408)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `glmmPQL`

Quadro 22a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
10	0,05	0,05	$-1,036$	1,034	1,008	0,043	0,049	732
			$(-1,128; -0,909)$	(0,979;1,285)	(0,911;1,157)	(0,021;0,054)	(0,000;0,108)	
		0,10	$-1,036$	1,015	1,001	0,046	0,101	770
			$(-1,171; -1,171)$	(0,899;1,227)	(0,893;1,145)	(0,021;0,053)	(0,032;0,183)	
		0,15	$-1,035$	1,021	1,001	0,047	0,151	835
			$(-1,192; -0,912)$	(0,969;1,202)	(0,881;1,149)	(0,022;0,054)	(0,088;0,257)	
		0,20	$-1,091$	1,035	1,016	0,040	0,200	1000
			$(-1,237; -0,933)$	(0,958;1,203)	(0,877;1,189)	(0,022;0,057)	(0,132;0,350)	
	0,10	0,05	$-1,006$	1,097	1,026	0,092	0,051	786
			$(-1,120; -0,870)$	(0,980;1,266)	(0,917;1,156)	(0,059;0,114)	(0,000;0,108)	
		0,10	$-0,961$	1,025	0,956	0,090	0,093	756
			$(-1,128; -0,843)$	(0,956;1,203)	(0,849;1,127)	(0,051;0,199)	(0,008;0,185)	
		0,15	$-1,020$	1,039	0,995	0,099	0,150	1000
			$(-1,172; -0,882)$	(0,930;1,231)	(0,861;1,183)	(0,062;0,125)	(0,068;0,247)	
		0,20	$-1,047$	1,038	1,015	0,094	0,201	1000
			$(-1,219; -0,877)$	(0,895;1,244)	(0,848;1,202)	(0,063;0,118)	(0,122;0,356)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 22b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
10	0,15	0,05	$-0,982$	1,040	1,020	0,150	0,044	1000
			$(-1,115; -0,850)$	(0,968;1,218)	(0,908;1,179)	(0,114;0,187)	(0,000;0,103)	
		0,10	$-0,996$	1,038	1,026	0,153	0,104	1000
			$(-1,150; -0,855)$	(0,957;1,220)	(0,872;1,168)	(0,118;0,185)	(0,023;0,170)	
		0,15	$-1,008$	1,036	1,016	0,154	0,151	1000
			$(-1,170; -0,843)$	(0,900;1,238)	(0,846;1,216)	(0,120;0,185)	(0,064;0,248)	
		0,20	$-1,033$	1,033	1,020	0,160	0,203	1000
			$(-1,207; -0,836)$	(0,879;1,259)	(0,830;1,241)	(0,120;0,195)	(0,133;0,305)	
	0,20	0,05	$-0,984$	1,040	1,030	0,203	0,042	1000
			$(-1,139; -0,838)$	(0,997;1,239)	(0,906;1,181)	(0,186;0,274)	(0,000;0,097)	
		0,10	$-1,003$	1,037	1,014	0,202	0,097	1000
			$(-1,173; -0,815)$	(0,928;1,242)	(0,860;1,202)	(0,188;0,274)	(0,011;0,168)	
		0,15	$-0,994$	1,027	1,000	0,204	0,163	718
			$(-1,180; -0,821)$	(0,873;1,257)	(0,840;1,222)	(0,189;0,271)	(0,064;0,240)	
		0,20	$-1,054$	1,020	1,038	0,205	0,202	1000
			$(-1,257; -0,835)$	(0,858;1,272)	(0,811;1,247)	(0,190;0,277)	(0,120;0,306)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 23a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
20	0,05	0,05	$-0,959$	1,034	1,013	0,042	0,049	1000
			$(-1,030; -0,888)$	(0,982;1,208)	(0,941;1,087)	(0,030;0,055)	(0,007;0,085)	
		0,10	$-0,978$	1,032	0,997	0,045	0,102	1000
			$(-1,058; -0,897)$	(0,980;1,168)	(0,922;1,101)	(0,031;0,057)	(0,051;0,149)	
		0,15	$-0,983$	1,027	1,011	0,045	0,148	1000
			$(-1,069; -0,898)$	(0,979;1,169)	(0,920;1,113)	(0,031;0,058)	(0,110;0,226)	
		0,20	$-1,002$	1,023	1,017	0,046	0,196	714
			$(-1,094; -0,888)$	(0,967;1,164)	(0,911;1,116)	(0,033;0,059)	(0,174;0,313)	
	0,10	0,05	$-0,944$	1,040	1,005	0,099	0,045	1000
			$(-1,032; -0,866)$	(1,033;1,211)	(0,928;1,109)	(0,078;0,120)	(0,003;0,082)	
		0,10	$-0,978$	1,037	1,003	0,100	0,094	1000
			$(-1,071; -0,876)$	(0,979;1,161)	(0,904;1,106)	(0,078;0,121)	(0,044;0,146)	
		0,15	$-0,988$	1,033	1,004	0,102	0,151	1000
			$(-1,079; -0,888)$	(0,940;1,155)	(0,893;1,119)	(0,082;0,121)	(0,100;0,219)	
		0,20	$-1,002$	1,036	1,008	0,108	0,203	754
			$(-1,113; -0,885)$	(0,948;1,171)	(0,936;1,149)	(0,089;0,123)	(0,173;0,273)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 23b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
20	0,15	0,05	$-0,953$	1,041	1,014	0,145	0,049	1000
			$(-1,051; -0,867)$	(0,986;1,200)	(0,934;1,116)	(0,139;0,191)	(0,006;0,087)	
		0,10	$-0,973$	1,039	0,999	0,146	0,102	1000
			$(-1,078; -0,875)$	(0,966;1,174)	(0,901;1,133)	(0,145;0,192)	(0,048;0,153)	
		0,15	$-1,008$	1,038	1,014	0,151	0,172	765
			$(-1,115; -0,901)$	(0,949;1,156)	(0,884;1,111)	(0,112;0,210)	(0,158;0,315)	
		0,20	$-0,988$	1,040	1,006	0,190	0,245	1000
			$(-1,120; -0,864)$	(0,904;1,186)	(0,875;1,157)	(0,144;0,203)	(0,138;0,192)	
	0,20	0,05	$-0,955$	1,030	1,005	0,198	0,046	1000
			$(-1,053; -0,849)$	(0,990;1,174)	(0,919;1,112)	(0,158;0,267)	(0,002;0,083)	
		0,10	$-0,978$	1,025	1,010	0,201	0,105	1000
			$(-1,091; -0,858)$	(0,934;1,174)	(0,894;1,142)	(0,141;0,224)	(0,049;0,154)	
		0,15	$-0,986$	1,024	1,006	0,204	0,166	1000
			$(-1,121; -0,860)$	(0,905;1,187)	(0,879;1,151)	(0,152;0,224)	(0,097;0,225)	
		0,20	$-1,005$	1,020	1,005	0,195	0,201	1000
			$(-1,140; -0,854)$	(0,868;1,195)	(0,850;1,170)	(0,171;0,276)	(0,153;0,313)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 24a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
50	0,05	0,05	$-0,906$	1,050	0,999	0,054	0,042	1000
			$(-0,972; -0,808)$	(0,990;1,146)	(0,948;1,052)	(0,043;0,066)	(0,021;0,064)	
		0,10	$-0,906$	1,038	0,998	0,053	0,094	1000
			$(-0,955; -0,828)$	(0,985;1,132)	(0,947;1,060)	(0,043;0,067)	(0,065;0,126)	
		0,15	$-0,909$	1,032	0,999	0,051	0,151	1000
			$(-0,969; -0,824)$	(0,971;1,127)	(0,944;1,060)	(0,044;0,068)	(0,113;0,191)	
		0,20	$-0,920$	1,025	0,998	0,051	0,201	1000
			$(-0,965; -0,835)$	(0,967;1,124)	(0,930;1,066)	(0,045;0,069)	(0,174;0,264)	
	0,10	0,05	$-0,900$	1,043	1,007	0,104	0,049	793
			$(-0,984; -0,836)$	0,981;1,135)	(0,952;1,067)	(0,062;0,136)	(0,025;0,073)	
		0,10	$-0,912$	1,028	1,002	0,103	0,097	1000
			$(-0,987; -0,854)$	(0,975;1,128)	(0,943;1,069)	(0,063;0,137)	(0,071;0,128)	
		0,15	$-0,924$	1,024	1,006	0,102	0,152	1000
			$(-0,992; -0,853)$	(0,964;1,132)	(0,938;1,086)	(0,61;0,138)	(0,121;0,204)	
		0,20	$-0,927$	1,017	0,996	0,103	0,202	1000
			$(-1,004; -0,858)$	(0,949;1,130)	(0,923;1,085)	(0,061;0,139)	(0,178;0,275)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 24b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
50	0,15	0,05	$-0,925$	1,038	1,000	0,153	0,047	1000
			$(-1,005; -0,850)$	(0,996;1,120)	(0,943;1,067)	(0,101;0,198)	(0,019;0,071)	
		0,10	$-0,936$	1,036	0,995	0,152	0,096	1000
			$(-1,015; -0,864)$	(0,972;1,108)	(0,930;1,071)	(0,106;0,199)	(0,065;0,130)	
		0,15	$-0,944$	1,036	1,004	0,151	0,150	1000
			$(-1,014; -0,874)$	(0,958;1,115)	(0,923;1,081)	(0,106;0,199)	(0,118;0,198)	
		0,20	$-0,946$	1,025	1,000	0,150	0,201	1000
			$(-1,023; -0,862)$	(0,945;1,117)	(0,923;1,089)	(0,109;0,200)	(0,185;0,284)	
	0,20	0,05	$-0,946$	1,034	1,006	0,201	0,047	1000
			$(-1,027; -0,866)$	(0,987;1,127)	(0,935;1,080)	(0,175;0,265)	(0,023;0,070)	
		0,10	$-0,957$	1,025	1,003	0,201	0,101	1000
			$(-1,033; -0,879)$	(0,952;1,122)	(0,925;1,088)	(0,177;0,265)	(0,069;0,132)	
		0,15	$-0,956$	1,028	1,004	0,202	0,165	1000
			$(-1,041; -0,877)$	(0,952;1,123)	(0,920;1,093)	(0,174;0,266)	(0,123;0,203)	
		0,20	$-0,959$	1,023	1,002	0,202	0,203	1000
			$(-1,054; -0,878)$	(0,927;1,126)	(0,910;1,096)	(0,175;0,266)	(0,186;0,284)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 25a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
100	0,05	0,05	$-1,000$	1,008	0,998	0,051	0,050	1000
			$(-1,041; -0,956)$	(0,973;1,048)	(0,962;1,037)	(0,044;0,059)	(0,031;0,070)	
		0,10	$-0,996$	1,010	1,002	0,051	0,100	1000
			$(-1,039; -0,956)$	(0,971;1,054)	(0,962;1,043)	(0,044;0,058)	(0,086;0,133)	
		0,15	$-1,011$	1,012	1,005	0,053	0,151	703
			$(-1,060; -0,962)$	(0,963;1,052)	(0,964;1,031)	(0,048;0,059)	(0,140;0,195)	
		0,20	$-0,999$	1,007	0,994	0,051	0,201	703
			$(-1,037; -0,953)$	(0,953;1,050)	(0,948;1,046)	(0,044;0,056)	(0,190;0,263)	
	0,10	0,05	$-0,994$	1,008	1,001	0,104	0,054	788
			$(-1,039; -0,952)$	(0,957;1,049)	(0,959;1,047)	(0,990;0,118)	(0,031;0,075)	
		0,10	$-0,991$	1,008	0,999	0,994	0,104	771
			$(-1,036; -0,953)$	(0,954;1,046)	(0,955;1,038)	(0,980;0,116)	(0,091;0,127)	
		0,15	$-0,988$	1,010	1,015	0,978	0,154	787
			$(-1,051; -0,956)$	(0,948;1,061)	(0,948;1,078)	(0,950;0,121)	(0,147;0,212)	
		0,20	$-0,987$	1,006	1,014	0,994	0,202	777
			$(-1,041; -0,956)$	(0,959;1,052)	(0,959;1,052)	(0,980;0,119)	(0,191;0,254)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 25b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
100	0,15	0,05	$-1,007$	1,029	1,009	0,154	0,057	770
			$(-1,061; -0,953)$	(0,978;1,054)	(0,977;1,051)	(0,143;0,161)	(0,037;0,074)	
		0,10	$-1,013$	1,018	1,017	0,152	0,109	787
			$(-1,069; -0,961)$	(0,975;1,071)	(0,967;1,053)	(0,143;0,158)	(0,088;0,140)	
		0,15	$-0,997$	1,009	1,009	0,151	0,176	789
			$(-1,051; -0,943)$	(0,968;1,047)	(0,956;1,066)	(0,143;0,185)	(0,134;0,168)	
		0,20	$-1,008$	1,017	1,009	0,153	0,256	754
			$(-1,048; -0,954)$	(0,956;1,065)	(0,965;1,081)	(0,140;0,150)	(0,190;0,226)	
	0,20	0,05	$-0,997$	1,007	1,003	0,201	0,053	1000
			$(-1,050; -0,945)$	(0,958;1,060)	(0,953;1,055)	(0,174;0,234)	(0,032;0,073)	
		0,10	$-1,002$	1,002	1,000	0,201	0,100	1000
			$(-1,056; -0,946)$	(0,949;1,055)	(0,952;1,053)	(0,174;0,232)	(0,080;0,121)	
		0,15	$-0,998$	1,009	1,000	0,204	0,148	1000
			$(-1,036; -0,956)$	(0,973;1,054)	(0,965;1,037)	(0,176;0,235)	(0,107;0,175)	
		0,20	$-1,001$	1,010	1,000	0,203	0,200	1000
			$(-1,041; -0,964)$	(0,971;1,047)	(0,964;1,034)	(0,176;0,234)	(0,175;0,244)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 26a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
5	0,05	0,05	$-1,054$	**1,277**	**1,106**	0,043	0,049	**604**
			$(-1,163; -0,987)$	(1,170;1,333)	(1,044;1,249)	(0,039;0,050)	(0,030;0,080)	
		0,10	$-1,072$	1,064	1,030	0,052	0,109	**611**
			$(-1,130; -0,929)$	(0,985;1,122)	(0,986;1,193)	(0,045;0,060)	(0,070;0,137)	
		0,15	$\mathbf{-1,120}$	**1,114**	**1,140**	0,053	0,160	**617**
			$(-1,183; -1,006)$	(0,986;1,236)	(0,984;1,190)	(0,047;0,067)	(0,127;0,230)	
		0,20	$-1,059$	**1,112**	**1,113**	0,048	0,206	**617**
			$(-1,173; -1,028)$	(0,901;1,167)	(0,900;1,220)	(0,041;0,057)	(0,162;0,243)	
	0,10	0,05	$-1,036$	1,085	1,071	0,094	0,044	**610**
			$(-1,131; -0,960)$	(1,064;1,266)	(1,023;1,226)	(0,089;0,108)	(0,034;0,066)	
		0,10	$-1,041$	**1,125**	1,041	0,099	0,108	**609**
			$(-1,085; -0,990)$	(1,003;1,289)	(0,973;1,149)	(0,095;0,104)	(0,111;0,198)	
		0,15	$-1,035$	**1,146**	1,047	0,091	0,141	**607**
			$(-1,095; -0,955)$	(1,013;1,246)	(0,996;1,108)	(0,071;0,099)	(0,099;0,173)	
		0,20	$-1,050$	1,023	1,064	0,091	0,194	**643**
			$(-1,204; -0,877)$	(0,890;1,181)	(0,914;1,216)	(0,082;0,099)	(0,153;0,246)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 26b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
5	0,15	0,05	$\mathbf{-1,128}$	**1,159**	1,020	0,147	0,046	**605**
			(−1,155; −1,081)	(1,020;1,165)	(0,997;1,202)	(0,143;0,150)	(0,037;0,064)	
		0,10	−1,022	0,967	0,982	0,139	0,087	**608**
			(−1,197; −0,874)	(0,789;1,180)	(0,883;1,275)	(0,131;0,158)	(0,060;0,125)	
		0,15	−0,976	1,026	1,047	0,141	**0,168**	**622**
			(−1,139; −0,823)	(0,905;1,220)	(0,892;1,180)	(0,116;0,153)	(0,106;0,198)	
		0,20	$\mathbf{-1,280}$	**1,250**	**1,150**	0,139	0,189	**611**
			(−1,357; −1,032)	(0,973;1,393)	(0,912;1,368)	(0,129;0,149)	(0,142;0,201)	
	0,20	0,05	−0,978	1,083	1,051	0,198	0,048	**604**
			(−1,045; −0,932)	(1,010;1,143)	(1,002;1,132)	(0,195;0,203)	(0,038;0,061)	
		0,10	$\mathbf{-1,153}$	1,082	1,083	0,211	0,107	**607**
			(−1,327; −0,993)	(1,064;1,449)	(0,978;1,352)	(0,176;0,214)	(0,068;0,149)	
		0,15	−1,005	1,016	1,050	0,186	0,149	**617**
			(−1,101; −0,818)	(0,921;1,262)	(0,957;1,277)	(0,169;0,199)	(0,109;0,180)	
		0,20	$\mathbf{-1,137}$	**1,106**	**1,152**	0,190	0,205	**614**
			(−1,294; −0,919)	(1,064;1,257)	(1,043;1,281)	(0,183;0,202)	(0,146;0,223)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 27a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
10	0,05	0,05	$-1,018$	1,095	1,086	0,047	0,063	**620**
			$(-1,076; -0,949)$	(0,953;1,122)	(0,998;1,187)	(0,040;0,056)	(0,041;0,083)	
		0,10	$-1,003$	1,004	1,033	0,050	0,104	**621**
			$(-1,123; -0,962)$	(0,947;1,099)	(0,997;1,124)	(0,044;0,056)	(0,090;0,116)	
		0,15	$-1,025$	1,015	1,013	0,049	0,140	**635**
			$(-1,100; -0,950)$	(0,935;1,128)	(0,929;1,128)	(0,041;0,055)	(0,108;0,177)	
		0,20	$-1,042$	1,039	1,028	0,047	0,197	715
			$(-1,121; -0,964)$	(0,932;1,144)	(0,935;1,151)	(0,039;0,055)	(0,158;0,233)	
	0,10	0,05	$-0,980$	1,048	1,033	0,090	0,051	**619**
			$(-1,041; -0,915)$	(0,971;1,104)	(0,991;1,108)	(0,086;0,101)	(0,040;0,078)	
		0,10	$-0,989$	1,039	0,997	0,100	0,096	**618**
			$(-1,034; -0,920)$	(0,917;1,125)	(0,912;1,071)	(0,088;0,104)	(0,053;0,117)	
		0,15	$-0,995$	1,000	1,060	0,100	0,151	**625**
			$(-1,084; -0,896)$	(0,939;1,127)	(0,957;1,118)	(0,089;0,112)	(0,123;0,182)	
		0,20	$-1,016$	1,043	1,021	0,093	0,196	**643**
			$(-1,071; -1,008)$	(0,961;1,150)	(0,945;1,142)	(0,083;0,101)	(0,158;0,226)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 27b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
10	0,15	0,05	$-1,002$	1,018	1,091	0,149	0,049	**614**
			$(-1,032; -0,954)$	(0,954;1,131)	(0,984;1,135)	(0,137;0,154)	(0,028;0,051)	
		0,10	$-1,011$	0,977	1,005	0,146	0,090	**611**
			$(-1,032; -0,939)$	(0,944;1,092)	(0,931;1,078)	(0,139;0,148)	(0,051;0,093)	
		0,15	$-0,982$	1,027	1,003	0,143	0,149	**640**
			$(-1,082; -0,907)$	(0,928;1,162)	(0,929;1,131)	(0,132;0,153)	(0,107;0,175)	
		0,20	$-1,002$	1,003	1,005	0,143	0,193	714
			$(-1,108; -0,939)$	(0,901;1,125)	(0,900;1,166)	(0,135;0,155)	(0,159;0,224)	
	0,20	0,05	$-0,921$	1,002	1,024	0,185	0,050	**604**
			$(-1,043; -0,888)$	(0,957;1,118)	(0,947;1,120)	(0,182;0,186)	(0,029;0,070)	
		0,10	$-0,996$	1,063	1,001	0,193	0,089	**617**
			$(-1,029; -0,948)$	(0,935;1,194)	(0,946;1,134)	(0,185;0,199)	(0,080;0,131)	
		0,15	$-1,065$	1,057	1,053	0,198	0,150	**610**
			$(-1,137; -0,892)$	0,990;1,193)	(0,929;1,117)	(0,175;0,211)	(0,118;0,178)	
		0,20	$-1,016$	1,026	1,024	0,194	0,185	**648**
			$(-1,127; -0,925)$	(0,910;1,183)	(0,899;1,166)	(0,183;0,204)	(0,152;0,223)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 28a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
20	0,05	0,05	$-0,988$	1,063	1,073	0,048	0,051	**663**
			$(-1,017; -0,948)$	(0,988;1,112)	(0,999;1,107)	(0,043;0,052)	(0,040;0,065)	
		0,10	$-0,998$	1,062	1,056	0,049	0,100	742
			$(-1,039; -0,945)$	(0,980;1,130)	(0,998;1,114)	(0,044;0,055)	(0,085;0,122)	
		0,15	$-0,994$	1,029	1,021	0,049	0,144	787
			$(-1,038; -0,956)$	(0,979;1,085)	(0,975;1,084)	(0,044;0,054)	(0,123;0,167)	
		0,20	$-1,012$	1,033	1,029	0,048	0,193	1000
			$(-1,060; -0,966)$	(0,969;1,102)	(0,964;1,100)	(0,043;0,054)	(0,167;0,216)	
	0,10	0,05	$-1,016$	1,071	1,075	0,102	0,049	**613**
			$(-1,062; -0,972)$	(0,996;1,084)	(0,095;1,122)	(0,086;0,103)	(0,032;0,069)	
		0,10	$-0,981$	1,066	1,063	0,098	0,104	733
			$(-1,030; -0,934)$	(0,991;1,128)	(0,999;1,121)	(0,091;0,105)	(0,083;0,119)	
		0,15	$-0,984$	1,040	1,030	0,097	0,145	725
			$(-1,030; -0,936)$	(0,969;1,109)	(0,963;1,112)	(0,091;0,104)	(0,124;0,169)	
		0,20	$-0,993$	1,035	1,030	0,097	0,193	1000
			$(-1,046; -0,936)$	(0,964;1,116)	(0,963;1,114)	(0,090;0,104)	(0,171;0,219)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 28b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
20	0,15	0,05	$-0,974$	1,056	1,048	0,145	0,051	**641**
			$(-1,012; -0,937)$	(1,000;1,106)	(1,003;1,123)	(0,141;0,152)	(0,036;0,074)	
		0,10	$-0,962$	1,015	1,020	0,147	0,103	**660**
			$(-1,039; -0,899)$	(0,964;1,086)	(0,966;1,092)	(0,136;0,159)	(0,079;0,123)	
		0,15	$-0,986$	1,035	1,025	0,147	0,146	730
			$(-1,039; -0,925)$	(0,954;1,112)	(0,947;1,101)	(0,139;0,154)	(0,122;0,170)	
		0,20	$-0,977$	1,028	1,024	0,145	0,195	1000
			$(-1,048; -0,915)$	(0,947;1,118)	(0,945;1,120)	(0,138;0,153)	(0,171;0,219)	
	0,20	0,05	$-0,970$	1,044	1,043	0,199	0,053	**638**
			$(-1,023; -0,934)$	(0,979;1,114)	(0,993;1,132)	(0,187;0,205)	(0,031;0,063)	
		0,10	$-0,972$	1,035	1,041	0,196	0,097	838
			$(-1,043; -0,911)$	(0,956;1,112)	(0,963;1,111)	(0,187;0,205)	(0,077;0,118)	
		0,15	$-0,980$	1,033	1,030	0,196	0,147	895
			$(-1,059; -0,912)$	(0,935;1,123)	(0,938;1,124)	(0,186;0,205)	(0,122;0,168)	
		0,20	$-0,983$	1,038	1,046	0,195	0,195	747
			$(-1,069; -0,900)$	(0,950;1,122)	(0,932;1,122)	(0,187;0,205)	(0,172;0,219)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 29a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
50	0,05	0,05	$-0,940$	1,004	1,061	0,050	0,048	798
			$(-0,973; -0,910)$	(0,995;1,096)	(0,998;1,096)	(0,047;0,053)	(0,038;0,059)	
		0,10	$-0,942$	1,035	1,054	0,049	0,097	1000
			$(-0,978; -0,911)$	(0,994;1,093)	(0,995;1,097)	(0,046;0,053)	(0,084;0,109)	
		0,15	$-0,944$	1,035	1,050	0,050	0,145	1000
			$(-0,980; -0,913)$	(0,993;1,094)	(0,991;1,090)	(0,046;0,053)	(0,130;0,158)	
		0,20	$-0,949$	1,024	1,045	0,050	0,191	1000
			$(-0,981; -0,915)$	(0,991;1,086)	(0,994;1,084)	(0,046;0,053)	(0,177;0,205)	
	0,10	0,05	$-0,929$	1,030	1,052	0,100	0,050	1000
			$(-0,967; -0,895)$	(0,990;1,091)	(0,990;1,095)	(0,095;0,105)	(0,040;0,060)	
		0,10	$-0,932$	1,025	1,050	0,100	0,097	1000
			$(-0,972; -0,894)$	(0,986;1,089)	(0,987;1,090)	(0,095;0,105)	(0,085;0,110)	
		0,15	$-0,937$	1,023	1,046	0,100	0,145	1000
			$(-0,976; -0,901)$	(0,984;1,093)	(0,986;1,094)	(0,095;0,105)	(0,131;0,158)	
		0,20	$-0,940$	1,020	1,043	0,101	0,192	1000
			$(-0,976; -0,898)$	(0,979;1,091)	(0,980;1,095)	(0,096;0,105)	(0,176;0,206)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 29b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
50	0,15	0,05	$-0,938$	1,049	1,049	0,151	0,050	779
			$(-0,99; -0,898)$	(0,998;1,092)	(0,998;1,098)	(0,146;0,157)	(0,040;0,059)	
		0,10	$-0,938$	1,045	1,044	0,151	0,097	1000
			$(-0,98; -0,897)$	(0,993;1,090)	(0,996;1,094)	(0,146;0,157)	(0,085;0,111)	
		0,15	$-0,946$	1,038	1,044	0,152	0,144	1000
			$(-0,99; -0,905)$	(0,994;1,092)	(0,996;1,093)	(0,146;0,158)	(0,130;0,159)	
		0,20	$-0,944$	1,034	1,033	0,152	0,190	1000
			$(-0,99; -0,899)$	(0,981;1,093)	(0,977;1,091)	(0,146;0,159)	(0,175;0,204)	
	0,20	0,05	$-0,950$	1,039	1,031	0,206	0,049	**626**
			$(-0,99; -0,929)$	(0,998;1,104)	(0,983;1,083)	(0,199;0,215)	(0,042;0,055)	
		0,10	$-0,946$	1,034	1,034	0,203	0,097	1000
			$(-0,99; -0,899)$	(0,986;1,086)	(0,982;1,083)	(0,195;0,211)	(0,083;0,111)	
		0,15	$-0,952$	1,037	1,036	0,203	0,146	1000
			$(-1,00; -0,902)$	(0,980;1,093)	(0,979;1,093)	(0,195;0,212)	(0,132;0,160)	
		0,20	$-0,957$	1,028	1,034	0,205	0,192	1000
			$(-1,00; -0,905)$	(0,970;1,092)	(0,972;1,096)	(0,197;0,213)	(0,176;0,208)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 30a. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
100	0,05	0,05	$-0,990$	0,999	1,003	0,049	0,044	**605**
			$(-1,020; -0,989)$	(0,974;1,017)	(0,995;1,008)	(0,047;0,049)	(0,040;0,066)	
		0,10	$-1,013$	0,984	0,993	0,049	0,097	**605**
			$(-1,029; -0,987)$	(0,973;1,005)	(0,978;1,037)	(0,048;0,050)	(0,092;0,105)	
		0,15	$-0,984$	0,995	0,966	0,049	0,142	**604**
			$(-0,996; -0,949)$	(0,973;1,001)	(0,955;1,021)	(0,048;0,050)	(0,135;0,156)	
		0,20	$-1,027$	1,002	0,984	0,050	0,188	**605**
			$(-1,030; -0,994)$	(0,976;1,006)	(0,969;0,997)	(0,046;0,051)	(0,178;0,192)	
	0,10	0,05	$-1,002$	1,007	1,017	0,101	0,054	**605**
			$(-1,071; -0,978)$	(0,952;1,013)	(0,993;1,020)	(0,099;0,104)	(0,051;0,058)	
		0,10	$-0,977$	1,000	1,007	0,105	0,112	**605**
			$(-1,001; -0,973)$	(0,953;1,012)	(0,991;1,009)	(0,098;0,106)	(0,104;0,115)	
		0,15	$-0,973$	0,984	0,940	0,097	0,149	**605**
			$(-0,976; -0,970)$	(0,968;0,989)	(0,936;1,011)	(0,095;0,098)	(0,148;0,158)	
		0,20	$-0,982$	1,030	1,021	0,100	0,213	**605**
			$(-1,021; -0,981)$	(1,021;1,042)	(0,994;1,026)	(0,098;0,101)	(0,196;0,219)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 30b. Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Estimação da razão da má classificação e dos coeficientes do modelo logístico. Valores verdadeiros: $\beta_0 = -1.0, \beta_1 = 1.0,$ e $\theta = \sigma^2_{u0} = 1.0$.(*)

n	Prob. erro		β_0	β_1	σ^2_{u0}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	
100	0,15	0,05	$-1,002$	1,006	1,004	0,150	0,101	1000
			$(-1,023; -0,984)$	(0,983;1,030)	(0,982;1,028)	(0,147;0,155)	(0,086;0,108)	
		0,10	$-1,002$	1,004	1,007	0,149	0,050	1000
			$(-1,023; -0,982)$	(0,985;1,027)	(0,986;1,029)	(0,144;0,155)	(0,048;0,052)	
		0,15	$-1,004$	1,005	1,004	0,150	0,148	1000
			$(-1,022; -0,981)$	(0,982;1,030)	(0,983;1,029)	(0,145;0,154)	(0,134;0,157)	
		0,20	$-1,002$	1,004	1,005	0,151	0,197	1000
			$(-1,024; -0,986)$	(0,984;1,030)	(0,983;1,025)	(0,147;0,153)	(0,188;0,210)	
	0,20	0,05	$-1,003$	1,002	1,004	0,199	0,049	1000
			$(-1,032; -0,973)$	(0,972;1,035)	(0,976;1,036)	(0,195;0,204)	(0,042;0,057)	
		0,10	$-0,999$	1,001	1,000	0,199	0,100	1000
			$(-1,032; -0,965)$	(0,968;1,035)	(0,966;1,035)	(0,195;0,204)	(0,089;0,109)	
		0,15	$-1,002$	1,002	1,004	0,200	0,149	1000
			$(-1,040; -0,969)$	(0,964;1,044)	(0,965;1,044)	(0,195;0,204)	(0,138;0,159)	
		0,20	$-1,002$	1,000	1,005	0,200	0,199	1000
			$(-1,041; -0,968)$	(0,960;1,044)	(0,957;1,051)	(0,195;0,204)	(0,186;0,210)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 31a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
5	0,05	0,05	$-1,074$	1,018	1,101	1,015	1,035	0,049	0,050	1000
			$(-1,281; -0,880)$	(0,863;1,207)	(0,959;1,271)	(0,886;1,171)	(0,907;1,183)	(0,024;0,078)	(0,000;0,107)	
		0,10	$\mathbf{-1,110}$	1,040	1,084	1,001	1,036	0,049	0,100	1000
			$(-1,332; -0,911)$	(0,827 ;1,232)	(0,935;1,281)	(0,844;1,206)	(0,882;1,224)	(0,022;0,076)	(0,019;0,180)	
		0,15	$-1,104$	0,987	1,057	1,000	1,004	0,0478	0,152	1000
			$(-1,357; -0,881)$	(0,808;1,231)	(0,900;1,280)	(0,841;1,230)	(0,855;1,206)	(0,022;0,074)	(0,042;0,263)	
		0,20	$\mathbf{-1,119}$	1,028	1,042	1,010	1,041	0,0456	0,209	**609**
			$(-1,391; -0,845)$	(0,768;1,255)	(0,861;1,283)	(0,815;1,244)	(0,861;1,137)	(0,019;0,073)	(0,067;0,335)	
	0,10	0,05	$-1,092$	1,045	**1,107**	1,026	1,022	0,105	0,050	931
			$(-1,301; -0,863)$	(0,885 ;1,261)	(0,959;1,270)	(0,873;1,195)	(0,877;1,210)	(0,064;0,142)	(0,000;0,111)	
		0,10	$-1,096$	1,019	1,083	1,018	1,000	0,100	0,101	936
			$(-1,334; -0,860)$	(0,831;1,274)	(0,928;1,294)	(0,858;1,227)	(0,855;1,201)	(0,061;0,146)	(0,028;0,201)	
		0,15	$-1,087$	1,007	1,065	1,016	1,025	0,094	0,151	940
			$(-1,383; -0,854)$	(0,790;1,292)	(0,864;1,325)	(0,811;1,259)	(0,812;1,188)	(0,056;0,138)	(0,062;0,213)	
		0,20	$\mathbf{-1,156}$	1,002	1,013	1,000	1,066	0,201	0,220	**600**
			$(-1,418; -0,876)$	(0,735;1,308)	(0,851;1,332)	(0,789;1,344)	(0,827;1,322)	(0,050; 0,138)	(0,077;0,201)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 31b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
5	0,15	0,05	$-1,064$	1,074	1,104	1,035	1,086	0,152	0,049	900
			$(-1,331; -0,834)$	(0,870;1,321)	(0,951;1,319)	(0,864; 1,245)	(0,890;1,164)	(0,104;0,202)	(0,000;0,110)	
		0,10	$-1,071$	1,027	1,034	1,019	1,018	0,152	0,095	1000
			$(-1,398; -0,877)$	(0,830;1,342)	(0,852;1,317)	(0,829;1,279)	(0,846;1,237)	(0,091;0,244)	(0,000;0,187)	
		0,15	$\mathbf{-1,309}$	**1,138**	**1,318**	**1,142**	1,014	0,157	0,152	800
			$(-1,592; -1,058)$	(0,937;1,750)	(1,026;1,584)	(0,940;1,512)	(0,911;1,231)	(0,113;0,196)	(0,117;0,187)	
		0,20	$-1,078$	0,927	1,070	1,008	1,061	0,151	0,231	1000
			$(-1,470; -0,827)$	(0,703;1,430)	(0,804 ; 1,277)	(0,701;1,313)	(0,788;1,348)	(0,106;0,196)	(0,072;0,331)	
	0,20	0,05	$-1,063$	**1,107**	**1,122**	1,037	1,033	0,204	0,051	900
			$(-1,346; -0,823)$	(0,860;1,376)	(0,931;1,352)	(0,875;1,256)	(0,916;1,227)	(0,138;0,270)	(0,000;0,112)	
		0,10	$\mathbf{-1,128}$	1,056	1,066	1,016	1,013	0,203	0,103	906
			$(-1,386; -0,785)$	(0,808;1,330)	(0,858;1,335)	(0,799;1,291)	(0,853;1,317)	(0,134;0,274)	(0,000;0,193)	
		0,15	$\mathbf{-1,124}$	1,052	1,087	1,031	1,082	0,202	0,152	964
			$(-1,498; -0,802)$	(0,798;1,366)	(0,843;1,323)	(0,796;1,315)	(0,884;1,453)	(0,135;0,269)	(0,031;0,273)	
		0,20	$-1,102$	1,051	1,063	1,041	1,087	0,201	0,203	1000
			$(-1,547; -0,950)$	(0,728;1,363)	(0,755;1,349)	(0,799;1,269)	(0,791;1,236)	(0,138;0,264)	(0,110;0,316)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 32a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
10	0,05	0,05	$-1,022$	0,990	1,070	0,978	0,913	0,053	0,057	1000
			$(-1,147; -0,899)$	(0,882;1,116)	(0,976;1,174)	(0,889;1,088)	(0,888;1,002)	(0,034;0,074)	(0,009;0,106)	
		0,10	$-1,033$	0,962	1,032	0,975	0,968	0,052	0,104	1000
			$(-1,168; -0,891)$	(0,844;1,092)	(0,924;1,159)	(0,871;1,093)	(0,900;1,072)	(0,034;0,072)	(0,060;0,178)	
		0,15	$-1,040$	0,953	1,026	0,983	1,046	0,052	0,153	1000
			$(-1,196; -0,896)$	(0,821;1,103)	(0,906;1,151)	(0,849;1,112)	(0,944;1,125)	(0,034;0,070)	(0,110;0,247)	
		0,20	$-1,021$	0,940	1,004	0,983	0,998	0,054	0,189	800
			$(-1,239; -1,009)$	(0,829;0,873)	(0,929;1,067)	(0,847;1,094)	(0,874; 1,142)	(0,032;0,069)	(0,137;0,240)	
	0,10	0,05	$-1,037$	0,995	1,001	0,985	0,980	0,103	0,054	1000
			$(-1,179; -0,889)$	(0,876;1,139)	(0,946;1,171)	(0,879;1,121)	(0,916;1,069)	(0,084;0,121)	(0,011;0,100)	
		0,10	$-1,041$	0,988	1,039	0,982	1,030	0,104	0,103	1000
			$(-1,213; -0,8948)$	(0,845;1,150)	(0,925;1,173)	(0,867;1,128)	(0,884;1,105)	(0,082;0,124)	(0,055;0,168)	
		0,15	$-1,034$	0,968	1,035	0,988	1,040	0,105	0,151	1000
			$(-1,205; -0,861)$	(0,825;1,136)	(0,884;1,189)	(0,852;1,144)	(0,939;1,123)	(0,081;0,138)	(0,109;0,204)	
		0,20	$-1,001$	0,963	1,016	0,978	1,071	0,104	0,205	1000
			$(-1,261; -0,874)$	(0,776;1,166)	(0,846;1,204)	(0,812;1,161)	(0,896;1,149)	(0,076;0,118)	(0,151;0,305)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 32b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
10	0,15	0,05	$-1,022$	0,990	1,010	0,978	0,961	0,145	0,052	1000
			$(-1,147; -0,899)$	(0,882;1,116)	(0,976;1,174)	(0,889;1,088)	(0,899;1,186)	(0,129;0,145)	(0,041;0,0,063)	
		0,10	$-1,043$	0,994	1,033	0,993	0,956	0,155	0,104	1000
			$(-1,221; -0,879)$	(0,851;1,181)	(0,892; 1,204)	(0,862;1,160)	(0,803;1,170)	(0,116;0,174)	(0,056;0,143)	
		0,15	$-1,055$	0,995	1,024	0,992	0,961	0,149	0,150	1000
			$(-1,275; -0,849)$	(0,817;1,203)	(0,873;1,203)	(0,840; 1,177)	(0,899;1,186)	(0,140;0,204)	(0,100;0,201)	
		0,20	$-1,031$	0,984	1,030	0,990	0,923	0,190	0,206	865
			$(-1,282; -0,847)$	(0,785;1,190)	(0,862;1,219)	(0,822;1,185)	(0,834;1,131)	(0,137;0,243)	(0,168;0,250)	
	0,20	0,05	$-1,032$	1,016	1,033	0,987	0,944	0,204	0,052	1000
			$(-1,214; -0,865)$	(0,873;1,200)	(0,916;1,18)	(0,878;1,134)	(0,869;0,976)	(0,163;0,235)	(0,006;0,096)	
		0,10	$-1,043$	1,000	1,023	0,995	0,925	0,201	0,103	1000
			$(-1,258; -0,876)$	(0,829;1,183)	(0,890;1,194)	(0,857;1,180)	(0,803;0,325)	(0,160;0,237)	(0,042;0,159)	
		0,15	$-1,075$	0,994	1,014	1,005	0,959	0,203	0,152	1000
			$(-1,296; -0,853)$	(0,801;1,230)	(0,853;1,221)	(0,821;1,211)	(0,752;1,095)	(0,158;0,243)	(0,102;0,219)	
		0,20	$-1,066$	0,993	1,013	0,994	1,008	0,202	0,209	808
			$(-1,313; -0,831)$	(0,779 ;1,239)	(0,849;1,243)	(0,813;1,220)	(0,823;1,940)	(0,163;0,236)	(0,167;0,249)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 33a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
20	0,05	0,05	$-0,969$	0,9611	1,044	0,984	0,965	0,053	0,052	1000
			$(-1,054; -0,884)$	(0,887;1,044)	(0,976;1,113)	(0,902;1,048)	(0,9497;1,0439)	(0,039;0,067)	(0,019;0,080)	
		0,10	$-0,979$	0,9418	1,018	0,973	0,978	0,053	0,105	1000
			$(-1,075; -0,888)$	(0,862;1,036)	(0,949;1,108)	(0,895;1,054)	(0,924;1,099)	(0,049;0,077)	(0,079;0,139)	
		0,15	$-0,979$	0,931	1,013	0,964	1,017	0,051	0,149	1000
			$(-1,084; -0,880)$	(0,837;1,043)	(0,921;1,110)	(0,877;1,063)	(0,9806;1,1249)	(0,046;0,074)	(0,131;0,171)	
		0,20	$-0,979$	0,929	1,006	0,973	0,996	0,049	0,203	1000
			$(-1,099; -0,868)$	(0,821;1,044)	(0,907;1,111)	(0,868;1,065)	(0,980;1,038)	(0,044;0,072)	(0,158;0,250)	
	0,10	0,05	$-0,983$	0,969	1,034	0,972	1,002	0,106	0,053	1000
			$(-1,09; -0,886)$	(0,888;1,064)	(0,952;1,106)	(0,894;1,058)	(0,8979;1,1016)	(0,084;0,136)	(0,027;0,068)	
		0,10	$-0,989$	0,949	1,005	0,960	1,020	0,103	0,100	1000
			$(-1,097; -0,891)$	(0,855;1,056)	(0,919;1,105)	(0,875;1,064)	(0,9675;1,092)	(0,084;0,135)	(0,073;0,129)	
		0,15	$-1,001$	0,946	1,002	0,975	0,957	0,103	0,153	1000
			$(-1,108; -0,885)$	(0,833;1,058)	(0,901;1,114)	(0,873;1,089)	(0,826;1,066)	(0,082;0,123)	(0,128;0,174)	
		0,20	$-1,008$	0,952	1,000	0,981	1,013	0,103	0,205	1000
			$(-1,130; -0,883)$	(0,840;1,072)	(0,896;1,126)	(0,871;1,087)	(0,964;1,073)	(0,082;0,122)	(0,191;0,223)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 33b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
20	0,15	0,05	$-0,999$	0,996	1,032	0,975	1,048	0,152	0,054	1000
			$(-1,112; -0,892)$	(0,890;1,096)	(0,948;1,115)	(0,889;1,077)	(0,995;1,106)	(0,124;0,184)	(0,023;0,085)	
		0,10	$-1,010$	0,961	1,014	0,972	0,987	0,152	0,102	1000
			$(-1,150; -0,884)$	(0,853;1,091)	(0,909;1,129)	(0,872;1,085)	(0,961;1,160)	(0,124;0,183)	(0,072;0,129)	
		0,15	$-1,014$	0,948	0,999	0,980	0,979	0,151	0,152	1000
			$(-1,149; -0,889)$	(0,839;1,089)	(0,892;1,127)	(0,867;1,1036)	(0,890;1,039)	(0,125;0,169)	(0,124;0,183)	
		0,20	$-1,004$	0,949	0,991	0,983	0,923	0,150	0,203	1000
			$(-1,171; -0,864)$	(0,824;1,088)	(0,885;1,135)	(0,868;1,113)	(0,834;1,131)	(0,124;0,166)	(0,180;0,224)	
	0,20	0,05	$-1,017$	0,983	1,014	0,974	1,006	0,204	0,050	1000
			$(-1,15; -0,890)$	(0,878;1,104)	(0,927;1,110)	(0,888;1,070)	(0,978;1,100)	(0,246;0,307)	(0,025;0,080)	
		0,10	$-1,024$	0,969	0,998	0,969	1,041	0,203	0,109	1000
			$(-1,168; -0,897)$	(0,846;1,098)	(0,889;1,112)	(0,866;1,088)	(0,954;1,080)	(0,245;0,301)	(0,070;0,151)	
		0,15	$-1,044$	0,964	1,003	0,979	0,972	0,202	0,147	1000
			$(-1,195; -0,884)$	(0,834 ;1,117)	(0,891; 1,141)	(0,865;1,112)	(0,957;1,008)	(0,243;0,304)	(0,124;0,176)	
		0,20	$-1,042$	0,961	0,994	0,981	0,959	0,201	0,204	1000
			$(-1,221; -0,883)$	(0,819;1,129)	(0,851;1,146)	(0,851;1,131)	(0,899;1,188)	(0,240;0,300)	(0,176;0,222)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 34a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
50	0,05	0,05	$-0,906$	0,951	1,023	0,962	1,004	0,053	0,053	1000
			$(-0,981; -0,826)$	(0,901;1,000)	(0,980;1,069)	(0,918;1,010)	(0,967;1,041)	(0,039;0,064)	(0,037;0,071)	
		0,10	$-0,905$	0,937	1,011	0,960	1,026	0,052	0,104	1000
			$(-0,972; -0,826)$	(0,880;0,999)	(0,964;1,070)	(0,909;1,016)	(0,957;1,081)	(0,039;0,062)	(0,083;0,130)	
		0,15	$-0,902$	0,938	1,017	0,969	1,004	0,052	0,151	1000
			$(-0,967; -0,829)$	(0,877;1,005)	(0,955;1,076)	(0,913;1,026)	(0,965;1,020)	(0,038;0,063)	(0,140;0,203)	
		0,20	$-0,903$	0,930	1,002	0,962	1,005	0,051	0,202	1000
			$(-0,980; -0,817)$	(0,865;0,999)	(0,943;1,071)	(0,900;1,021)	(0,977; 1,037)	(0,038;0,062)	(0,200;0,276)	
	0,10	0,05	$-0,934$	0,952	1,011	0,962	1,006	0,104	0,053	1000
			$(-1,018; -0,853)$	(0,892;1,017)	(0,958;1,075)	(0,910;1,019)	(0,964;1,035)	(0,084;0,119)	(0,034;0,071)	
		0,10	$-0,936$	0,956	1,011	0,976	1,029	0,103	0,100	1000
			$(-1,018; -0,861)$	(0,895;1,026)	(0,954;1,074)	(0,915;1,032)	(0,986;1,103)	(0,084;0,117)	(0,074;0,125)	
		0,15	$-0,934$	0,944	0,996	0,966	1,024	0,102	0,152	1000
			$(-1,013; -0,852)$	(0,871;1,017)	(0,939;1,063)	(0,909;1,027)	(0,928;1,054)	(0,082;0,113)	(0,123;0,183)	
		0,20	$-0,923$	0,936	0,998	0,961	0,962	0,103	0,201	1000
			$(-1,007; -0,850)$	(0,856; 1,010)	(0,927;1,075)	(0,900;1,035)	(0,922;1,016)	(0,084;0,116)	(0,181;0,221)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 34b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
50	0,15	0,05	$-0,965$	0,962	1,003	0,962	0,990	0,154	0,052	1000
			$(-1,052; -0,885)$	(0,897;1,025)	(0,949; 1,061)	(0,910;1,022)	(0,935;1,038)	(0,122;0,178)	(0,035;0,072)	
		0,10	$-0,965$	0,953	0,996	0,964	1,003	0,152	0,100	1000
			$(-1,051; -0,885)$	(0,884;1,023)	(0,931;1,062)	(0,907;1,036)	(0,956;1,062)	(0,123;0,171)	(0,083;0,134)	
		0,15	$-0,965$	0,943	0,989	0,965	1,009	0,151	0,150	1000
			$(-1,045; -0,886)$	(0,864;1,029)	(0,927;1,065)	(0,900;1,035)	(0,936;1,106)	(0,012;0,169)	(0,140;0,200)	
		0,20	$-0,960$	0,938	0,992	0,962	1,048	0,150	0,204	1000
			$(-1,066; -0,872)$	(0,856;1,030)	(0,917;1,075)	(0,892;1,049)	(0,982;1,095)	(0,123;0,168)	(0,201;0,281)	
	0,20	0,05	$-1,004$	0,972	1,004	0,973	1,041	0,204	0,053	1000
			$(-1,098; -0,918)$	(0,902;1,044)	(0,942;1,070)	(0,909;1,034)	(0,968;1,053)	(0,175;0,224)	(0,034;0,072)	
		0,10	$-0,999$	0,957	0,998	0,972	0,954	0,203	0,108	1000
			$(-1,099; -0,904)$	(0,882;1,042)	(0,924;1,066)	(0,903;1,039)	(0,877;1,042)	(0,173;0,237)	(0,084;0,133)	
		0,15	$-1,002$	0,954	0,991	0,976	0,954	0,201	0,152	1000
			$(-1,098; -0,904)$	(0,874;1,053)	(0,917;1,080)	(0,897;1,052)	(0,918;1,043)	(0,176;0,238)	(0,127;0,174)	
		0,20	$-0,992$	0,946	0,988	0,967	0,998	0,202	0,202	1000
			$(-1,098; -0,899)$	(0,857;1,043)	(0,905;1,081)	(0,879;1,055)	(0,961;1,048)	(0,176;0,238)	(0,164;0,243)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 35a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
100	0,05	0,05	$-1,004$	0,905	0,969	0,962	1,019	0,053	0,054	1000
			$(-1,052; -0,958)$	(0,870;0,946)	(0,937;1,008)	(0,929;0,996)	(0,980;1,046)	(0,039;0,055)	(0,041;0,059)	
		0,10	$-0,992$	0,906	0,974	0,963	0,982	0,052	0,101	1000
			$(-1,042; -0,946)$	(0,868;0,946)	(0,937;1,012)	(0,930;1,000)	(0,960;1,012)	(0,036;0,063)	(0,071;0,125)	
		0,15	$-0,982$	0,907	0,973	0,965	1,017	0,051	0,154	1000
			$(-1,036; -0,934)$	(0,864;0,950)	(0,931;1,014)	(0,924;1,005)	(0,990;1,033)	(0,035;0,059)	(0,134;0,168)	
		0,20	$-0,979$	0,901	0,968	0,958	1,006	0,005	0,201	1000
			$(-1,033; -0,921)$	(0,859;0,943)	(0,923;1,009)	(0,922;1,004)	(0,966;1,054)	(0,043;0,057)	(0,171;0,227)	
	0,10	0,05	$-1,023$	0,926	0,974	0,967	0,991	0,104	0,051	1000
			$(-1,074; -0,972)$	(0,883;0,966)	(0,936;1,009)	(0,931;1,004)	(0,963;1,017)	(0,091;0,110)	(0,040;0,070)	
		0,10	$-1,016$	0,920	0,971	0,929	0,992	0,103	0,103	1000
			$(-1,061; -0,967)$	(0,877;0,963)	(0,932;1,0120)	(0,929;1,005)	(0,969;1,045)	(0,088;0,107)	(0,087;0,123)	
		0,15	$-1,007$	0,910	0,967	0,964	1,037	0,102	0,152	1000
			$(-1,063; -0,947)$	(0,863;0,964)	(0,923;1,014)	(0,918;1,009)	(1,017;1,080)	(0,094;0,113)	(0,135;0,165)	
		0,20	$-1,004$	0,916	0,966	0,964	1,016	0,101	0,203	1000
			$(-1,062; -0,943)$	(0,864;0,962)	(0,920; 1,017)	(0,917;1,012)	(0,922;1,064)	(0,093;0,112)	(0,1840;0,221)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 35b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo logístico hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
100	0,15	0,05	$-1,043$	0,934	0,971	0,969	1,003	0,153	0,005	1000
			$(-1,104; -0,991)$	(0,890;0,982)	(0,934;1,014)	(0,929;1,005)	(0,978;1,046)	(0,139;0,162)	(0,037;0,067)	
		0,10	$-1,030$	0,925	0,966	0,964	1,015	0,152	0,101	1000
			$(-1,096; -0,977)$	(0,875;0,977)	(0,922;1,013)	(0,919;1,010)	(0,993;1,031)	(0,134;0,168)	(0,081;0,118)	
		0,15	$-1,024$	0,926	0,970	0,969	0,989	0,151	0,150	1000
			$(-1,090; -0,963)$	(0,869; 0,980)	(0,920;1,019)	(0,917;1,017)	(0,959;1,070)	(0,131;0,165)	(0,125;0,172)	
		0,20	$-1,027$	0,920	0,961	0,960	0,981	0,150	0,201	1000
			$(-1,095; -0,954)$	(0,860;0,983)	(0,907;1,0170)	(0,912;1,015)	(0,961;1,014)	(0,138;0,161)	(0,173;0,208)	
	0,20	0,05	$-1,062$	0,943	0,975	0,970	1,016	0,204	0,050	1000
			$(-1,125; -1,001)$	(0,895;0,998)	(0,930; 1,020)	(0,926;1,014)	(0,997;1,067)	(0,183;0,213)	(0,036;0,064)	
		0,10	$-1,04$	0,935	0,969	0,970	0,980	0,203	0,101	1000
			$(-1,109; -0,982)$	(0,877; 0,991)	(0,920;1,020)	(0,915;1,019)	(0,957;1,000)	(0,190;0,206)	(0,082;0,119)	
		0,15	$-1,044$	0,932	0,964	0,966	1,014	0,202	0,151	1000
			$(-1,110; -0,974)$	(0,871;0,991)	(0,916;1,021)	(0,914;1,025)	(0,982;1,043)	(0,187;0,204)	(0,123;0,168)	
		0,20	$-1,037$	0,929	0,972	0,968	1,002	0,201	0,201	1000
			$(-1,116; -0,961)$	(0,859;1,000)	(0,909;1,028)	(0,907;1,023)	(0,968;1,036)	(0,182;0,209)	(0,174;0,206)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 36a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
5	0,05	0,05	$-1,017$	1,032	0,988	**1,113**	1,027	0,043	0,057	975
			$(-1,050; -0,977)$	(0,969;1,126)	(0,948;1,074)	(0,980;1,145)	(0,948;1,173)	(0,035;0,050)	(0,038;0,067)	
		0,10	$-1,065$	1,049	1,084	1,087	1,085	0,047	0,091	1000
			$(-1,232; -0,993)$	(0,986;1,181)	(0,963;1,186)	(0,973;1,194)	(0,954;1,203)	(0,040;0,053)	(0,079;0,100)	
		0,15	$-1,068$	1,010	1,066	1,030	1,050	0,050	0,141	1000
			$(-1,197; -0,994)$	(0,921;1,173)	(0,967;1,165)	(0,932;1,189)	(0,968;1,185)	(0,040;0,061)	(0,115;0,177)	
		0,20	$-1,081$	1,061	1,049	1,036	1,066	0,049	0,184	1000
			$(-1,158; -1,006)$	(0,939;1,171)	(0,897;1,169)	(0,892;1,200)	(0,958;1,221)	(0,040;0,056)	(0,146;0,234)	
	0,10	0,05	$-0,979$	1,099	**1,112**	1,063	1,049	**0,085**	0,056	1000
			$(-1,092; -0,905)$	(0,979;1,182)	(0,972;1,185)	(0,961;1,158)	(0,971;1,166)	(0,077;0,100)	(0,036;0,075)	
		0,10	$-1,025$	1,016	0,997	1,020	0,984	0,090	0,081	1000
			$(-1,133; -0,918)$	(0,904;1,121)	(0,918;1,096)	(0,911;1,135)	(0,889;1,097)	(0,081;0,100)	(0,063;0,103)	
		0,15	$-1,062$	1,042	1,017	1,018	0,985	0,090	**0,203**	800
			$(-1,146; -0,905)$	(1,004;1,085)	(0,986;1,083)	(0,997;1,084)	(0,898;1,068)	(0,080;0,097)	(0,142;0,211)	
		0,20	$-1,066$	1,039	1,079	1,035	1,063	0,087	0,197	1000
			$(-1,187; -0,902)$	(0,888;1,220)	(0,918;1,188)	(0,890;1,197)	(0,953;1,243)	(0,074;0,101)	(0,167;0,225)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 36b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
5	0,15	0,05	$-1,007$	0,994	1,102	1,074	1,050	0,141	0,039	975
			$(-1,143; -0,905)$	(0,929;1,185)	(0,985;1,127)	(0,950;1,161)	(0,974;1,179)	(0,123;0,156)	(0,023;0,060)	
		0,10	$-1,022$	1,033	0,915	0,951	0,984	0,158	0,115	**517**
			$(-1,119; -0,959)$	(0,888;1,111)	(0,835;1,091)	(0,862;1,058)	(0,887;1,219)	(0,139;0,172)	(0,069;0,147)	
		0,15	$-0,993$	0,976	1,012	0,998	0,984	0,142	0,145	1000
			$(-1,127; -0,892)$	(0,858;1,189)	(0,902;1,166)	(0,907;1,141)	(0,887;1,219)	(0,128;0,156)	(0,114;0,173)	
		0,20	$-1,060$	1,082	1,095	1,025	1,054	**0,134**	**0,173**	1000
			$(-1,190; -0,902)$	(0,862;1,248)	(0,937;1,279)	(0,888;1,211)	(0,907;1,261)	(0,123;0,150)	(0,143;0,217)	
	0,20	0,05	$-1,037$	1,043	1,037	1,038	0,955	0,182	0,036	957
			$(-1,126; -0,874)$	(0,898;1,125)	(0,907;1,185)	(0,950;1,169)	(0,908;1,126)	(0,169;0,202)	(0,028;0,055)	
		0,10	$-1,041$	1,023	1,088	**1,112**	0,997	0,178	0,093	987
			$(-1,153; -0,856)$	(0,926;1,215)	(0,973;1,249)	(0,933;1,287)	(0,880;1,254)	(0,164;0,200)	(0,077;0,113)	
		0,15	$-1,052$	**1,230**	1,056	**1,119**	1,035	0,180	0,135	987
			$(-1,302; -0,952)$	(0,850;1,329)	(0,940;1,180)	(0,929;1,186)	(0,984;1,286)	(0,171;0,206)	(0,118;0,162)	
		0,20	$-1,100$	1,066	1,005	1,047	1,098	0,181	0,202	1000
			$(-1,368; -0,889)$	(0,880 ;1,369)	(0,928;1,411)	(0,916;1,430)	(0,883;1,466)	(0,173;0,200)	(0,157;0,227)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 37a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
10	0,05	0,05	$-0,952$	0,977	0,997	0,985	1,045	0,049	0,051	800
			$(-1,047;-0,926)$	(0,947;1,108)	(0,934;1,085)	(0,960;1,055)	(0,984;1,085)	(0,036;0,051)	(0,040;0,059)	
		0,10	$-1,046$	0,976	0,985	1,005	0,993	0,044	0,094	1000
			$(-1,083;-0,976)$	(0,971;1,098)	(0,926;1,097)	(0,922;1,060)	(0,917;1,056)	(0,035;0,050)	(0,083;0,104)	
		0,15	$-1,059$	0,975	1,067	1,001	0,980	0,047	0,137	1000
			$(-1,140;-0,999)$	(0,936;1,005)	(0,999;1,108)	(0,915;1,135)	(0,925;1,125)	(0,042;0,052)	(0,123;0,155)	
		0,20	$-1,070$	1,032	1,048	1,041	1,019	0,048	0,201	1000
			$(-1,087;-0,966)$	(0,995;1,070)	(0,986;1,130)	(0,998;1,095)	(0,976;1,053)	(0,045;0,056)	(0,188;0,208)	
	0,10	0,05	$-1,024$	1,078	1,037	1,050	1,143	0,089	0,058	1000
			$(-1,056;-0,970)$	(0,995;1,199)	(0,931;1,173)	(0,976;1,108)	(1,047;1,194)	(0,085;0,100)	(0,047;0,067)	
		0,10	$-0,987$	1,036	1,055	1,056	1,038	0,088	0,115	810
			$(-1,056;-0,933)$	(0,995;1,130)	(0,941;1,147)	(0,998;1,076)	(0,988;1,115)	(0,085;0,100)	(0,087;0,124)	
		0,15	$-1,013$	0,999	1,027	1,011	0,996	0,101	0,147	1000
			$(-1,108;-0,954)$	(0,940;1,016)	(0,947;1,065)	(0,895;1,112)	(0,945;1,104)	(0,091;0,103)	(0,126;0,159)	
		0,20	$-1,000$	0,997	0,985	0,989	1,031	0,094	0,178	1000
			$(-1,036;-0,955)$	(0,977;1,053)	(0,939;1,025)	(0,918;1,011)	(0,955;1,105)	(0,091;0,100)	(0,158;0,204)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 37b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
10	0,15	0,05	$-0,951$	1,004	1,036	0,999	0,987	0,135	0,052	810
			$(-1,001; -0,901)$	(0,957;1,082)	(0,999;1,094)	(0,946;1,055)	(0,967;1,057)	(0,129;0,151)	(0,041;0,063)	
		0,10	$-1,078$	1,032	1,009	1,034	1,086	0,138	0,109	1000
			$(-1,144; -0,996)$	(0,973;1,091)	(0,958;1,141)	(0,938;1,130)	(0,998;1,174)	(0,136;0,152)	(0,088;0,118)	
		0,15	$-1,048$	1,072	1,015	1,022	1,008	0,132	0,139	800
			$(-1,126; -0,970)$	(0,927;1,217)	(0,998;1,135)	(0,999;1,158)	(0,989;1,204)	(0,123;0,151)	(0,119;0,150)	
		0,20	$-0,956$	1,003	1,016	0,970	1,025	0,138	0,191	795
			$(-1,022; -0,842)$	(0,863;1,052)	(0,942;1,036)	(0,950;1,074)	(0,965;1,064)	(0,127;0,150)	(0,176;0,224)	
	0,20	0,05	$-1,014$	1,032	1,006	1,030	1,016	0,190	0,054	1000
			$(-1,138; -0,883)$	(0,968;1,148)	(0,951;1,121)	(1,000;1,167)	(0,940;1,128)	(0,181;0,200)	(0,044;0,062)	
		0,10	$-1,006$	1,016	1,071	1,056	0,937	0,190	0,098	1000
			$(-1,130; -0,802)$	(0,869;1,127)	(0,980;1,110)	(0,871;1,128)	(0,861;1,130)	(0,176;0,209)	(0,084;0,102)	
		0,15	$-1,059$	1,068	1,078	1,051	1,092	0,181	0,150	1000
			$(-1,108; -0,972)$	(0,996;1,084)	(0,988;1,168)	(0,999;1,079)	(0,901;1,128)	(0,179;0,200)	(0,115;0,170)	
		0,20	$-1,115$	1,041	1,021	1,032	1,071	0,185	0,196	900
			$(-1,142; -0,895)$	(0,946; 1,140)	(0,994;1,238)	(0,930;1,241)	(0,902;1,166)	(0,178;200)	(0,178;0,233)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 38a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
20	0,05	0,05	$-0,983$	1,030	1,057	1,005	0,999	0,044	0,052	1000
			$(-1,020; -0,968)$	(0,977;1,054)	(1,000;1,090)	(0,994;1,024)	(0,973;1,014)	(0,043;0,050)	(0,0422;0,056)	
		0,10	$-1,025$	1,019	0,989	1,032	1,006	0,047	0,095	1000
			$(-1,073; -0,958)$	(0,982;1,058)	(0,937;1,042)	(0,990;1,053)	(0,979;1,085)	(0,044;0,053)	(0,083;0,102)	
		0,15	$-1,033$	1,021	1,060	1,042	1,026	0,050	0,141	1000
			$(-1,075; -0,977)$	(0,932;1,096)	(0,990;1,013)	(0,993;1,096)	(0,966;1,115)	(0,046;0,057)	(0,126;0,151)	
		0,20	$-0,999$	1,007	1,017	0,974	1,018	0,046	0,190	1000
			$(-1,048; -0,942)$	(0,912;1,083)	(0,977;1,085)	(0,933;1,112)	(0,884;1,124)	(0,041;0,053)	(0,189;0,200)	
	0,10	0,05	$-1,016$	1,063	1,060	1,078	1,048	0,097	0,038	1000
			$(-1,031; -0,997)$	(1,000;1,090)	(0,970;1,150)	(0,952;1,204)	(0,953;1,143)	(0,093;0,101)	(0,035;0,050)	
		0,10	$-1,060$	1,018	1,077	1,041	1,010	0,100	0,090	1000
			$(-1,071; -0,928)$	(0,958;1,078)	(0,960;1,109)	(0,947;1,194)	(0,985;1,050)	(0,091;0,103)	(0,085;0,108)	
		0,15	$-0,980$	0,993	1,003	0,999	0,983	0,093	0,134	1000
			$(-1,082; -0,904)$	(0,948;1,033)	(0,958;1,085)	(0,939;1,039)	(0,968;1,010)	(0,086;0,100)	(0,118;0,150)	
		0,20	$-0,931$	0,962	1,003	1,029	0,986	0,094	0,190	1000
			$(-1,000; -0,903)$	(0,939;1,086)	(0,950;1,030)	(0,943;1,075)	(0,927;1,104)	(0,089;0,100)	(0,188;0,200)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 38b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
20	0,15	0,05	$-1,005$	1,007	1,001	1,091	1,097	0,142	0,055	1000
			$(-1,080; -0,976)$	(0,954;1,060)	(0,983;1,019)	(0,925;1,319)	(0,928;1,266)	(0,139;0,152)	(0,046;0,060)	
		0,10	$-0,990$	1,019	0,990	1,041	0,961	0,142	0,090	1000
			$(-1,041; -0,895)$	(0,980;1,071)	(0,974;1,034)	(0,972;1,068)	(0,915;1,039)	(0,139;0,150)	(0,077;0,105)	
		0,15	$-1,014$	0,979	1,022	0,998	1,001	0,143	0,140	1000
			$(-1,032; -0,997)$	(0,919;1,047)	(0,970;1,037)	(0,956;1,013)	(0,961;1,055)	(0,141;0,150)	(0,133;0,150)	
		0,20	$-0,932$	1,042	1,057	1,029	1,051	0,145	0,204	1000
			$(-1,061; -0,889)$	(0,985;1,127)	(0,980;1,134)	(0,906;1,152)	(0,999;1,103)	(0,138;0,150)	(0,195;0,215)	
	0,20	0,05	$-0,953$	1,037	1,025	1,002	1,010	0,196	0,046	1000
			$(-1,005; -0,909)$	(0,930;1,074)	(0,938;1,134)	(0,942;1,090)	(0,948;1,130)	(0,188;0,204)	(0,027;0,055)	
		0,10	$-1,046$	1,030	1,048	1,041	1,053	0,192	0,085	1000
			$(-1,108; -0,926)$	(0,922;1,104)	(0,932;1,078)	(0,966;1,099)	(0,934;1,124)	(0,187;0,200)	(0,0753;0,100)	
		0,15	$-1,017$	1,097	1,042	1,010	1,011	0,193	0,145	1000
			$(-1,124; -0,961)$	(0,983;1,139)	(0,977;1,142)	(0,996;1,024)	(0,993;1,029)	(0,191;0,200)	(0,130;0,164)	
		0,20	$-1,064$	1,011	0,963	1,031	1,040	0,197	0,187	1000
			$(-1,090; -0,937)$	(0,960;1,073)	(0,943;1,083)	(0,962;1,116)	(0,927;1,090)	(0,194;0,200)	(0,176;0,200)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 39a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
50	0,05	0,05	$-0,928$	1,035	1,034	1,017	1,027	0,048	0,049	1000
			$(-1,002; -0,909)$	(0,999;1,049)	(0,991;1,064)	(1,000;1,077)	(0,998;1,056)	(0,045;0,054)	(0,047;0,053)	
		0,10	$-0,986$	1,028	1,066	1,024	1,014	0,052	0,095	1000
			$(-1,052; -0,947)$	(0,981;1,084)	(0,983;1,149)	(0,987;1,061)	(0,987;1,030)	(0,046;0,054)	(0,092;0,100)	
		0,15	$-0,933$	0,991	1,007	1,027	1,000	0,045	0,142	1000
			$(-1,061; -0,928)$	(0,964;1,048)	(0,998;1,009)	(0,999;1,068)	(0,976;1,017)	(0,044;0,050)	(0,134;0,152)	
		0,20	$-0,935$	1,042	1,031	1,041	1,041	0,048	0,186	1000
			$(-1,046; -0,903)$	(0,997;1,087)	(0,993; 1,069)	(0,994;1,088)	(0,992;1,121)	(0,046;0,051)	(0,181;0,200)	
	0,10	0,05	$-0,970$	1,010	1,025	1,025	1,020	0,099	0,046	1000
			$(-1,092; -0,907)$	(0,970;1,024)	(0,970;1,037)	(0,968;1,033)	(0,981;1,023)	(0,095;0,105)	(0,042;0,052)	
		0,10	$-0,945$	0,999	0,999	1,029	0,978	0,102	0,091	1000
			$(-1,061; -0,905)$	(0,958;1,038)	(0,949;1,041)	(0,991;1,048)	(0,958;1,029)	(0,101;0,104)	(0,084;0,100)	
		0,15	$-0,948$	1,005	1,054	1,090	1,070	0,098	0,149	1000
			$(-1,069; -0,925)$	(0,954;1,146)	(0,969;1,139)	(0,979;1,201)	(0,982;1,158)	(0,095;0,105)	(0,146;0,157)	
		0,20	$-0,961$	1,040	1,050	1,080	1,054	0,103	0,181	1000
			$(-1,002; -0,923)$	(0,961;1,119)	(1,036;1,113)	(0,998;1,162)	(0,996;1,1,112)	(0,099;0,109)	(0,177;0,200)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 39b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
50	0,15	0,05	$-0,938$	1,018	1,037	1,032	1,021	0,146	0,045	1000
			$(-1,003; -0,905)$	(0,998;1,038)	(0,9779;1,104)	(0,992;1,068)	(0,968;1,054)	(0,142;0,152)	(0,034;0,050)	
		0,10	$-0,908$	1,044	1,058	1,033	1,053	0,145	0,098	1000
			$(-1,000; -0,879)$	(0,979;1,108)	(0,974;1,106)	(0,986;1,087)	(0,992;1,090)	(0,139;0,158)	(0,094;0,103)	
		0,15	$-0,970$	1,044	1,070	1,053	1,053	0,150	0,143	1000
			$(-1,003; -0,922)$	(0,999;1,089)	(0,983;1,157)	(0,987;1,109)	(0,992;1,114)	(0,143;0,158)	(0,135;0,154)	
		0,20	$-0,926$	0,948	0,993	0,965	0,977	0,151	0,188	1000
			$(-1,007; -0,904)$	(0,923;0,968)	(0,931;1,040)	(0,953;0,976)	(0,903;0,993)	(0,146;0,157)	(0,172;0,200)	
	0,20	0,05	$-0,929$	1,020	1,040	1,049	1,011	0,207	0,048	1000
			$(-1,003; -0,909)$	(0,992;1,051)	(0,999;1,065)	(0,988;1,110)	(0,967;1,055)	(0,195;0,207)	(0,040;0,050)	
		0,10	$-0,964$	1,044	1,059	1,056	1,050	0,197	0,098	1000
			$(-1,000; -0,934)$	(0,990;1,098)	(0,985;1,150)	(0,985;1,150)	(0,978;1,120)	(0,192;0,215)	(0,096;0,104)	
		0,15	$-0,952$	1,041	1,052	1,025	1,022	0,201	0,149	1000
			$(-1,000; -0,903)$	(0,986;1,049)	(0,978;1,090)	(0,957;1,093)	(0,969;1,063)	(0,192;0,208)	(0,144;0,153)	
		0,20	$-0,917$	0,983	1,002	0,988	1,016	0,199	0,186	1000
			$(-1,000; -0,903)$	(0,969;1,041)	(0,962;1,045)	(0,947;1,035)	(0,961;1,032)	(0,196;0,208)	(0,176;0,200)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 40a Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
100	0,05	0,05	$-1,000$	0,990	1,015	0,999	1,008	0,049	0,048	1000
			$(-1,014; -0,971)$	(0,977;1,017)	(0,989;1,039)	(0,983;1,030)	(0,979;1,044)	(0,048;0,050)	(0,044;0,051)	
		0,10	$-1,009$	1,011	1,012	1,008	1,014	0,050	0,098	1000
			$(-1,022; -0,997)$	(0,987;1,023)	(0,986;1,040)	(0,981;1,021)	(0,999;1,025)	(0,046;0,051)	(0,093;0,104)	
		0,15	$-1,010$	1,011	0,997	0,997	1,006	0,050	0,152	1000
			$(-1,016; -1,002)$	(0,989;1,032)	(0,986;1,017)	(0,980;1,023)	(0,982;1,015)	(0,048;0,052)	(0,141;0,158)	
		0,20	$-1,004$	0,980	0,985	0,963	0,971	0,052	0,194	1000
			$(-1,019; -0,989)$	(0,942;1,017)	(0,956;0,995)	(0,958;1,007)	(0,932;0,999)	(0,048;0,054)	(0,185;0,201)	
	0,10	0,05	$-0,997$	1,016	1,009	1,016	1,009	0,099	0,052	1000
			$(-1,017; -0,971)$	(0,978;1,035)	(0,984;1,030)	(0,996;1,044)	(0,989;1,029)	(0,096;0,103)	(0,049;0,056)	
		0,10	$-1,000$	0,996	1,013	0,991	1,006	0,099	0,099	1000
			$(-1,027; -0,980)$	(0,970;1,016)	(0,982;1,024)	(0,971;1,026)	(0,966;1,013)	(0,095;0,102)	(0,094;0,102)	
		0,15	$-0,986$	1,000	0,999	1,005	0,980	0,097	0,149	1000
			$(-1,008; -0,978)$	(0,985;1,008)	(0,982;1,008)	(0,986;1,018)	(0,969;1,015)	(0,096;0,100)	(0,139;0,157)	
		0,20	$-1,000$	0,994	0,987	0,998	0,994	0,100	0,200	1000
			$(-1,015; -0,988)$	(0,972;1,040)	(0,952;1,017)	(0,949;1,044)	(0,990;1,022)	(0,097;0,104)	(0,193;0,208)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

Quadro 40b Mediana observada e intervalo interquartílico dos coeficientes de regressão do modelo probit hierárquico, para dados agrupados, simulados com resposta binária má classificada. Valores verdadeiros: $\beta_0 = -1, \beta_1 = 1, \beta_2 = 1$ e $\theta = \sigma^2_{u0} = \sigma^2_{u1} = 1$. (*)

n	Prob. erro		β_0	β_1	β_2	σ^2_{u0}	σ^2_{u1}	$\hat{\tau}_0$	$\hat{\tau}_1$	Conv.
	τ_0	τ_1	Est.	Est.	Est.	Est.	Est.	Est.	Est.	
100	0,15	0,05	$-1,019$	1,019	1,026	1,052	1,022	0,150	0,053	1000
			$(-1,030; -1,008)$	(0,990;1,046)	(0,980;1,072)	(1,031;1,063)	(0,999;1,035)	(0,146;0,152)	(0,050;0,057)	
		0,10	$-1,008$	1,000	0,993	0,990	0,996	0,150	0,104	1000
			$(-1,030; -0,987)$	(0,964;1,049)	(0,983;1,046)	(0,976;1,037)	(0,977;1,053)	(0,148;0,151)	(0,094;0,107)	
		0,15	$-0,992$	1,003	1,009	1,007	0,996	0,150	0,144	1000
			$(-1,004; -0,966)$	(0,976;1,064)	(1,000;1,020)	(0,144;1,025)	(0,968;1,014)	(0,144;0,152)	(0,143;0,156)	
		0,20	$-1,003$	1,010	1,006	1,013	0,992	0,152	0,198	1000
			$(-1,018; -0,967)$	(0,971;1,055)	(0,999;1,050)	(0,961;1,036)	(0,962;1,032)	(0,147;0,156)	(0,191;0,205)	
	0,20	0,05	$-1,000$	1,003	1,001	0,991	0,995	0,199	0,048	1000
			$(-1,012; -0,989)$	(0,993;1,025)	(0,974;1,039)	(0,977;1,015)	(0,967;1,005)	(0,196;0,203)	(0,043;0,054)	
		0,10	$-1,015$	0,995	0,985	1,025	0,982	0,203	0,100	1000
			$(-1,027; -0,994)$	(0,973;1,017)	(0,948;1,020)	(0,976;1,057)	(0,963;1,014)	(0,198;0,205)	(0,092;0,105)	
		0,15	$-0,996$	0,986	0,992	1,004	0,991	0,198	0,145	1000
			$(-1,050; -0,958)$	(0,959;1,017)	(0,944;1,012)	(0,941;1,026)	(0,944;1,021)	(0,193;0,201)	(0,137;0,153)	
		0,20	$-1,027$	0,996	0,989	0,986	0,997	0,200	0,197	1000
			$(-1,049; -0,9565)$	(0,979;1,016)	(0,972;1,020)	(0,965;1,001)	(0,944;1,027)	(0,197;0,202)	(0,191;0,200)	

(*) Prob. erro = probabilidade do erro; Est.= estimativa; Conv.= número de de convergência pelo procedimento `gnlmm`

C. Programa de simulação no R

C.1 - Programa do primeiro estudo de simulação

```
/*********************************************************************
* PROGRAMA: Simulacao para avaliar a magnitude do vies para resposta   *
*           ma classificada do modelo  logistico multinivel (4.2) e    *
*           para avaliar a qualidade da aproximacao (4.17) para beta1  *
*           diferente de zero.                                         *
* ********************************************************************/
/*******************************
*   Bibliotecas necessarias    *
*******************************/
library (MASS)
library (boot)
/*******************************
*  Parametros para simulacao   *
*******************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
b0<--1      /* Valor verdadeiro de beta0 */
b1<-1       /* Valor verdadeiro de beta1 */
R<-1000     /* Numero de replicas */
/*****************************************************************
*                            Corpo                               *
*****************************************************************/
/* Declaracao de variaveis */
dt<-matrix(0,R,2)
gt<-matrix(0,R,1)
dy<-matrix(0,R,2)
gy<-matrix(0,R,1)
/* Inicio do loop da simulacao */
for (r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x<-matrix(0,nj,j)
u<-matrix(0,nj,j)
T<-matrix(0,k,1)
test<-matrix(0,k,3)
ct<-matrix(0,2,1)
proby=matrix(0,k,1)
Y=matrix(0,k,1)
cy<-matrix(0,2,1)
test2<-matrix(0,k,3)
for(i in 1:j){
    x[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
}
x1<-data.frame(x) /* Converte a matriz x em dados */
b<-names(x1)      /* Mostra os nomes das variaveis dos dados a1 */
u1<-data.frame(u) /* Converte a matriz u em dados */
c<-names(u1)      /* Mostra os nomes das variaveis dos dados u1 */
for(i in 1:j){
    names(x1)[i]<-i /* Muda o nome da variavel para numeros */
    names(u1)[i]<-i /* Muda o nome da variavel para numeros */
}
vec1<-stack(x1)  /* Empilha cada coluna de x1 em um vetor */
vec2<-stack(u1)  /* Empilha cada coluna de u1 em um vetor */
x<-vec1[,1]      /* Armazena vec1 em x */
U<-vec2[,1]      /* Armazena vec2 em U */
J<-vec1[,2]      /* Obtem a coluna dos grupos */
```

```
X<-cbind(1,x)    /* Obtem a  matriz de regressores */
/* Geracao de eta com base nos valores verdadeiros dos parametros:
beta_0=-1; beta_1=1; e sigma_0u=1 */
eta<--1*X[,1]+1*X[,2]+U
/* Obter a probabilidade inicial em funcao de eta */
p<-inv.logit(eta) /* Obtem a probabilidade inicial de t*/
/* Gerar T verdadeiro a partir da dist. binomial(1,p) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i])
}
/* Ajustar o MMLG para resposta sem correcao de
   erros T com ligacaologit via funcao glmmPQL  */
test[,1]<-T
test[,2]<-x
test[,3]<-J
dados<-data,frame(test)  /* Converte informacoes matriz num conj. dados */
for(i in 1:3){
    if (i==1){
    names(dados)[1]<-"t" /* Muda o nome da primeira coluna para t */
    }
    if (i==2){
    names(dados)[2]="x1" /* Muda o nome da segunda coluna para x1 */
    }
    if (i==3){
    names(dados)[3]="grup" /* Muda o nome da terceira coluna para grupo*/
    }
}
/* Ajute para a resposta t via procedimento QVP */
ajustet<-glmmPQL(t ~ x1, random = ~ 1 | grup,
                family = binomial (link=logit),data=dados)
bt<-cbind(fixed.effects(ajustet)) /* Estimativas dos efeitos fixos */
ct[,1]<-bt                        /* Armazena as estimativas em um vetor */
dt[r,1]<-ct[1,1]                  /* Armazena a estimativa de beta0 */
dt[r,2]<-ct[2,1]                  /* Armazena a estimativa de beta1 */
et<-cbind(getVarCov(ajustet))     /* Estimativa do efeito aleatorio */
gt[r,1]<-et                       /* Armazena a estimativa do efeito aleatorio */
/* Obter a probabilidade de y com a correcao do erro a partir de t */
for (k in 1:k){
    proby[k]<-(1-tau1-tau0)*p[k]+tau0
}
/* Gerar y a partir da dist binomial(1,proby) */
for(w in 1: k){
    Y[w]<-rbinom(1,1,proby[w])
}
/* Ajustar um MMLG para a resposta y via procedimento QVP */
test2[,1]<-Y
test2[,2]<-x
test2[,3]<-J
dados2<-data.frame(test2)   /* Transforma informacoes y banco dados */
for(i in 1:3){
    if (i==1){
    names(dados2)[1]<-"y"   /* Muda nome da coluna para y */
    }
    if (i==2){
    names(dados2)[2]="x2"   /* Muda nome da coluna para x2 */
    }
    if (i==3){
    names(dados2)[3]="grup2" /* Muda nome da coluna para grup2 */
    }
}

/* Ajuste para a resposta y via procedimento QVP */
ajustey<-glmmPQL(y ~ x2, random = ~ 1 | grup2,
                family = binomial (link=logit),data=dados2)
```

```
by<-cbind(fixed.effects(ajustey))/* Estimativas dos efeitos fixos */
cy[,1]<-by                       /* Armazena as estimativas em um vetor */
dy[r,1]<-cy[1,1]                 /* Armazena a estimativa de beta0 */
dy[r,2]<-cy[2,1]                 /* Armazena a estimativa de beta1 */
ey<-cbind(getVarCov(ajustey))    /* Estimativa do efeito aleatorio */
gy[r,1]<-ey                      /* Armazena a estimativa do efeito aleatorio */
} /* Final do loop */
/* Verificar o numero total de convergencia */
conv<-0  /* Declaracao da variavel conv */
for (r in 1:R){
  if(dt[r,1]!=0){
    conv<-conv+1 /* Conta o numero de convergencia */
  }
}
/* Impressao dos resultados */
nj     /* Imprime o tamanho da amostra no grupo */
tau0  /* Imprime o valor de tau0 */
tau1  /* Imprime o valor de tau0 */
conv   /* Imprime o total de convergencia */
esttb0<-mean(dt[1:conv,1]) /* Media das estimativas beta0 de t */
sdtb0<-sd(dt[1:conv,1])    /* Desvio padrao das estimativas beta0 de t */
esttb0                     /* Imprime  media de beta0 de t  */
sdtb0                      /* Imprime  desvio padrao de beta0 de t */
esttb1<-mean(dt[1:conv,2]) /* Media das estimativas beta1 de t*/
sdtb1<-sd(dt[1:conv,2])    /* Desvio padrao das estimativas beta1 de t */
esttb1                     /* Imprime  media de beta1 de t */
sdtb1                      /* Imprime  media de beta1  de t*/
esttu0<-mean(gt[1:conv,1]) /* Media das estimativas u0 de t */
sdtbu0<-sd(gt[1:conv,1])   /* Desvio padrao das estimativas u0 de t */
esttu0                     /* Imprime  media de u0 de t */
sdtbu0                     /* Imprime  desvio padrao de u0 de t */

estyb0<-mean(dy[1:conv,1]) /* Media das estimativas beta0 de y */
sdyb0<-sd(dy[1:conv,1])    /* Desvio padrao das estimativas beta0 de y */
vyb0<-estyb0-b0            /* Vies de beta0 de y */
estyb0                     /* Imprime  media de beta0 de y  */
sdyb0                      /* Imprime  desvio padrao de beta0 de y  */
vyb0                       /* Imprime  vies de beta0 de y  */
estyb1<-mean(dy[1:conv,2]) /* Media das estimativas bet10 de y */
sdyb1<-sd(dy[1:conv,2])    /* Desvio padrao das estimativas beta1 de y */
vyb1<-estyb1-b1            /* Vies de beta01de y */
estyb1                     /* Imprime  media de beta1 de y  */
sdyb1                      /* Imprime  desvio padrao de beta1 de y  */
vyb1                       /* Imprime  vies de beta1 de y  */
estyu0<-mean(gy[1:conv,1]) /* Media das estimativas u0 de y */
sdybu0<-sd(gy[1:conv,1])   /* Desvio padrao das estimativas u0 de y */
estyu0                     /* Imprime  media de u0 de y  */
sdybu0                     /* Imprime  desvio padrao de u0 de y  */

/* Calcular beta_1 pela formula aproximada da eq. (4.17) */
/* Precisa beta_0; beta_1; e sigma_u0 do modelo ajustado para y */
p<-inv.logit(estyb0+ estyu0)
a<-(1-tau1-tau0)*{p*(1-p)}
b<-{(1-tau1-tau0)*p +tau0}
c<-{(1-tau1-tau0)*(1-p)+tau1}
d<-estyb1
beta11<- (d*a)/(b*c)   /* Calcula beta1 do modelo corrigido */
beta11                /* Mostra o valor de beta11y corrigido*/
```

```
/*********************************************************************
* PROGRAMA: Simulacao para avaliar a magnitude do vies para resposta   *
*           ma classificada do modelo probit multinivel (4.3) e        *
*           para avaliar a qualidade da aproximacao (4.18) para beta1  *
*           diferente de zero.                                         *
* ********************************************************************/
/*******************************
*   Bibliotecas necessarias    *
*******************************/
library (MASS)
library (boot)
/*******************************
*  Parametros para simulacao   *
*******************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
b0<--1      /* Valor verdadeiro de beta0 */
b1<-1       /* Valor verdadeiro de beta1 */
R<-1000     /* Numero de replicas */
/*********************************************************
*                            Corpo                       *
*********************************************************/
/* Declaracao de variaveis */
dt<-matrix(0,R,2)
gt<-matrix(0,R,1)
dy<-matrix(0,R,2)
gy<-matrix(0,R,1)
/* Inicio do loop da simulacao */
for (r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x<-matrix(0,nj,j)
u<-matrix(0,nj,j)
T<-matrix(0,k,1)
test<-matrix(0,k,3)
ct<-matrix(0,2,1)
proby=matrix(0,k,1)
Y=matrix(0,k,1)
cy<-matrix(0,2,1)
test2<-matrix(0,k,3)
for(i in 1:j){
    x[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */

}
x1<-data.frame(x)  /* Converte a matriz x em dados */
b<-names(x1)       /* Mostra os nomes das variaveis dos dados x1 */
u1<-data.frame(u) /* Converte a matriz u em dados */
c<-names(u1)     /* Mostra os nomes das variaveis dos dados u1 */
for(i in 1:j){
    names(x1)[i]<-i /* Muda o nome da variavel para numeros */
    names(u1)[i]<-i /* Muda o nome da variavel para numeros */
}
vec1<-stack(x1) /* Empilha cada coluna de x1 em um vetor */
vec2<-stack(u1) /* Empilha cada coluna de u1 em um vetor */
x<-vec1[,1]     /* Armazena vec1 em x */
U<-vec2[,1]     /* Armazena vec2 em U */
J<-vec1[,2]     /* Obtem a coluna dos grupos */
X<-cbind(1,x)   /* Obtem a  matriz de regressores */
/* Geracao de eta com base nos valores verdadeiros dos parametros:
beta_0=-1; beta_1=1; e sigma_0u=1 */
eta<--1*X[,1]+1*X[,2]+U
```

```
/* Obter a probabilidade inicial em funcao de eta */
p<-pnorm(eta) /* Obtem a probabilidade inicial de t*/
/* Gerar T verdadeiro a partir da dist. binomial(1,tp1) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i])
}
/* Ajustar o MMLG para resposta sem correcao de
   erros T com ligacao probit via funcao glmmPQL  */
test[,1]<-T
test[,2]<-x
test[,3]<-J
dados<-data.frame(test)   /* Converte informacoes matriz num conj. dados */
for(i in 1:3){
    if (i==1){
    names(dados)[1]<-"t"  /* Muda o nome da primeira coluna para t */
    }
    if (i==2){
    names(dados)[2]="x1"  /* Muda o nome da segunda coluna para x1 */
    }
    if (i==3){
    names(dados)[3]="grup" /* Muda o nome da segunda coluna para grupo*/
    }
}
/* Ajute para a resposta t via procedimento QVP */
ajustet<-glmmPQL(t ~ x1, random = ~ 1  grup,
                family = binomial (link=probit),data=dados)
bt<-cbind(fixed.effects(ajustet)) /* Estimativas dos efeitos fixos */
ct[,1]<-bt                        /* Armazena as estimativas em um vetor */
dt[r,1]<-ct[1,1]                  /* Armazena a estimativa de beta0 */
dt[r,2]<-ct[2,1]                  /* Armazena a estimativa de beta1 */
et<-cbind(getVarCov(ajustet))     /* Estimativa do efeito aleatorio */
gt[r,1]<-et                       /* Armazena a estimativa do efeito aleatorio */
/* Obter a probabilidade de y com a correcao do erro a partir de t */
for (k in 1:k){
    proby[k]<-(1-tau1-tau0)*p[k]+tau0
}
/* Gerar y a partir da dist binomial(1,proby) */
for(w in 1: k){
    Y[w]<-rbinom(1,1,proby[w])
}
/* Ajustar um MMLG para a resposta y via procedimento QVP */
test2[,1]<-Y
test2[,2]<-x
test2[,3]<-J
dados2<-data,frame(test2) /* Transforma informacoes y banco dados */
for(i in 1:3){
    if (i==1){
    names(dados2)[1]<-"y"   /* Muda nome da coluna para y */
    }
    if (i==2){
    names(dados2)[2]="x2"   /* Muda nome da coluna para x2 */
    }
    if (i==3){
    names(dados2)[3]="grup2" /* Muda nome da coluna para grup */
    }
}
/* Ajuste para a resposta y via procedimento QVP */
ajustey<-glmmPQL(y ~ x2, random = ~ 1  grup2,
                family = binomial (link=probit),data=dados2)
by<-cbind(fixed.effects(ajustey)) /* Estimativas dos efeitos fixos */
cy[,1]<-by                        /* Armazena as estimativas em um vetor */
dy[r,1]<-cy[1,1]                  /* Armazena a estimativa de beta0 */
dy[r,2]<-cy[2,1]                  /* Armazena a estimativa de beta1 */
ey<-cbind(getVarCov(ajustey))     /* Estimativa do efeito aleatorio */
```

```
gy[r,1]<-ey                       /* Armazena a estimativa do efeito aleatorio */
} /* Final do loop */
/* Verificar o numero total de convergencia */
conv<-0  /* Declaracao da variavel conv */
for (r in 1:R){
  if(dt[r,1]!=0){
    conv<-conv+1  /* Conta o numero de convergencia */
  }
}
/* Impressao dos resultados */
nj     /* Imprime o tamanho da amostra no grupo */
tau0  /* Imprime o valor de tau0 */
tau1  /* Imprime o valor de tau0 */
conv   /* Imprime o total de convergencia */
esttb0<-mean(dt[1:conv,1])/* Media das estimativas beta0 de t */
sdtb0<-sd(dt[1:conv,1])   /* Desvio padrao das estimativas beta0 de t */
esttb0                    /* Imprime  media de beta0 de t  */
sdtb0                     /* Imprime  desvio padrao de beta0 de t */
esttb1<-mean(dt[1:conv,2])/* Media das estimativas beta1 de t*/
sdtb1<-sd(dt[1:conv,2])   /* Desvio padrao das estimativas beta1 de t */
esttb1                    /* Imprime  media de beta1 de t */
sdtb1                     /* Imprime  media de beta1  de t*/
esttu0<-mean(gt[1:conv,1])/* Media das estimativas u0 de t */
sdtbu0<-sd(gt[1:conv,1])  /* Desvio padrao das estimativas u0 de t */
esttu0                    /* Imprime  media de u0 de t */
sdtbu0                    /* Imprime  desvio padrao de u0 de t */

estyb0<-mean(dy[1:conv,1])/* Media das estimativas beta0 de y */
sdyb0<-sd(dy[1:conv,1])   /* Desvio padrao das estimativas beta0 de y */
vyb0<-estyb0-b0           /* Vies de beta0 de y */
estyb0                    /* Imprime  media de beta0 de y  */
sdyb0                     /* Imprime  desvio padrao de beta0 de y  */
vyb0                      /* Imprime  vies de beta0 de y  */
estyb1<-mean(dy[1:conv,2])/* Media das estimativas bet10 de y */
sdyb1<-sd(dy[1:conv,2])   /* Desvio padrao das estimativas beta1 de y */
vyb1<-estyb1-b1           /* Vies de beta01de y */
estyb1                    /* Imprime  media de beta1 de y  */
sdyb1                     /* Imprime  desvio padrao de beta1 de y  */
vyb1                      /* Imprime  vies de beta1 de y  */
estyu0<-mean(gy[1:conv,1])/* Media das estimativas u0 de y */
sdybu0<-sd(gy[1:conv,1])  /* Desvio padrao das estimativas u0 de y */
estyu0                    /* Imprime  media de u0 de y  */
sdybu0                    /* Imprime  desvio padrao de u0 de y  */

/* Calcular beta_1 pela formula aproximada da eq. (4.18) */
/* Precisa beta_0; beta_1 e sigma_u0 do modelo ajustado para y */
a<-pnorm( estyb0+ estyu0)
b<-1/pnorm(a)
c<-dnorm(b)
d<-(1-tau1-tau0)*c
e<-(1-tau1-tau0)*a+tau0
f<-1/pnorm(e)
g<-dnorm(f)
h<-esttb1
beta11<- (h*d)/g   /* Calcula beta1 do modelo corrigido */
beta11             /* Mostra o valor de beta11y corrigido*/
```

C.2 - Programa do segundo estudo de simulação.

```
/*************************************************************************
* PROGRAMA: Simulacao para avaliar a magnitude do vies para resposta     *
*           ma classificada do modelo logistico multinivel (5.1) e       *
*           para avaliar a qualidade da aproximacao (5.3) para beta1     *
*           diferente de zero.                                           *
*************************************************************************/
/*******************************
*   Bibliotecas necessarias    *
*******************************/
library (MASS)
library (boot)
/*******************************
*  Parametros para simulacao   *
*******************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
b0<--1      /* Valor verdadeiro de beta0 */
b1<-1       /* Valor verdadeiro de beta1 */
b2<-1       /* Valor verdadeiro de beta2 */
R<-1000     /* Numero de replicas */
/*********************************************************
*                           Corpo                        *
*********************************************************/
/* Declaracao de variaveis */
dt<-matrix(0,R,3)
gt<-matrix(0,R,2)
dy<-matrix(0,R,3)
gy<-matrix(0,R,2)
/* Inicio do loop da simulacao */
for (r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x1<-matrix(0,nj,j)
x2<-matrix(0,nj,j)
u0<-matrix(0,nj,j)
u1<-matrix(0,nj,j)
T<-matrix(0,k,1)
test<-matrix(0,k,4)
ct<-matrix(0,3,1)
proby=matrix(0,k,1)
Y=matrix(0,k,1)
cy<-matrix(0,3,1)
test2<-matrix(0,k,4)
for(i in 1:n){
    x1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    x2[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u0[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
}
a1<-data.frame(x1) /* Converte a matriz x1 em dados */
b1<-names(a1)      /* Mostra os nomes das variaveis dos dados a1 */
a2<-data.frame(x2) /* Converte a matriz x2 em dados */
b2<-names(a2)      /* Mostra os nomes das variaveis dos dados a2 */
a3<-data.frame(u0) /* Converte a matriz u0 em dados */
b3<-names(a3)      /* Mostra os nomes das variaveis dos dados a3 */
a4<-data.frame(u1) /* Converte a matriz u1 em dados */
b4<-names(a4)      /* Mostra os nomes das variaveis dos dados a4 */
for(i in 1:n){
    names(a1)[i]<-i /* Muda o nome da variavel para numeros */
    names(a2)[i]<-i /* Muda o nome da variavel para numeros */
```

```
    names(a3)[i]<-i /* Muda o nome da variavel para numeros */
    names(a4)[i]<-i/* Muda o nome da variavel para numeros */
}
vec1<-stack(a1)     /* Empilha cada coluna de a1 em um vetor */
vec2<-stack(a2)     /* Empilha cada coluna de a2 em um vetor */
vec3<-stack(a3)     /* Empilha cada coluna de a3 em um vetor */
vec4<-stack(a4)     /* Empilha cada coluna de x1 em um vetor */
x11<-vec1[,1]       /* Armazena vec1 em x11 */
x22<-vec2[,1]       /* Armazena vec2 em x22 */
u0j<-vec3[,1]       /* Armazena vec3 em u0j */
u1j<-vec4[,1]       /* Armazena vec4 em u1j */
J<-vec1[,2]         /* Obtem a coluna dos grupos */
X<-cbind(1,x11,x22) /* Obtem a  matriz de regressores */
/* Geracao de eta com base nos valores verdadeiros dos parametros:
beta_0=-1; beta_1=1; beta_2=1; sigma_0u=1 e sigma_1u=1 */
eta<--1*X[,1]+1*X[,2]+1*X[,3]+u0j+X[,2]*u1j
/* Obter a probabilidade inicial em funcao de eta */
p<-inv.logit(eta) /* Obtem a probabilidade inicial de t*/
/* Gerar T verdadeiro a partir da dist. binomial(1,p) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i]) #gera T verdadeiro a partir da dist. binomial(1,tp)
}
/* Ajustar o MMLG para resposta sem correcao de
   erros T com ligacao logit via funcao glmmPQL  */
test[,1]<-T
test[,2]<-x11
test[,3]<-x22
test[,4]<-J
dados<-data.frame(test) /* Converte informacoes matriz num conj. dados */
for(i in 1:4){
    if (i==1){
    names(dados)[1]<-"t" /* Muda o nome da primeira coluna para t */
    }
    if (i==2){
    names(dados)[2]="x1" /* Muda o nome da segunda coluna para x1 */
    }
    if (i==3){
    names(dados)[3]="x2" /* Muda o nome da terceira coluna para x2 */
    }
    if (i==4){
    names(dados)[4]="grup" /* Muda o nome da quarta coluna para grup */
    }
}
/* Ajute para a resposta t via procedimento QVP */
ajustet<-glmmPQL(t ~ x1+x2, random = ~ 1+x1  grup,
               family = binomial (link=logit),data=dados)
bt<-cbind(fixed.effects(ajustet))  /* Estimativas dos efeitos fixos */
ct[,1]<-bt                         /* Armazena as estimativas em um vetor */
dt[r,1]<-ct[1,1]                   /* Armazena a estimativa de beta0 */
dt[r,2]<-ct[2,1]                   /* Armazena a estimativa de beta1 */
dt[r,3]<-ct[3,1]                   /* Armazena a estimativa de beta2 */
et<-cbind(getVarCov(ajustet))      /* Estimativa do efeito aleatorio */
gt[r,1]<-et[1,1]                   /* Armazena a estimativa do efeito aleatorio u0j*/
gt[r,2]<-et[2,2]                   /* Armazena a estimativa do efeito aleatorio u1j*/
/* Obter a probabilidade de y com a correcao do erro a partir de t */
for (k in 1:k){
    proby[k]<-(1-tau1-tau0)*p[k]+tau0
}
/* Gerar y a partir da dist binomial(1,proby) */
for(w in 1: k){
    Y[w]<-rbinom(1,1,proby[w])
}
/* Ajustar um MMLG para a resposta y via procedimento QVP */
test2[,1]<-Y
```

```
test2[,2]<-x11
test2[,3]<-x22
test2[,4]<-J
dados2<-data.frame(test2)  /* Transforma informacoes y banco dados */
for(i in 1:4){
    if (i==1){
    names(dados2)[1]<-"y" /* Muda nome da coluna para y */
    }
    if (i==2){
    names(dados2)[2]="x1" /* Muda nome da coluna para x1 */
    }
    if (i==3){
    names(dados2)[3]="x2" /* Muda nome da coluna para x2 */
    }
    if (i==4){
    names(dados2)[4]="grup" /* Muda nome da coluna para grup */
    }
}
/* Ajuste para a resposta y via procedimento QVP */
ajustey<-glmmPQL(y ~ x1+x2, random = ~ 1+x1  grup,
               family = binomial (link=logit),data=dados2)
by<-cbind(fixed.effects(ajustey))/* Estimativas dos efeitos fixos */
cy[,1]<-by                       /* Armazena as estimativas em um vetor */
dy[r,1]<-cy[1,1]                 /* Armazena a estimativa de beta0 */
dy[r,2]<-cy[2,1]                 /* Armazena a estimativa de beta1 */
dy[r,3]<-cy[3,1]                 /* Armazena a estimativa de beta2 */
ey<-cbind(getVarCov(ajustey))    /* Estimativa do efeito aleatorio */
gy[r,1]<-ey[1,1]                 /* Armazena a estimativa do efeito aleatorio u0j */
gy[r,2]<-ey[2,2]                 /* Armazena a estimativa do efeito aleatorio u1j*/
}  /* Final do loop */
/* Verificar o numero total de convergencia */
conv<-0  /* Declaracao da variavel conv */
for (r in 1:R){
   if(dt[r,1]!=0){
    conv<-conv+1  /* Conta o numero de convergencia */
  }
}
/* Impressao dos resultados */
nj     /* Imprime o tamanho da amostra no grupo */
tau0  /* Imprime o valor de tau0 */
tau1 /* Imprime o valor de tau0 */
conv  /* Imprime o total de convergencia */

esttb0<-mean(dt[1:conv,1]) /* Media das estimativas beta0 de t */
sdtb0<-sd(dt[1:conv,1])    /* Desvio padrao das estimativas beta0 de t */
esttb0                     /* Imprime  media de beta0 de t  */
sdtb0                      /* Imprime  desvio padrao de beta0 de t */
esttb1<-mean(dt[1:conv,2]) /* Media das estimativas beta1 de t*/
sdtb1<-sd(dt[1:conv,2])    /* Desvio padrao das estimativas beta1 de t */
esttb1                     /* Imprime  media de beta1 de t */
sdtb1                      /* Imprime desvio padrao de beta1  de t*/
esttb2<-mean(dt[1:conv,3]) /* Media das estimativas beta2 de t*/
sdtb2<-sd(dt[1:conv,3])    /* Desvio padrao das estimativas beta2 de t */
esttb2                     /* Imprime  media de beta2 de t */
sdtb2                      /* Imprime desvio padrao de beta2  de t*/
esttu0<-mean(gt[1:conv,1]) /* Media das estimativas u0 de t */
sdtbu0<-sd(gt[1:conv,1])   /* Desvio padrao das estimativas u0 de t */
esttu0                     /* Imprime  media de u0 de t */
sdtbu0                     /* Imprime  desvio padrao de u0 de t */
esttu1<-mean(gt[1:conv,2])  /* Media das estimativas u1 de t */
sdtbu1<-sd(gt[1:conv,2])    /* Desvio padrao das estimativas u1 de t */
esttu1                      /* Imprime  media de u1 de t */
sdtbu1                      /* Imprime  desvio padrao de u1 de t */
```

```
estyb0<-mean(dy[1:conv,1])  /* Media das estimativas beta0 de y */
sdyb0<-sd(dy[1:conv,1])     /* Desvio padrao das estimativas beta0 de y */
vyb0<- estyb0-b0            /* Vies de beta0 de y */
estyb0                      /* Imprime  media de beta0 de y  */
sdyb0                       /* Imprime  desvio padrao de beta0 de y  */
vyb0                        /* Imprime  vies de beta0 de y  */
estyb1<-mean(dy[1:conv,2])  /* Media das estimativas beta1 de y */
sdyb1<-sd(dy[1:conv,2])     /* Desvio padrao das estimativas beta1 de y */
vyb1<- estyb1-b1            /* Vies de beta1 de y */
estyb1                      /* Imprime  media de beta1 de y  */
sdyb1                       /* Imprime  desvio padrao de beta1 de y  */
vyb1                        /* Imprime  vies de beta1 de y  */
estyb2<-mean(dy[1:conv,3])  /* Media das estimativas beta2 de y */
sdyb2<-sd(dy[1:conv,3])     /* Desvio padrao das estimativas beta2 de y */
vyb2<- estyb2-b2            /* Vies de beta2 de y */
estyb2                      /* Imprime  media de beta2 de y  */
sdyb2                       /* Imprime  desvio padrao de beta2 de y  */
vyb2                        /* Imprime  vies de beta2 de y  */
estyu0<-mean(gy[1:conv,1])  /* Media das estimativas u0 de y */
sdybu0<-sd(gy[1:conv,1])    /* Desvio padrao das estimativas u0 de y */
estyu0                      /* Imprime  media de u0 de y  */
sdybu0                      /* Imprime  desvio padrao de u1 de y  */
estyu1<-mean(gy[1:conv,2])  /* Media das estimativas u1 de y */
sdybu1<-sd(gy[1:conv,2])    /* Desvio padrao das estimativas u1 de y */
estyu1                      /* Imprime  media de u1 de y  */
sdybu1                      /* Imprime  desvio padrao de u1 de y  */


/* Calcular beta_1 pela formula aproximada da eq. (5.3) */
/* Precisa beta_0; beta_1; beta_2; sigma_u0 e sigma_u1 do modelo ajustado para y */
p<-inv.logit(estyb0  + estyu0 + estyu1)
a<-(1-tau1-tau0)*{p*(1-p)}
b<-{(1-tau1-tau0)*p +tau0}
c<-{(1-tau1-tau0)*(1-p)+tau1}
d<- estyb1
e<- estyb2
beta11<- (d*a)/(b*c) /* Calcula beta1 do modelo corrigido */
beta11               /* Mostra o valor de beta11y corrigido*/
beta22<- (e*a)/(b*c) /* Calcula beta2 do modelo corrigido */
beta22               /* Mostra o valor de beta22y corrigido*/


/*********************************************************************
* PROGRAMA: Simulacao para avaliar a magnitude do vies para resposta   *
*           ma classificada do modelo probit multinivel (5.2) e        *
*           para avaliar a qualidade da aproximacao (5.4) para beta1   *
*           diferente de zero.                                         *
* *******************************************************************/
/*******************************
*   Bibliotecas necessarias    *
*******************************/
library (MASS)
library (boot)
/*******************************
*  Parametros para simulacao   *
*******************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
b0<--1      /* Valor verdadeiro de beta0 */
b1<-1       /* Valor verdadeiro de beta1 */
b2<-1       /* Valor verdadeiro de beta2 */
```

```
R<-1000     /* Numero de replicas */
/*********************************************************
*                            Corpo                       *
*********************************************************/
/* Declaracao de variaveis */
dt<-matrix(0,R,3)
gt<-matrix(0,R,2)
dy<-matrix(0,R,3)
gy<-matrix(0,R,2)
/* Inicio do loop da simulacao */
for (r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x1<-matrix(0,nj,j)
x2<-matrix(0,nj,j)
u0<-matrix(0,nj,j)
u1<-matrix(0,nj,j)
T<-matrix(0,k,1)
test<-matrix(0,k,4)
ct<-matrix(0,3,1)
proby=matrix(0,k,1)
Y=matrix(0,k,1)
cy<-matrix(0,3,1)
test2<-matrix(0,k,4)
for(i in 1:n){
    x1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    x2[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u0[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
}
a1<-data.frame(x1) /* Converte a matriz x1 em dados */
b1<-names(a1)      /* Mostra os nomes das variaveis dos dados a1 */
a2<-data.frame(x2) /* Converte a matriz x2 em dados */
b2<-names(a2)      /* Mostra os nomes das variaveis dos dados a2 */
a3<-data.frame(u0) /* Converte a matriz u0 em dados */
b3<-names(a3)      /* Mostra os nomes das variaveis dos dados a3 */
a4<-data.frame(u1) /* Converte a matriz u1 em dados */
b4<-names(a4)      /* Mostra os nomes das variaveis dos dados a4 */
for(i in 1:n){
    names(a1)[i]<-i /* Muda o nome da variavel para numeros */
    names(a2)[i]<-i /* Muda o nome da variavel para numeros */
    names(a3)[i]<-i /* Muda o nome da variavel para numeros */
    names(a4)[i]<-i/* Muda o nome da variavel para numeros */
}
vec1<-stack(a1)     /* Empilha cada coluna de a1 em um vetor */
vec2<-stack(a2)     /* Empilha cada coluna de a2 em um vetor */
vec3<-stack(a3)     /* Empilha cada coluna de a3 em um vetor */
vec4<-stack(a4)     /* Empilha cada coluna de x1 em um vetor */
x11<-vec1[,1]       /* Armazena vec1 em x11 */
x22<-vec2[,1]       /* Armazena vec2 em x22 */
u0j<-vec3[,1]       /* Armazena vec3 em u0j */
u1j<-vec4[,1]       /* Armazena vec4 em u1j */
J<-vec1[,2]         /* Obtem a coluna dos grupos */
X<-cbind(1,x11,x22) /* Obtem a  matriz de regressores */
/* Geracao de eta com base nos valores verdadeiros dos parametros:
beta_0=-1; beta_1=1; beta_2=1; sigma_0u=1 e sigma_1u=1 */
eta<--1*X[,1]+1*X[,2]+1*X[,3]+u0j+X[,2]*u1j
/* Obter a probabilidade inicial em funcao de eta */
p<-pnorm(eta) /* Obtem a probabilidade inicial de t*/
/* Gerar T verdadeiro a partir da dist. binomial(1,p) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i])
}
/* Ajustar o MMLG para resposta sem correcao de
   erros T com ligacao probit via funcao glmmPQL  */
```

```
test[,1]<-T
test[,2]<-x11
test[,3]<-x22
test[,4]<-J
dados<-data.frame(test)  /* Converte informacoes matriz num conj. dados */
for(i in 1:4){
    if (i==1){
    names(dados)[1]<-"t" /* Muda o nome da primeira coluna para t */
    }
    if (i==2){
    names(dados)[2]="x1" /* Muda o nome da segunda coluna para x1 */
    }
    if (i==3){
    names(dados)[3]="x2" /* Muda o nome da terceira coluna para x2 */
    }
    if (i==4){
    names(dados)[4]="grup" /* Muda o nome da quarta coluna para grup */
    }
}
/* Ajute para a resposta t via procedimento QVP */
ajustet<-glmmPQL(t ~ x1+x2, random = ~ 1+x1  grup,
                family = binomial (link=probit),data=dados)

bt<-cbind(fixed.effects(ajustet)) /* Estimativas dos efeitos fixos */
ct[,1]<-bt                        /* Armazena as estimativas em um vetor */
dt[r,1]<-ct[1,1]                  /* Armazena a estimativa de beta0 */
dt[r,2]<-ct[2,1]                  /* Armazena a estimativa de beta1 */
dt[r,3]<-ct[3,1]                  /* Armazena a estimativa de beta2 */
et<-cbind(getVarCov(ajustet))     /* Estimativa do efeito aleatorio */
gt[r,1]<-et[1,1]                  /* Armazena a estimativa do efeito aleatorio u0j*/
gt[r,2]<-et[2,2]                  /* Armazena a estimativa do efeito aleatorio u1j*/
/* Obter a probabilidade de y com a correcao do erro a partir de t */
for (k in 1:k){
    proby[k]<-(1-tau1-tau0)*p[k]+tau0
}
/* Gerar y a partir da dist binomial(1,proby) */
for(w in 1: k){
    Y[w]<-rbinom(1,1,proby[w])
}
/* Ajustar um MMLG para a resposta y via procedimento QVP */
test2[,1]<-Y
test2[,2]<-x11
test2[,3]<-x22
test2[,4]<-J
dados2<-data.frame(test2)  /* Transforma informacoes y banco dados */
for(i in 1:4){
    if (i==1){
    names(dados2)[1]<-"y" /* Muda nome da coluna para y */
    }
    if (i==2){
    names(dados2)[2]="x1" /* Muda nome da coluna para x1 */
    }
    if (i==3){
    names(dados2)[3]="x2" /* Muda nome da coluna para x2 */
    }
    if (i==4){
    names(dados2)[4]="grup" /* Muda nome da coluna para grup */
    }
}
/* Ajuste para a resposta y via procedimento QVP */
ajustey<-glmmPQL(y ~ x1+x2, random = ~ 1+x1  grup,
                family = binomial (link=probit),data=dados2)
by<-cbind(fixed.effects(ajustey))/* Estimativas dos efeitos fixos */
cy[,1]<-by                       /* Armazena as estimativas em um vetor */
```

```
dy[r,1]<-cy[1,1]                /* Armazena a estimativa de beta0 */
dy[r,2]<-cy[2,1]                /* Armazena a estimativa de beta1 */
dy[r,3]<-cy[3,1]                /* Armazena a estimativa de beta2 */
ey<-cbind(getVarCov(ajustey))   /* Estimativa do efeito aleatorio */
gy[r,1]<-ey[1,1]                /* Armazena a estimativa do efeito aleatorio u0j */
gy[r,2]<-ey[2,2]                /* Armazena a estimativa do efeito aleatorio u1j*/
}  /* Final do loop */
/* Verificar o numero total de convergencia */
conv<-0  /* Declaracao da variavel conv */
for (r in 1:R){
   if(dt[r,1]!=0){
    conv<-conv+1  /* Conta o numero de convergencia */
  }
}
/* Impressao dos resultados */
nj     /* Imprime o tamanho da amostra no grupo */
tau0   /* Imprime o valor de tau0 */
tau1  /* Imprime o valor de tau1 */
conv  /* Imprime o total de convergencia */

esttb0<-mean(dt[1:conv,1]) /* Media das estimativas beta0 de t */
sdtb0<-sd(dt[1:conv,1])    /* Desvio padrao das estimativas beta0 de t */
esttb0                     /* Imprime  media de beta0 de t  */
sdtb0                      /* Imprime  desvio padrao de beta0 de t */
esttb1<-mean(dt[1:conv,2]) /* Media das estimativas beta1 de t*/
sdtb1<-sd(dt[1:conv,2])    /* Desvio padrao das estimativas beta1 de t */
esttb1                     /* Imprime  media de beta1 de t */
sdtb1                      /* Imprime desvio padrao de beta1  de t*/
esttb2<-mean(dt[1:conv,3]) /* Media das estimativas beta2 de t*/
sdtb2<-sd(dt[1:conv,3])    /* Desvio padrao das estimativas beta2 de t */
esttb2                     /* Imprime  media de beta2 de t */
sdtb2                      /* Imprime desvio padrao de beta2  de t*/
esttu0<-mean(gt[1:conv,1]) /* Media das estimativas u0 de t */
sdtbu0<-sd(gt[1:conv,1])   /* Desvio padrao das estimativas u0 de t */
esttu0                     /* Imprime  media de u0 de t */
sdtbu0                     /* Imprime  desvio padrao de u0 de t */
esttu1<-mean(gt[1:conv,2]) /* Media das estimativas u1 de t */
sdtbu1<-sd(gt[1:conv,2])   /* Desvio padrao das estimativas u1 de t */
esttu1                     /* Imprime  media de u1 de t */
sdtbu1                     /* Imprime  desvio padrao de u1 de t */

estyb0<-mean(dy[1:conv,1])  /* Media das estimativas beta0 de y */
sdyb0<-sd(dy[1:conv,1])     /* Desvio padrao das estimativas beta0 de y */
vyb0<- estyb0-b0            /* Vies de beta0 de y */
estyb0                      /* Imprime  media de beta0 de y  */
sdyb0                       /* Imprime  desvio padrao de beta0 de y  */
vyb0                        /* Imprime  vies de beta0 de y  */
estyb1<-mean(dy[1:conv,2])  /* Media das estimativas beta1 de y */
sdyb1<-sd(dy[1:conv,2])     /* Desvio padrao das estimativas beta1 de y */
vyb1<- estyb1-b1            /* Vies de beta1 de y */
estyb1                      /* Imprime  media de beta1 de y  */
sdyb1                       /* Imprime  desvio padrao de beta1 de y  */
vyb1                        /* Imprime  vies de beta1 de y  */
estyb2<-mean(dy[1:conv,3])  /* Media das estimativas beta2 de y */
sdyb2<-sd(dy[1:conv,3])     /* Desvio padrao das estimativas beta2 de y */
vyb2<- estyb2-b2            /* Vies de beta2 de y */
estyb2                      /* Imprime  media de beta2 de y  */
sdyb2                       /* Imprime  desvio padrao de beta2 de y  */
vyb2                        /* Imprime  vies de beta2 de y  */
estyu0<-mean(gy[1:conv,1])  /* Media das estimativas u0 de y */
sdybu0<-sd(gy[1:conv,1])    /* Desvio padrao das estimativas u0 de y */
estyu0                      /* Imprime  media de u0 de y  */
sdybu0                      /* Imprime  desvio padrao de u1 de y  */
estyu1<-mean(gy[1:conv,2])  /* Media das estimativas u1 de y */
```

```
sdybu1<-sd(gy[1:conv,2])    /* Desvio padrao das estimativas u1 de y */
estyu1                      /* Imprime  media de u1 de y  */
sdybu1                      /* Imprime  desvio padrao de u1 de y  */

/* Calcular beta_1 pela formula aproximada da eq. (5.4) */
/* Precisa beta_0; beta_1; beta_2; sigma_u0 e sigma_u1 do modelo ajustado para y */
a<-pnorm( estyb0 +estyu0+estyu1)
b<-1/pnorm(a)
c<-dnorm(b)
d<-(1-tau1-tau0)*c
e<-(1-tau1-tau0)*a+tau0
f<-1/pnorm(e)
g<-dnorm(f)
h<- estyb1
i<- estyb2
beta11<- (h*d)/g /* Calcula beta1 do modelo corrigido */
beta11           /* Mostra o valor de beta11y */
beta22<- (i*d)/g /* Calcula beta2 do modelo corrigido */
beta22           /* Mostra o valor de beta22y */
```

C.3 - Programa do terceiro estudo de simulação.

```
 /*************************************************************************
* PROGRAMA: Simulacao para estimar as probabilidades dos erros tau0 e    *
*           tau1, bem como os parametros do modelo logistico hierarquico *
*           (4.2) utilizando uma parametrizacao para os componetes       *
*            de variancia.                                               *
* ************************************************************************/
/*****************************
*   Bibliotecas necessarias   *
*****************************/
library(boot)
library(gnlm)
library(repeated)
/*****************************
*  Parametros para simulacao   *
*****************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
sens=0.95   /* Sensibilidade */
espec=0.95  /* Especificidade */
b0<--1      /* Valor verdadeiro de beta0 */
b1<-1       /* Valor verdadeiro de beta1 */
sig<-1      /* Valor verdadeiro de sigma_u0*/
R<-1000     /* Numero de replicas */
/*****************************************************
*                           Corpo                    *
*****************************************************/
/* Declaracao de variaveis */
l_t0<-log(tau0/(1-tau0))
l_t1<-log(tau1/(1-tau1))
l_t0
l_t1
l_sig<-log(sig)
d<-matrix(0,R,5)
/* Inicio do loop da simulacao */
for(r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x<-matrix(0,nj,j)
u<-matrix(0,nj,j)
T<-matrix(0,k,1)
y<-matrix(0,k,1)
for(i in 1:j){
    x[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
}
x1<-data.frame(x)   /* Converte a matriz x em dados */
b<-names(x1)        /* Mostra os nomes das variaveis dos dados x1 */
u0<-data.frame(u)   /* Converte a matriz u0 em dados */
c<-names(u0)        /* Mostra os nomes das variaveis dos dados u0 */
for(i in 1:nj){
    names(a1)[i]<-i /* Muda o nome da variavel para numeros */
    names(u0)[i]<-i /* Muda o nome da variavel para numeros */
}
vec1<-stack(x1)     /* Empilha cada coluna de x1 em um vetor */
vec2<-stack(u0)     /* Empilha cada coluna de u0 em um vetor */
x11<-vec1[,1]       /* Armazena vec1 em x11 */
U<-vec2[,1]         /* Armazena vec2 em U */
J<-vec1[,2]         /* Obtem a coluna dos grupos */
X<-cbind(1,x11)     /* Obtem a  matriz de regressores */
```

```
/* Geracao de eta com base nos valores verdadeiros dos parametros:
beta_0=-1; beta_1=1 e sigma_0u=1 */
eta<--1*X[,1]+1*X[,2]+1*U

/* Obter a probabilidade inicial em funcao de eta */
p<-inv.logit(eta)

/* Gerar T verdadeiro a partir da dist. binomial(1,p) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i])
}

/* Gerar y a partir de T */
for(i in 1: k){
    if (T[i]==1){
       y[i]<-rbinom(1,1,sens)
    }
    else{
       y[i]<-1-rbinom(1,1,espec)
    }
}
/* Ajustar um MMNLG para a resposta y via procedimento
Quadratura de Gauss-Hermite */
mu<-function (p) (1-(exp(p[4])/(1+exp(p[4])))-(exp(p[5])/(1+exp(p[5]))))
                 *(exp(p[1]+p[2]*x11+exp(p[3])*u)/(1+exp(p[1]+p[2]*x11+exp(p[3])*u)))
                 +(exp(p[4])/(1+exp(p[4])))
print(z<-gnlr(y, dist="binomial", mu=mu, pmu=c(b0,b1,l_sig,l_t0,l_t1), pshape=1))
/* Ajuste para a resposta y via procedimento Quadratura de Gauss-Hermite */
ajustey<-gnlmm(y, dist="binomial", mu=mu, nest=J, pmu=z$coef[1:5],
               pshape=z$coef[6], psd=1, points=5)
b<-rbind(coef(ajustey)) /* Estimativas dos coeficientes e dos erros */
d[r,]<-b[1:5]           /* Armazena as estimativas em um vetor */
} /* Final do loop */
/* Verificar o numero total de convergencia */
conv<-0          /* Declaracao da variavel conv */
for (r in 1:R){
  if(d[r,1]!=0){
    conv<-conv+1 /* Conta o numero de convergencia */
  }
}
/* Impressao dos resultados */
nj     /* Imprime o tamanho da amostra no grupo */
tau0   /* Imprime o valor de tau0 */
tau1   /* Imprime o valor de tau1 */
conv    /* Imprime o total de convergencia */

b0<-summary(d[1:conv,1]) /* Descritiva das estimativas beta0 de y */
b0                       /* Imprime a descritiva das est. de beta0 de y*/
b1<-summary(d[1:conv,2]) /* Descritiva das estimativas beta1 de y */
b1                       /* Imprime a descritiva das est. de beta1 de y*/

lu0<-cbind(summary(d[1:conv,3]))/* Descritiva das estimativas l_sig de y */
lt0<-cbind(summary(d[1:conv,4]))/* Descritiva das estimativas l_t0 de y */
lt1<-cbind(summary(d[1:conv,5]))/* Descritiva das estimativas l_t1 de y */

u0<-exp(lu0[3,1])                    /* Mediana da estimativa u0 de y*/
u0                                   /* Imprime mediana de u0 de y */
infu0<-exp(lu0[2,1])                 /* Primeiro quartil de u0 de y*/
infu0                                /* Imprime primeiro quartil de u0 de y*/
supu0<-exp(lu0[5,1])                 /* Terceiro quartil de u0 de y*/
supu0                                /* Imprime terceiro quartil de u0 de y*/

t0<-exp(lt0[3,1])/(1+exp(lt0[3,1]))  /* Mediana da estimativa t0 de y*/
t0                                   /* Imprime mediana de t0 de y */
```

```
inft0<-exp(lt0[2,1])/(1+exp(lt0[2,1])) /* Primeiro quartil de t0 de y*/
inft0                                  /* Imprime primeiro quartil de t0 de y*/
supt0<-exp(lt0[5,1])/(1+exp(lt0[5,1])) /* Terceiro quartil de t0 de y*/
supt0                                  /* Imprime terceiro quartil de t0 de y*/

t1<-exp(lt1[3,1])/(1+exp(lt1[3,1]))    /* Mediana da estimativa t1 de y*/
t1                                     /* Imprime mediana de t1 de y */
inft1<-exp(lt1[2,1])/(1+exp(lt1[2,1]))  /* Primeiro quartil de t1 de y*/
inft1                                  /* Imprime primeiro quartil de t1 de y*/
supt1<-exp(lt1[5,1])/(1+exp(lt1[5,1]))  /* Terceiro quartil de t1 de y*/
supt1                                  /* Imprime terceiro quartil de t1 de y*/


 /*************************************************************************
* PROGRAMA: Simulacao para estimar as probabilidades dos erros tau0 e     *
*           tau1, bem como os parametros do modelo probit hierarquico     *
*           (4.3) utilizando uma parametrizacao para os componetes        *
*            de variancia.                                                *
* *************************************************************************/
/*******************************
*   Bibliotecas necessarias    *
*******************************/
library(boot)
library(gnlm)
library(repeated)
/*******************************
*  Parametros para simulacao   *
*******************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
sens=0.95   /* Sensibilidade */
espec=0.95  /* Especificidade */
b0<--1      /* Valor verdadeiro de beta0 */
b1<-1       /* Valor verdadeiro de beta1 */
sig<-1      /* Valor verdadeiro de sigma_u0*/
R<-1000     /* Numero de replicas */
/*****************************************************
*                           Corpo
*****************************************************/
/* Declaracao de variaveis */
d<-matrix(0,R,5)
/* Inicio do loop da simulacao */
for(r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x<-matrix(0,nj,j)
u<-matrix(0,nj,j)
T<-matrix(0,k,1)
y<-matrix(0,k,1)
for(i in 1:n){
    x[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
}
x1<-data.frame(x) /* Converte a matriz x em dados */
b<-names(x1) /* Mostra os nomes das variaveis dos dados x1 */
u0<-data.frame(u) /* Converte a matriz u0 em dados */
c<-names(u0)  /* Mostra os nomes das variaveis dos dados u0 */
for(i in 1:n){
    names(x1)[i]<-i /* Muda o nome da variavel para numeros */
    names(u0)[i]<-i /* Muda o nome da variavel para numeros */
}
vec1<-stack(x1)  /* Empilha cada coluna de x1 em um vetor */
```

```
vec2<-stack(u0)  /* Empilha cada coluna de u0 em um vetor */
x11<-vec1[,1]    /* Armazena vec1 em x11 */
U<-vec2[,1]      /* Armazena vec2 em U */
J<-vec1[,2]      /* Obtem a coluna dos grupos */
X<-cbind(1,x11)  /* Obtem a  matriz de regressores */
/* Geracao de eta com base nos valores verdadeiros dos parametros:
beta_0=-1; beta_1=1 e sigma_0u=1 */
eta<--1*X[,1]+1*X[,2]+1*U
/* Obter a probabilidade inicial em funcao de eta */
p<-pnorm(eta)
/* Gerar T verdadeiro a partir da dist. binomial(1,p) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i])
}
/* Gerar y a partir de T */
for(i in 1: k){
    if (T[i]==1){
       y[i]<-rbinom(1,1,sens)
    }
    else{
       y[i]<-1-rbinom(1,1,espec)
    }
}
 /* Ajustar um MMNLG para a resposta y via procedimento
Quadratura de Gauss-Hermite */
mu<-function (p) (1-p[4]-p[5])*(pnorm(p[1]+p[2]*x+p[3]*u))+p[4]
print(z<-gnlr(y, dist="binomial", mu=mu, pmu=c(-1,1,sig,tau0,tau1), pshape=1))
/* Ajuste para a resposta y via procedimento Quadratura de Gauss-Hermite */
ajustey<-gnlmm(y, dist="binomial", mu=mu, nest=J, pmu=z$coef[1:5],
               pshape=z$coef[6], psd=1, points=5)
b<-rbind(coef(ajustey))/* Estimativas dos coeficientes e dos erros */
d[r,]<-b[1:5]          /* Armazena as estimativas em um vetor */
}/* Final do loop */
/* Verificar o numero total de convergencia */
conv<-0   /* Declaracao da variavel conv */
for (r in 1:R){
  if(d[r,1]!=0){
    conv<-conv+1 /* Conta o numero de convergencia */
  }
}

/* Impressao dos resultados */

nj     /* Imprime o tamanho da amostra no grupo */
tau0   /* Imprime o valor de tau0 */
tau1   /* Imprime o valor de tau1 */
conv    /* Imprime o total de convergencia */

b0<-summary(d[1:conv,1]) /* Descritiva das estimativas beta0 de y */
b0                       /* Imprime a descritiva das est. de beta0 de y*/
b1<-summary(d[1:conv,2]) /* Descritiva das estimativas beta1 de y */
b1                       /* Imprime a descritiva das est. de beta1 de y*/

u0<-(summary(d[1:conv,3]))/* Descritiva das estimativas u0 de y */
t0<-(summary(d[1:conv,4]))/* Descritiva das estimativas t0 de y */
t1<-(summary(d[1:conv,5]))/* Descritiva das estimativas t1 de y */
u0                        /* Imprime a descritiva de u0 de y */
t0                        /* Imprime a descritiva de t0 de y */
t1                        /* Imprime a descritiva de t1 de y */
```

C.4 - Programa do quarto estudo de simulação.

```
/*************************************************************************
* PROGRAMA: Simulacao para estimar as probabilidades dos erros tau0 e    *
*           tau1, bem como os parametros do modelo logistico hierarquico *
*           (5.2) utilizando uma parametrizacao para os componetes       *
*           de variancia.                                                 *
* ***********************************************************************/
/*******************************
*   Bibliotecas necessarias
*******************************/
library(boot)
library(gnlm)
library(repeated)
/*******************************
*  Parametros para simulacao
*******************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
/*1-gamma0=spec*/
/*1-gamma1=sens*/
sens<-1-gama1   /* Sensibilidade */
espec<-1-gama0  /* Especificidade */
b0<--1          /* Valor verdadeiro de beta0 */
b1<-1          /* Valor verdadeiro de beta1 */
b2<-1          /* Valor verdadeiro de beta2 */
sigu0<-1        /* Valor verdadeiro de sigma_u0*/
sigu1<-1        /* Valor verdadeiro de sigma_u1*/
R<-1000       /* Numero de replicas */
/*****************************************************
*                            Corpo
*****************************************************/
/* Declaracao de variaveis */
l_t0<-log(tau0/(1-tau0))
l_t1<-log(tau1/(1-tau1))
l_t0
l_t1
l_sigu0<-log(sigu0)
l_sigu1<-log(sigu1)
d<-matrix(0,R,7)
/* Inicio do loop da simulacao */
for(r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x1<-matrix(0,nj,j)
x2<-matrix(0,nj,j)
u0<-matrix(0,nj,j)
u1<-matrix(0,nj,j)
T<-matrix(0,k,1)
y<-matrix(0,k,1)
for(i in 1:j){
    x1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    x2[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u0[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
}
x11<-data.frame(x1) /* Converte a matriz x1 em dados */
x22<-data.frame(x2) /* Converte a matriz x2 em dados */
b11<-names(x11) /* Mostra os nomes das variaveis dos dados x1 */
b22<-names(x22)  /* Mostra os nomes das variaveis dos dados x2 */
u00<-data.frame(u0) /* Converte a matriz u0 em dados */
u11<-data.frame(u1)  /* Converte a matriz u1 em dados */
```

```
c00<-names(u0)       /* Mostra os nomes das variaveis dos dados u0 */
c11<-names(u1)    /* Mostra os nomes das variaveis dos dados u1 */
for(i in 1:nj){
    names(x11)[i]<-i /* Muda o nome da variavel para numeros */
    names(x22)[i]<-i /* Muda o nome da variavel para numeros */
    names(u00)[i]<-i /* Muda o nome da variavel para numeros */
    names(u11)[i]<-i /* Muda o nome da variavel para numeros */
}
vec11<-stack(x11) /* Empilha cada coluna de x11 em um vetor */
vec12<-stack(x22) /* Empilha cada coluna de x22 em um vetor */
vec21<-stack(u00) /* Empilha cada coluna de u00 em um vetor */
vec22<-stack(u11)  /* Empilha cada coluna de u11 em um vetor */
x1<-vec11[,1]      /* Armazena vec1 em x1 */
x2<-vec12[,1]      /* Armazena vec1 em x2 */
U0<-vec21[,1]      /* Armazena vec1 em U0 */
U1<-vec22[,1]       /* Armazena vec1 em U1 */
J<-vec11[,2]         /* Obtem a coluna dos grupos */
X<-cbind(1,x1,x2)  /* Obtem a  matriz de regressores */
/* Geracao de eta com base nos valores verdadeiros dos parametros: */
/* beta_0=-1; beta_1=1; beta_2=1 e sigma_u0=sigma_u1 */
eta<--1*X[,1]+1*X[,2]+1*X[,3]+U0+X[,2]*U1
/* Obter a probabilidade inicial em funcao de eta */
p<-inv.logit(eta)
/* Gerar T verdadeiro a partir da dist. binomial(1,p) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i])
}
/* Gerar y a partir de T */
for(i in 1: k){
    if (T[i]==1){
       y[i]<-rbinom(1,1,sens)
    }
    else{
       y[i]<-1-rbinom(1,1,espec)
    }
}
/* Estimar os erros ajustando-se um MMNLG para a resposta y via procedimento
Quadratura de Gauss-Hermite */
mu<-function (p) (1-(exp(p[6])/(1+exp(p[6])))-(exp(p[7])/(1+exp(p[7]))))*
             ((exp(p[1]+p[2]*x1+p[3]*x2+exp(p[4])*U0+exp(p[5])*U1*x1)/
             (1+exp(p[1]+p[2]*x1+p[3]*x2+exp(p[4])*U0+exp(p[5])*U1*x1))))+
              (exp(p[7])/(1+exp(p[7])))

print(z<-gnlr(y, dist="binomial", mu=mu, pmu=c(b0,b1,b2,l_sigu0,l_sigu1,l_t1,l_t0), pshape=1))
/* Ajuste para a resposta y via procedimento Quadratura de Gauss-Hermite */
ajustey<-gnlmm(y, dist="binomial", mu=mu, nest=J, pmu=z$coef[1:7],
               pshape=z$coef[8], psd=1, points=5)
b<-rbind(coef(ajustey)) /* Estimativas dos coeficientes e dos erros */
d[r,]<-b[1:7]           /* Armazena as estimativas em um vetor */
} /* Final do loop */

/* Verificar o numero total de convergencia */
conv<-0          /* Declaracao da variavel conv */
for (r in 1:R){
  if(d[r,1]!=0){
    conv<-conv+1 /* Conta o numero de convergencia */
  }
}

/* Impressao dos resultados */
nj     /* Imprime o tamanho da amostra no grupo */
tau0   /* Imprime o valor de tau0 */
tau1   /* Imprime o valor de tau1 */
conv    /* Imprime o total de convergencia */
```

```
b0<-summary(d[1:conv,1])/* Descritiva das estimativas beta0 de y */
b0                      /* Imprime a descritiva das est. de beta0 de y*/

b1<-summary(d[1:conv,2]) /* Descritiva das estimativas beta1 de y */
b1                       /* Imprime a descritiva das est. de beta1 de y*/

b2<-summary(d[1:conv,3]) /* Descritiva das estimativas beta2 de y */
b2                        /* Imprime a descritiva das est. de beta2 de y*/

lu0<-cbind(summary(d[1:conv,4]))/* Descritiva das estimativas l_sigu0 de y */
lu1<-cbind(summary(d[1:conv,5]))/* Descritiva das estimativas l_sigu1 de y */

lt1<-cbind(summary(d[1:conv,6]))/* Descritiva das estimativas l_t1 de y */
lt0<-cbind(summary(d[1:conv,7]))/* Descritiva das estimativas l_t0 de y */

u0<-exp(lu0[3,1])                    /* Mediana da estimativa u0 de y*/
u0                                  /* Imprime mediana de u0 de y */
infu0<-exp(lu0[2,1])                 /* Primeiro quartil de u0 de y*/
infu0                                /* Imprime primeiro quartil de u0 de y*/
supu0<-exp(lu0[5,1])                  /* Terceiro quartil de u0 de y*/
supu0

u1<-exp(lu1[3,1])                    /* Mediana da estimativa u1 de y*/
u1                                   /* Imprime mediana de u1 de y */
infu1<-exp(lu1[2,1])                 /* Primeiro quartil de u1 de y*/
infu1                                /* Imprime primeiro quartil de u1 de y*/
supu1<-exp(lu1[5,1])                 /* Terceiro quartil de u1 de y*/
supu1                                /* Imprime terceiro quartil de u1 de y*/

t0<-exp(lt0[3,1])/(1+exp(lt0[3,1]))   /* Mediana da estimativa t0 de y*/
t0                                   /* Imprime mediana de t0 de y */
inft0<-exp(lt0[2,1])/(1+exp(lt0[2,1])) /* Primeiro quartil de t0 de y*/
inft0                                 /* Imprime primeiro quartil de t0 de y*/
supt0<-exp(lt0[5,1])/(1+exp(lt0[5,1])) /* Terceiro quartil de t0 de y*/
supt0                                /* Imprime terceiro quartil de t0 de y*/

t1<-exp(lt1[3,1])/(1+exp(lt1[3,1]))  /* Mediana da estimativa t1 de y*/
t1                                  /* Imprime mediana de t1 de y */
inft1<-exp(lt1[2,1])/(1+exp(lt1[2,1]))/* Primeiro quartil de t1 de y*/
inft1                                /* Imprime primeiro quartil de t1 de y*/
supt1<-exp(lt1[5,1])/(1+exp(lt1[5,1]))/* Terceiro quartil de t1 de y*/
supt1                                /* Imprime terceiro quartil de t1 de y*/


/*************************************************************************
* PROGRAMA: Simulacao para estimar as probabilidades dos erros tau0 e    *
*           tau1, bem como os parametros do modelo probit hierarquico    *
*           (5.3) utilizando uma parametrizacao para os componetes       *
*            de variancia.                                               *
* ************************************************************************/
/*******************************
*   Bibliotecas necessarias    *
*******************************/
library(boot)
library(gnlm)
library(repeated)
/*******************************
*  Parametros para simulacao   *
*******************************/
j<-200      /* Numero de grupos */
nj<-5       /* Tamanho da amostra por grupo */
k<-nj*j     /* Numero total de observacoes */
tau0=0.05  /* Probabilidade de y=1 dado que t=0 */
```

```
tau1=0.05  /* Probabilidade de y=0 dado que t=1 */
/*1-gamma0=spec*/
/*1-gamma1=sens*/
sens<-1-gama1   /* Sensibilidade */
espec<-1-gama0  /* Especificidade */
b0<--1          /* Valor verdadeiro de beta0 */
b1<-1          /* Valor verdadeiro de beta1 */
b2<-1          /* Valor verdadeiro de beta2 */
sigu0<-1        /* Valor verdadeiro de sigma_u0*/
sigu1<-1        /* Valor verdadeiro de sigma_u1*/
R<-1000       /* Numero de replicas */
/*******************************************************
*                            Corpo                     *
*******************************************************/
/* Declaracao de variaveis */
d<-matrix(0,R,7)
/* Inicio do loop da simulacao */
for(r in 1:R){
/* Geracao da covariavel independentemente dentro de cada grupo */
x1<-matrix(0,nj,j)
x2<-matrix(0,nj,j)
u0<-matrix(0,nj,j)
u1<-matrix(0,nj,j)
T<-matrix(0,k,1)
y<-matrix(0,k,1)
for(i in 1:j){
    x1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    x2[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u0[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
    u1[,i]<-rnorm(nj) /* Gera nj valores da dist. normal padrao por grupo */
}
x11<-data.frame(x1) /* Converte a matriz x1 em dados */
x22<-data.frame(x2) /* Converte a matriz x2 em dados */
b11<-names(x11) /* Mostra os nomes das variaveis dos dados x1 */
b22<-names(x22)  /* Mostra os nomes das variaveis dos dados x2 */
u00<-data.frame(u0) /* Converte a matriz u0 em dados */
u11<-data.frame(u1)  /* Converte a matriz u1 em dados */
c00<-names(u0)       /* Mostra os nomes das variaveis dos dados u0 */
c11<-names(u1)    /* Mostra os nomes das variaveis dos dados u1 */
for(i in 1:nj){
    names(x11)[i]<-i /* Muda o nome da variavel para numeros */
    names(x22)[i]<-i /* Muda o nome da variavel para numeros */
    names(u00)[i]<-i /* Muda o nome da variavel para numeros */
    names(u11)[i]<-i /* Muda o nome da variavel para numeros */
}
vec11<-stack(x11) /* Empilha cada coluna de x11 em um vetor */
vec12<-stack(x22) /* Empilha cada coluna de x22 em um vetor */
vec21<-stack(u00) /* Empilha cada coluna de u00 em um vetor */
vec22<-stack(u11)  /* Empilha cada coluna de u11 em um vetor */
x1<-vec11[,1]      /* Armazena vec1 em x1 */
x2<-vec12[,1]      /* Armazena vec1 em x2 */
U0<-vec21[,1]      /* Armazena vec1 em U0 */
U1<-vec22[,1]       /* Armazena vec1 em U1 */
J<-vec11[,2]         /* Obtem a coluna dos grupos */
X<-cbind(1,x1,x2)  /* Obtem a  matriz de regressores */
/* Geracao de eta com base nos valores verdadeiros dos parametros: */
/* beta_0=-1; beta_1=1; beta_2=1 e sigma_u0=sigma_u1 */
eta<--1*X[,1]+1*X[,2]+1*X[,3]+U0+X[,2]*U1
/* Obter a probabilidade inicial em funcao de eta */
p<-pnorm(eta)
/* Gerar T verdadeiro a partir da dist. binomial(1,p) */
for(i in 1:k){
    T[i]<-rbinom(1,1,p[i])
}
```

```
/* Gerar y a partir de T */
for(i in 1: k){
    if (T[i]==1){
       y[i]<-rbinom(1,1,sens)
    }
    else{
       y[i]<-1-rbinom(1,1,espec)
    }
}
/* Estimar os erros ajustando-se um MMNLG para a resposta y via procedimento
Quadratura de Gauss-Hermite */
mu<-function (p) (1-p[6]-p[7])*(pnorm(p[1]+p[2]*x1+p[3]*x2+p[4]*U0+p[5]*U1*x1))+p[7]

print(z<-gnlr(y, dist="binomial", mu=mu, pmu=c(b0,b1,b2,sigu0,sigu1,tau1,tau0), pshape=1))
/* Ajuste para a resposta y via procedimento Quadratura de Gauss-Hermite */
ajustey<-gnlmm(y, dist="binomial", mu=mu, nest=J, pmu=z$coef[1:7],
               pshape=z$coef[8], psd=1, points=5)
b<-rbind(coef(ajustey)) /* Estimativas dos coeficientes e dos erros */
d[r,]<-b[1:7]           /* Armazena as estimativas em um vetor */
} /* Final do loop */

/* Verificar o numero total de convergencia */
conv<-0          /* Declaracao da variavel conv */
for (r in 1:R){
  if(d[r,1]!=0){
    conv<-conv+1 /* Conta o numero de convergencia */
  }
}

/* Impressao dos resultados */
nj     /* Imprime o tamanho da amostra no grupo */
tau0   /* Imprime o valor de tau0 */
tau1   /* Imprime o valor de tau1 */
conv    /* Imprime o total de convergencia */

b0<-summary(d[1:conv,1])/* Descritiva das estimativas beta0 de y */
b0                       /* Imprime a descritiva das est. de beta0 de y*/

b1<-summary(d[1:conv,2]) /* Descritiva das estimativas beta1 de y */
b1                       /* Imprime a descritiva das est. de beta1 de y*/

b2<-summary(d[1:conv,3]) /* Descritiva das estimativas beta2 de y */
b2                        /* Imprime a descritiva das est. de beta2 de y*/


u0<-(summary(d[1:conv,4]))/* Descritiva das estimativas u0 de y */
u1<-(summary(d[1:conv,5]))/* Descritiva das estimativas u1 de y */
t1<-(summary(d[1:conv,6]))/* Descritiva das estimativas tau1 de y */
t0<-(summary(d[1:conv,7]))/* Descritiva das estimativas tau0 de y */

u0    /* Imprime a descritiva das est. de u0 de y*/
u1    /* Imprime a descritiva das est. de u1 de y*/
t0    /* Imprime a descritiva das est. de tau0 de y*/
t1    /* Imprime a descritiva das est. de tau1 de y*/
```

D. Programa de aplicação no R

```
sandra2<-read.table("C://Userroot/Alunos/Sandrarj/tese/viviana/
sandra2.txt", header = TRUE,  na.strings = "NA" )

attach(sandra2)
UF<-sandra2$uf
damost<-sandra2$damost
sexo<-sandra2$sexo
IDCOD<-sandra2$idcod
cor<-sandra2$cor
estconj<-sandra2$estconj
classe<-sandra2$classe
escol<-sandra2$escol
religat<-sandra2$religat
EXPSEXO<-sandra2$camis
PERANALF<-sandra2$panalf
PERENSME<-sandra2$pensme
PERENSFU<-sandra2$pensup
PER18SUP<-sandra2$p18sup
PER25SUP<-sandra2$p25sup
PERAGU<-sandra2$pagua
PERENE<-sandra2$penerg
PERSUBN<-sandra2$psubn
IDHEDU<-sandra2$idhedu
IDHLONG<-sandra2$idhlong
IDHREND<-sandra2$idhrend
IDHMUNI<-sandra2$idhmuni
id2<-sandra2$id2
id3<-sandra2$id3
est2<-sandra2$est2
esco2<-sandra2$esco2
rel2<-sandra2$rel2
cla2<-sandra2$cla2

library(boot)
library(gnlm)
library(repeated)
U0<-rnorm(3315,0,1)

#***** Todas variaveis fixas *******

mu<-function (p) (1-(exp(p[1])/(1+exp(p[1])))-(exp(p[2])/(1+exp(p[2]))))*
             ((exp(p[3]+(p[4]*id2+p[5]*id3+p[6]*est2+p[7]*esco2+p[8]*cla2+p[9]*sexo+p[10]*rel2)
             +exp(p[11])*U0)/(1+exp(p[3]+(p[4]*id2+p[5]*id3+p[6]*est2+p[7]*esco2+p[8]*cla2+p[9]*sexo
             +p[10]*rel2)+exp(p[11])*U0))))+(exp(p[2])/(1+exp(p[2])))

print(z<-gnlr(EXPSEXO, dist="binomial", mu=mu, pmu=c(2,2,-1,3,3,8,1,1,-1,-1,-1), pshape=1))
t<-gnlmm(EXPSEXO, dist="binomial", mu=mu, nest=UF, pmu=z$coef[1:11], pshape=z$coef[12], psd=1, points=5)
#Wald
estati<-(t$coef[1:11]/t$se[1:11])^2
cbind(estati,1-pchisq(estati,1))

#****** Escolaridade aleatoria ************

mu<-function (p) (1-(exp(p[1])/(1+exp(p[1])))-(exp(p[2])/(1+exp(p[2]))))*
             ((exp(p[3]+(p[4]*id2+p[5]*id3+p[6]*est2+p[7]*cla2+p[8]*sexo+p[9]*rel2)
             +exp(p[10])*U0+exp(p[11])*U1*esco2)/(1+exp(p[3]+(p[4]*id2+p[5]*id3+p[6]*est2+
             p[7]*cla2+p[8]*sexo+p[9]*rel2)+exp(p[10])*U0+exp(p[11])*U1*esco2))))+
             (exp(p[2])/(1+exp(p[2])))

print(z<-gnlr(EXPSEXO, dist="binomial", mu=mu, pmu=c(2,2,-1,3,3,8,0,0,-1,-1,-10), pshape=1))
```

```
t<-gnlmm(EXPSEXO, dist="binomial", mu=mu, nest=UF, pmu=z$coef[1:11], pshape=z$coef[12],
psd=1, points=5)

estati<-(t$coef[1:11]/t$se[1:11])^2
cbind(estati,1-pchisq(estati,1))
```

E. Resultados da aplicação no R

```
#***** M1) Considerando todas as variveis fixas ********

> t
Call:
gnlmm(EXPSEXO, dist = "binomial", mu = mu, nest = UF, pmu = z$coef[1:11],
    pshape = z$coef[12], psd = 1, points = 5)

binomial distribution

 with normal mixing distribution on logit scale
 (5 point Gauss-Hermite integration)
% \headline{\hfill}
\headline{\hfil\tenpoint {\it  E. Resultados da aplica\cao no $\tt R$}}
Log likelihood function:
{
    m <- mu4(p)
    -wt * (y[, 1] * log(m) + y[, 2] * log(1 - m))
}

Location function:
(1 - (exp(p[1])/(1 + exp(p[1])))) - (exp(p[2])/(1 +
    exp(p[2])))) * ((exp(p[3] + (p[4] * id2 + p[5] * id3 + p[6] *
    est2 + p[7] * esco2 + p[8] * cla2 + p[9] * sexo + p[10] *
    rel2) + exp(p[11]) * U0)/(1 + exp(p[3] + (p[4] * id2 + p[5] *
    id3 + p[6] * est2 + p[7] * esco2 + p[8] * cla2 + p[9] * sexo +
    p[10] * rel2) + exp(p[11]) * U0)))) + (exp(p[2])/(1 + exp(p[2])))

-Log likelihood    1490.161
Degrees of freedom 3303
AIC                1502.161
Iterations         59

Location parameters:
       estimate      se
p[1]    -1.9629  0.1080
p[2]    -1.6279  0.1578
p[3]    -0.7252  0.8910
p[4]     2.3431  0.9661
p[5]     2.0808  0.9435
p[6]     7.1610  1.5722
p[7]     0.2411  0.3867
p[8]    -0.5118  0.4541
p[9]    -1.5115  0.5161
p[10]   -1.1224  0.4683
p[11]   -2.0044  1.3904

Mixing standard deviation:
   estimate      se
     0.1122  0.0804

> cbind(estati,1-pchisq(estati,1))
           estati
 [1,] 330.2139182 0.000000e+00
 [2,] 106.4744836 0.000000e+00
 [3,]   0.6623774 4.157219e-01
 [4,]   5.8820998 1.529557e-02
 [5,]   4.8643152 2.741768e-02
 [6,]  20.7466156 5.242415e-06
 [7,]   0.3887132 5.329766e-01
 [8,]   1.2702103 2.597272e-01
 [9,]   8.5782894 3.401948e-03
```

```
[10,]   5.7446763 1.653870e-02
[11,]   2.0783425 1.494023e-01

#***** M2) Considerando Percentual de analfabetismo e IDHRENDA
> t
Call:
gnlmm(EXPSEXO, dist = "binomial", mu = mu, nest = UF, pmu = z$coef[1:11],
    pshape = z$coef[12], psd = 1, points = 5)

binomial distribution

 with normal mixing distribution on logit scale
 (5 point Gauss-Hermite integration)

Log likelihood function:
{
    m <- mu4(p)
    -wt * (y[, 1] * log(m) + y[, 2] * log(1 - m))
}

Location function:
(1 - (exp(p[1])/(1 + exp(p[1]))) - (exp(p[2])/(1 +
    exp(p[2])))) * ((exp(p[3] + (p[4] * id2 + p[5] * id3 + p[6] *
    est2 + p[7] * PERANALF + p[8] * IDHREND + p[9] * sexo + p[10] *
    rel2) + exp(p[11]) * U0)/(1 + exp(p[3] + (p[4] * id2 + p[5] *
    id3 + p[6] * est2 + p[7] * PERANALF + p[8] * IDHREND + p[9] *
    sexo + p[10] * rel2) + exp(p[11]) * U0)))) + (exp(p[2])/(1 +
    exp(p[2])))

-Log likelihood    1490.705
Degrees of freedom 3303
AIC                1502.705
Iterations         63

Location parameters:
       estimate        se
p[1]   -1.95722   0.12259
p[2]   -1.60177   0.15236
p[3]   -6.30023   7.54060
p[4]    2.42622   1.18789
p[5]    2.24103   1.14031
p[6]    7.39040   1.90757
p[7]    0.05636   0.09285
p[8]    6.53898   8.74297
p[9]   -1.68704   0.59048
p[10]  -1.15832   0.49512
p[11]  -4.16142  13.52843

Mixing standard deviation:
   estimate       se
      0.123  0.08334

> estati<-(t$coef[1:11]/t$se[1:11])^2
> cbind(estati,1-pchisq(estati,1))
           estati
 [1,] 254.8939869 0.0000000000
 [2,] 110.5212006 0.0000000000
 [3,]   0.6980745 0.4034314341
 [4,]   4.1716318 0.0411061917
 [5,]   3.8623287 0.0493817438
 [6,]  15.0098497 0.0001069515
 [7,]   0.3684193 0.5438671276
 [8,]   0.5593729 0.4545130207
 [9,]   8.1627591 0.0042759273
```

```
[10,]   5.4732207 0.0193100087
[11,]   0.0946213 0.7583823335

#**** M3) Todas as variaveis fixas sem considerar nivel de instru. e classe soc. ***
> t
Call:
gnlmm(EXPSEXO, dist = "binomial", mu = mu, nest = UF, pmu = z$coef[1:9],
    pshape = z$coef[1:10], psd = 1, points = 5)

binomial distribution

 with normal mixing distribution on logit scale
 (5 point Gauss-Hermite integration)

Log likelihood function:
{
    m <- mu4(p)
    -wt * (y[, 1] * log(m) + y[, 2] * log(1 - m))
}

Location function:
(1 - (exp(p[1])/(1 + exp(p[1])))) - (exp(p[2])/(1 +
    exp(p[2])))) * ((exp(p[3] + (p[4] * id2 + p[5] * id3 + p[6] *
    est2 + p[7] * sexo + p[8] * rel2) + exp(p[9]) * U0)/(1 +
    exp(p[3] + (p[4] * id2 + p[5] * id3 + p[6] * est2 + p[7] *
        sexo + p[8] * rel2) + exp(p[9]) * U0)))) + (exp(p[2])/(1 +
    exp(p[2])))

-Log likelihood    1491.011
Degrees of freedom 3305
AIC                1501.011
Iterations         45

Location parameters:
      estimate       se
p[1]   -1.9601   0.1067
p[2]   -1.6200   0.1650
p[3]   -0.8718   0.8755
p[4]    2.2910   0.9816
p[5]    2.0890   0.9750
p[6]    7.1990   1.5808
p[7]   -1.6001   0.5469
p[8]   -1.1424   0.4736
p[9]  -10.4995  56.9435

Mixing standard deviation:
   estimate       se
    -0.1156  0.07974

> estati<-(t$coef[1:10]/t$se[1:10])^2
> cbind(estati,1-pchisq(estati,1))
           estati
 [1,] 337.41525318 0.000000e+00
 [2,]  96.36183217 0.000000e+00
 [3,]   0.99158455 3.193554e-01
 [4,]   5.44721331 1.959957e-02
 [5,]   4.59064364 3.214694e-02
 [6,]  20.74043521 5.259361e-06
 [7,]   8.56083560 3.434717e-03
 [8,]   5.81779953 1.586478e-02
 [9,]   0.03399734 8.537125e-01
[10,]   2.10021236 1.472787e-01

**** M4) Considerando escolaridade aleatoria
```

```
> t
Call:
gnlmm(EXPSEXO, dist = "binomial", mu = mu, nest = UF, pmu = z$coef[1:11],
    pshape = z$coef[12], psd = 1, points = 5)

binomial distribution

 with normal mixing distribution on logit scale
 (5 point Gauss-Hermite integration)

Log likelihood function:
{
    m <- mu4(p)
    -wt * (y[, 1] * log(m) + y[, 2] * log(1 - m))
}

Location function:
(1 - (exp(p[1])/(1 + exp(p[1])))) - (exp(p[2])/(1 +
    exp(p[2])))) * ((exp(p[3] + (p[4] * id2 + p[5] * id3 + p[6] *
    est2 + p[7] * cla2 + p[8] * sexo + p[9] * rel2) + exp(p[10]) *
    U0 + exp(p[11]) * U1 * esco2)/(1 + exp(p[3] + (p[4] * id2 +
    p[5] * id3 + p[6] * est2 + p[7] * cla2 + p[8] * sexo + p[9] *
    rel2) + exp(p[10]) * U0 + exp(p[11]) * U1 * esco2)))) + (exp(p[2])/(1 +
    exp(p[2])))

-Log likelihood    1489.071
Degrees of freedom 3303
AIC                1501.071
Iterations         62

Location parameters:
       estimate      se
p[1]    -1.9854  0.1168
p[2]    -1.6773  0.1894
p[3]    -0.6426  0.7600
p[4]     2.0975  0.8595
p[5]     1.8484  0.8398
p[6]     6.7283  1.4304
p[7]    -0.3577  0.3975
p[8]    -1.4154  0.4964
p[9]    -1.0014  0.4565
p[10]   -3.4191  5.9451
p[11]   -0.9045  0.5535

Mixing standard deviation:
   estimate       se
     0.1108  0.08134

> estati<-(t$coef[1:11]/t$se[1:11])^2
> cbind(estati,1-pchisq(estati,1))
           estati
 [1,] 288.8909529 0.000000e+00
 [2,]  78.4326571 0.000000e+00
 [3,]   0.7147765 3.978627e-01
 [4,]   5.9549829 1.467575e-02
 [5,]   4.8437737 2.774611e-02
 [6,]  22.1266051 2.552476e-06
 [7,]   0.8098266 3.681715e-01
 [8,]   8.1290835 4.356076e-03
 [9,]   4.8117917 2.826564e-02
[10,]   0.3307419 5.652225e-01
[11,]   2.6701865 1.022440e-01
```

Referências bibliográficas

[1] Aitkin, M., Anderson, D., Hinde, J. (1981). Statistical modelling of data on teaching styles. *Journal of the Royal Statistical Society*. Series A, v.144, p.148-161.

[2] Aitkin, M., Longford, N. (1986). Statistical modelling in school effectiveness studies *Journal of the Royal Statistical Society*. Series A, v.149, p.1-43.

[3] Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*. Second International Symposium on Information Theory. B. N. Petrov and F. Czáki (eds), p.267-281. Budapest: Akademiai Kiadó.

[4] Anderson, D. A., Aitkin, M. (1985). Variance component models with binary response. *Journal of the Royal Statistical Society*. Series B, v.47, p.203-210.

[5] Bennett, N. (1976). *Teaching styles and pupil progress*. Cambrige: Harvard University Press.

[6] Bergamo, G. C. (2002). *Aplicação de modelos multiníveis na análise de dados de medidas repetidas no tempo*. Dissertação de mestrado apresentada na Escola Superior de Agricultura Luiz de Queiroz, Universidade se São Paulo.

[7] Breslow, N. (2003). Whither PQL?. *http://www.bepress.com/uwbiostat/paper192.pdf*.

[8] Breslow, N. E., Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*. v.88, p.9-25.

[9] Breslow, N. E., Lin, X. (1995). Bias correction in generalized linear mixed models with a single components of dispersion. *Biometrika*. v.82, p.81-91.

[10] Bryk, A., Raudenbush, S. (2002). *Hierarchical Linear Models*. Newbury Park, California, Sage.

[11] Dempster, A. P., Laird, N. M., Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*. Series B. v.39, p.1-8.

[12] Gibbons, R. D., Hedeker, D. (1997). Random effects probit and logistic regression models for three level data. *Biometrics*. v.53, p.1527-1537.

[13] Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*. v.73, p.43-56.

[14] Goldstein, H. (1991). Nonlinear multilevel models, with an application to discrete response data. *Biometrika*. v.78, p.45-51.

[15] Goldstein, H. (1995). *Multilevel Statistical Models*. 2ª. ed. London: Institute of Education. University of London.

[16] Goldstein, H., Rasbash, J. (1996). Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society*. Serie A. v.150, p.505-513.

[17] Green, P. J. (1987). Penalized likelihood for general semi-parametric regression models. *International Statistical Review*. v.55, p.245-259.

[18] Guo, G., Zhao, H. (2000). Multilevel modeling for binary data. *Annu. Rev. Sociol.* v.26, p.441-462.

[19] Hedeker, D., Gibbons, R. D. (1996a). MIXOR: A computer program for mixed-effects ordinal regression analysis. *Biometrics*. v.50, p.933-944.

[20] Hedeker, D., Gibbons, R. D. (1996b). MIXREG: A computer program for mixed-effects regression analysis with autocorrelated errors. *Computer Methods and Programs in Biomedicine*. v.49, p.229-252.

[21] Hox, J. J. (1995). *Applied Multilevel Analysis*. Amsterdam: TT - Publikaties.

[22] Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics*. v.1. L. M. Le Cam and J. Neyman (eds). p.221-233. Berkeley: University of California Press.

[23] Ihaka, R., Gentleman, R. (1996). R a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*. v.5, p.299-314.

[24] Im, S., Gianola, D. (1988). Mixed models for binomial data with an application to lamb mortality. *Appl. Statist.* v.37, p.196-204.

[25] Jennrich R. I., Schluchter M. D. (1986). Unbalanced repeated-measure models with structured covariance matrices. *Biometrics*. v.42, p.805-820.

[26] Kuk, A. Y. C. (1995). Asymptotically unbiased estimation in generalised linear models with random effects. *Journal of the Royal Statistical Society*. Series B. v.57, p.395-407.

[27] Kullback, S. (1959). *Information Theory and Statistics*. New York: John Wiley.

[28] Laird, N. M. (1978). Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association*. v.73, p.805-811.

[29] Laird, N., Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*. v.38, p.963-974.

[30] Li, K. C., Duan, N. (1989). Regression analysis under link violation. *Annals of Statistics*. v.17, p.1009-1052.

[31] Lillard, L. A., Panis, C. W. A. (2003). *aML multilevel multiprocess statistical software version 2.0*. EconWare. Los Angeles. California.

[32] Lin, X., Breslow, N. E. (1996). Bias correction in generalized linear mixed models with multiple components of dispersion. *Journal of the American Statistical Association*. v.91, p.1007-1016.

[33] Lindley, D. V., Smith, A. F. M. (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society*. Series B. v.34, p.1-41.

[34] Longford, N. T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*. v.74, p.817-827.

[35] Longford, N. T. (1993a). *Random Coefficient Models*. Oxford: Clarendon Press.

[36] Longford, N. T. (1993b). *VARCL: Software for variance component analysis of data with nested random effects (maximun likehood)*. Manual, Groningen: Programma.

[37] Longford, N. T. (1994). Logistic regression with random coefficients. *J. Comput. Statist. Data Anal.* v.17, p.1-15.

[38] Mason, W. M. Wong, G, M., Entwistle B. (1983). Contextual analysis through the multilevel linear model. In S. Leinhardt (Ed.). *Sociological methodology*. v.13, p. 72-103. San Francisco: Jossey-Bass.

[39] McCullagh, P., Nelder, J. A. (1989). *Generalized Linear Models*. London: Chapman and Hall.

[40] Neuhaus, J. M. (1999). Bias and efficiency loss due to misclassified responses in binary regression. *Biometrika*. v.86, p.843-855.

[41] Neuhaus, J. M. (2002). Analysis of clustered and longitudinal binary data subject to response misclassification. *Biometrics*. v.58, p.675-683.

[42] Paula, G. A. (2003). Modelos de regressão com apoio cumputacional. São Paulo: IME/USP.

[43] Raudenbush, S. W., Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education*. v.59, p.1-17.

[44] Renard, D. (2002). Topics in modeling multilevel and longitudinal data. *http:// www.luc.ac.be/censtat/pub/dr.pdf* .

[45] Rodriguez, G., Goldman, N. (1995). An assessment of estimation procedures for multilevel models with binary responses. *Journal of the Royal Statistical Society*. Serie A. v.158, p.73-89.

[46] Schall, R. (1991). Estimation in generalized linear models with random effects. *Biometrika*. v.78, p.719-727.

[47] Smith, A. F. M. (1973). A general Bayesian linear model. *Journal of the Royal Statistical Society*. Series B. v.35, p.61-75.

[48] Snijders T., Bosker R. (1999). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. Sage: Publications.

[49] Stiratelli, N. L., Ware, J. H. (1984). Random-effects models for serial observations with binary response. *Biometrics*. v.40, p.961-971.

[50] Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, v.61, p.439-447.

[51] White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*. v.50, p.1-25.

[52] Wolfinger, R. (1993). Laplace's approximation for nonlinear mixed models. *Biometrika*. v.80, p.791-795.

[53] Yau, K. K. W. (2001). Multilevel models for survival analysis with random effects. *Biometrics*. v.57, p.96-102.