

Modelos de Mistura para a Avaliação de Produtos

Démerson André Polli

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação
em Estatística (PIPGES)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Démerson André Polli

Modelos de Mistura para a Avaliação de Produtos

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Doutor em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística. *VERSÃO REVISADA*

Área de Concentração: Estatística

Orientador: Prof. Dr. Carlos Alberto de Ribeiro Diniz

USP – São Carlos
Janeiro de 2020

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

P774m Polli, Démerson André
Modelos de Mistura para Avaliação de Produtos /
Démerson André Polli; orientador Carlos Alberto
Ribeiro Diniz. -- São Carlos, 2020.
133 p.

Tese (Doutorado - Programa Interinstitucional de
Pós-graduação em Estatística) -- Instituto de Ciências
Matemáticas e de Computação, Universidade de São
Paulo, 2019.

1. Avaliação de produtos. 2. Modelos de mistura.
3. Distribuições discretas. I. Diniz, Carlos Alberto
Ribeiro, orient. II. Título.

Démerson André Polli

Mixture Models for Product Evaluation

Doctoral dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC- USP and to the Department of Statistics – DEs- UFSCar, in partial fulfillment of the requirements for the degree of the Doctorate Interagency Program Graduate in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Carlos Alberto de Ribeiro Diniz

USP – São Carlos
January 2020

Agradecimentos

Antes de agradecer a qualquer pessoa que tenha contribuído com este trabalho, é importante observar que nada disto seria possível se não fosse da vontade de Deus, à quem se deve toda a gratidão e honra.

Dito isto, agradeço ao meu orientador, professor Carlos Alberto Ribeiro Diniz, pelas preciosas dicas e contribuições no decorrer do trabalho. Muitos dos acertos que aqui estão se devem às dicas dele e, obviamente, muitas das possíveis falhas são de inteira responsabilidade do autor que vos fala.

Agradeço muito à minha esposa, Jadasa Emerick, por ter sido uma companhia sempre presente ao longo da elaboração deste trabalho. A sua presença tornou tudo muito mais fácil, pode ter certeza.

Agradeço aos professores Donald Pianto, Bernardo Borba, Eduardo Nakano e André Cançado e, especialmente, George von Borries pelo incentivo que me deram para encarar o desafio de fazer um doutorado. Confesso que muitas vezes já cogitei em não assumir tal desafio, mas o incentivo de vocês (em especial do professor George) foi primordial.

Por fim, agradeço meus amigos do Ministério da Economia, em especial Alex Lima, Sérgio Gadelha, Marco Cavalcanti e Adolfo Sachsida. O ambiente de amizade e profissionalismo que compartilho com vocês foi um divisor de águas.

Muito obrigado.

Resumo

POLLI, D. A. Modelos de mistura para a avaliação de produtos. 2020. 133 p. Tese (Doutorado em Ciências – Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

A avaliação de produtos (ou serviços) é uma necessidade em muitas etapas do desenvolvimento e lançamento destes produtos (serviços). A evolução do produto (serviço) e a definição da fatia de mercado são duas etapas que envolvem algum tipo de avaliação. No entanto, na avaliação pode ocorrer a atribuição de notas não relacionadas com a percepção de qualidade do produto (ou serviço) – pela atribuição de notas ao acaso. Este trabalho apresenta 4 novos modelos, 2 deles baseados no modelo CUB (*Combined Uniform and shifted Binomial* – Uniforme (discreta) combinada com binomial deslocada), para a avaliação de itens quando existem dois ou mais níveis de preferência entre os avaliadores e pode haver avaliadores atribuindo notas ao acaso.

Palavras-chave: avaliação de produtos. modelos de mistura finita discretas. distribuições discretas.

Abstract

POLLI, D. A. Modelos de mistura para a avaliação de produtos. 2020. 133 p. Tese (Doutorado em Ciências – Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

The product (or services) evaluation is a necessity in many steps of the development and release of such products (services). The product evolution and the market share definition are two steps that evolves some kind of evaluation. Although, the evaluation occur by assigning grades not related with the perception of the product (service) quality – assigning values at random. This work presents 4 new models, 2 of them based on the CUB (*Combined Uniform and shifted Binomial*) model, for evaluation of items where there are two or more levels of preference by the evaluators and may have evaluators assigning grades at random.

Keywords: product evaluation. finite mixture models. discrete distributions.

Lista de ilustrações

Figura 1 – Estimativas dos parâmetros de mistura para 500 amostras com escala de 20 notas, amostra de tamanho 500 e preditor linear do segundo componente binomial variando entre 0.50 e 0.75 – efeito do preditor linear no modelo binomial redimensionado e deslocado.	72
Figura 2 – Estimativas dos parâmetros de mistura para 500 amostras com escala de 20 notas, tamanho de amostra variando entre 100 e 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho amostral no modelo binomial redimensionado e deslocado.	73
Figura 3 – Estimativas dos parâmetros de mistura para 500 amostras com escala de notas variando entre 10 e 20, tamanho de amostra igual a 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho da escala de notas no modelo binomial redimensionado e deslocado. . .	74
Figura 4 – Estimativas dos parâmetros de regressão para 500 amostras com escala de 20 notas, amostra de tamanho 500 e preditor linear do segundo componente binomial variando entre 0.50 e 0.75 – efeito do preditor linear no modelo binomial redimensionado e deslocado.	75
Figura 5 – Estimativas dos parâmetros de regressão para 500 amostras com escala de 20 notas, tamanho de amostra variando entre 100 e 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho amostral no modelo binomial redimensionado e deslocado.	76
Figura 6 – Estimativas dos parâmetros de mistura para 500 amostras com escala de notas variando entre 10 e 20, tamanho de amostra igual a 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho da escala de notas no modelo binomial redimensionado e deslocado. . .	77
Figura 7 – Estimativas dos parâmetros de mistura para 500 amostras com escala de 10 notas – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.	79
Figura 8 – Estimativas dos parâmetros de regressão do primeiro componente da mistura para 500 amostras com escala de 10 notas – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.	80
Figura 9 – Estimativas dos parâmetros de regressão do segundo componente da mistura para 500 amostras com escala de 10 notas – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.	81
Figura 10 – Estimativas dos parâmetros de mistura para 500 amostras de tamanho 100 – efeito do tamanho da escala de notas para o modelo binomial redimensionado e deslocado.	82

Figura 11 – Estimativas dos parâmetros de regressão do primeiro componente da mistura para 500 amostras de tamanho 100 – efeito do tamanho da escala de notas para o modelo binomial redimensionado e deslocado.	83
Figura 12 – Estimativas dos parâmetros de regressão do segundo componente da mistura para 500 amostras de tamanho 100 notas – efeito do tamanho escala de notas para o modelo binomial redimensionado e deslocado.	84
Figura 13 – Estimativas dos parâmetros de mistura, referente ao termo de avaliação ao acaso e inflacionada, para 500 amostras com 10% de inflação na maior nota da escala – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.	85
Figura 14 – Estimativas dos parâmetros de mistura, referente ao termo de avaliação ao acaso e inflacionada, para 500 amostras com 25% de inflação na maior nota da escala – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.	86
Figura 15 – Estimativas dos parâmetros de mistura, referente ao termo de avaliação ao acaso e inflacionada, para 500 amostras com 50% de inflação na maior nota da escala – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.	87
Figura 16 – Estimativas dos parâmetros de mistura, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.	88
Figura 17 – Estimativas dos parâmetros de variância, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.	89
Figura 18 – Estimativas dos parâmetros de regressão, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.	90
Figura 19 – Estimativas dos parâmetros de mistura, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.	91
Figura 20 – Estimativas dos parâmetros de variância, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.	92

Figura 21 – Estimativas dos parâmetros de regressão, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.	93
Figura 22 – Histograma dos escores de ansiedade manifestada de Taylor.	95
Figura 23 – Medidas de alavanca para os dados de ansiedade manifestada de Taylor – modelo binomial redimensionado e deslocado.	101
Figura 24 – Medidas de alavanca para os dados de ansiedade manifestada de Taylor – modelo binomial duplo.	102
Figura 25 – Medidas de alavanca para os dados de ansiedade manifestada de Taylor – modelo binomial multiplicativo.	103
Figura 26 – Medidas de influência para os dados de ansiedade manifestada de Taylor – modelo binomial redimensionado e deslocado.	104
Figura 27 – Medidas de influência para os dados de ansiedade manifestada de Taylor – modelo binomial duplo.	105
Figura 28 – Medidas de influência para os dados de ansiedade manifestada de Taylor – modelo binomial multiplicativo.	106

Lista de tabelas

Tabela 1 – Preditor linear e probabilidade das distribuições binomiais redimensionadas e deslocadas – simulação do modelo sem inflação.	78
Tabela 2 – Descrição da amostra de ansiedade manifestada de Taylor. Conjunto de dados atualizado até 22 de julho de 2012.	94
Tabela 3 – Ajuste do modelo binomial redimensionado e deslocado sem inflação com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.	96
Tabela 4 – Ajuste do modelo binomial redimensionado e deslocado sem inflação com 1 componente binomial aos dados de ansiedade manifestada de Taylor.	96
Tabela 5 – Ajuste do modelo binomial redimensionado e deslocado com inflação com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.	97
Tabela 6 – Ajuste do modelo binomial duplo com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.	98
Tabela 7 – Ajuste do modelo binomial duplo com 1 componente binomial aos dados de ansiedade manifestada de Taylor.	99
Tabela 8 – Ajuste do modelo binomial multiplicativo com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.	100
Tabela 9 – Ajuste do modelo binomial multiplicativo com 1 componente binomial aos dados de ansiedade manifestada de Taylor.	100

Sumário

1	INTRODUÇÃO	17
2	EXTENSÃO DO MODELO CUB	23
2.1	Os modelos lineares	23
2.2	Modelos para diferentes tipos de avaliadores	26
2.3	Inflação em um valor da escala	30
3	MODELOS PARA SOBREDISPERSÃO	35
3.1	O modelo binomial duplo	35
3.1.1	Incorporando covariáveis ao modelo	37
3.1.2	Modelo para avaliação de produtos	37
3.2	O modelo binomial multiplicativo	39
3.2.1	Incorporando covariáveis ao modelo	40
3.2.2	Modelo para avaliação de produtos	40
4	INFERÊNCIA PARA OS MODELOS	45
4.1	Qualidade de ajuste para os modelos	45
4.1.1	Ponto de alavanca	45
4.1.2	Influência	47
4.1.3	Influência local	49
4.2	Testes de hipóteses	51
4.2.1	Teste de razão de verossimilhanças	52
4.2.2	Teste de Wald	57
4.2.3	Teste escore	60
4.2.4	Teste gradiente	64
5	APLICAÇÃO DOS MODELOS	71
5.1	Estudos de simulação	71
5.2	Aplicação a dados reais	93
6	CONCLUSÃO	107
	REFERÊNCIAS	109

ANEXOS

113

ANEXO A – CÓDIGOS PARA ESTIMAÇÃO DOS MODELOS . . 115

1 Introdução

O sucesso de um produto ou de um serviço é resultado de uma série de fatores, dentre os quais a percepção da qualidade pelos clientes ou usuários é um dos mais importantes. Um produto ou serviço que não atenda às necessidades dos clientes em geral não permanece no mercado por muito tempo. Desta forma, a avaliação destes produtos e serviços é uma necessidade para se traçar estratégias de melhoria continuada, sobretudo quando se planeja um novo serviço, se desenvolve novos produtos ou, até mesmo, com o objetivo de direcionar campanhas de marketing. A avaliação da percepção da qualidade por parte dos clientes permite às empresas identificar pontos de melhoria.

É possível definir a qualidade como uma relação entre a expectativa que determinado produto ou serviço desperta nos clientes e usuários e a percepção de quanto tal expectativa é satisfeita. Um produto ou um serviço que se parece igual ou melhor que a expectativa é, em geral, considerado um produto de qualidade. Por outro lado, um produto que não atende todas as expectativas pode ser considerado de menor qualidade, mesmo que comparado a outros produtos cuja expectativa é menor. Desta forma, um produto de luxo pode fornecer uma percepção de qualidade inferior à de um produto mais popular. É importante observar que estamos mensurando algo subjetivo. Um produto pode ser considerado de qualidade por alguns clientes enquanto recebe uma avaliação pior de outros. Além disto, diversos fatores podem influenciar na expectativa que se tem do produto ou do serviço e, por consequência, na percepção da qualidade.

O preço de um produto ou de um serviço é um dos fatores que podem influenciar na percepção de qualidade (Dodds; Monroe; Grewal, 1979; Zeithaml, 1988; Homburg; Hoyer; Koschate, 2005) ou na percepção do valor do produto ou do serviço (Sweeney; Soutar, 2001), em que preço e valor são conceitos distintos – preço está associado com a grandeza monetária enquanto que valor se relaciona com a expectativa que se tem do produto ou do serviço – um produto ou serviço custa caro quando o preço é superior ao valor percebido; o preço é uma medida objetiva enquanto que valor é uma medida subjetiva e, assim, o preço de um produto pode ser considerado excessivo por alguns clientes enquanto que é considerado justo por outros. Observe que valor e qualidade possuem uma relação estrita.

Alguns aspectos psicológicos também podem influenciar a percepção de qualidade de um produto ou serviço. Um aspecto interessante é o conceito de dissonância, o efeito da percepção da qualidade de produtos ou serviços concorrentes na mesma percepção do produto ou serviço adquirido (Cohen; Goldberd, 1970). Além disto, algumas áreas podem ter particularidades que influenciam nas avaliações. Larsen et al. (1979) estuda alguns vícios que ocorrem na avaliação de serviços de saúde, como, por exemplo, a tendência

dos pacientes sempre atribuir notas altas, a dificuldade de mensuração e interpretação da satisfação do cliente que se confunde com outros fatores e a grande variabilidade na satisfação em função aos diferentes estágios dos tratamentos de saúde. [Taylor e Baker \(1994\)](#) avaliam o efeito da satisfação do cliente na intenção de novas compras de serviços, em particular, de saúde, recreação (parques de diversão), viagens aéreas e telefonia de longa distância. [Hui e Tse \(1996\)](#) estuda o efeito do tempo de espera para atendimento na avaliação da qualidade de serviços.

A percepção de qualidade, bem como a própria ideia de qualidade, sofre influências psicológicas e subjetivas e, desta forma, é coerente pensar que existe uma grande variação na avaliação de qualidade entre diferentes avaliadores. Da mesma forma, é coerente que os avaliadores sejam provenientes de grupos que possuem expectativas distintas, inclusive ponderando características do produto de forma diferente. Por exemplo, ao se avaliar a qualidade de um automóvel, um grupo de pessoas pode ter grandes expectativas quanto à potência do motor enquanto ignora o conforto interno, enquanto outro grupo de pessoas pode ter grandes expectativas quanto a fatores estéticos enquanto consideram de pouca importância o espaço interno. Esta diferença na definição de expectativas entre faz com que os avaliadores tenham que ser vistos como pertencentes a diferentes populações.

Muitos trabalhos encontrados na literatura a respeito da avaliação da qualidade de serviços, inclusive os citados acima, aplicam métodos de análise descritiva, análise fatorial, componentes principais e regressões lineares para se fazer a análise dos dados observados. Neste cenário, os avaliadores são considerados como provenientes de uma única população. Esta homogeneidade pode não ser razoável.

A avaliação de produtos e serviços pode ser feita de diferentes maneiras. Uma delas, usual na literatura, consiste de ordenar os itens em avaliação com relação ao nível de preferência ou de qualidade percebida. Tal ordenação pode ser realizada comparando os objetos aos pares ou testando um rol de itens e, posteriormente, colocando-os em ordem de preferência ou de qualidade e atribuindo uma posição (posto) aos itens. Diversos modelos que permitem a análise de dados em postos são descritos por [Fligner e Verducci \(1993\)](#) e [Marden \(1995\)](#). No entanto, os modelos consideram homogeneidade entre avaliadores.

[Bailey e Terry \(1952\)](#) apresenta um modelo para a ordenação de um rol de itens (produtos) em que as comparações são realizadas aos pares: os itens sob avaliação são apresentados aos pares para os avaliadores que selecionam o objeto preferido. No modelo original não existe a necessidade de que todos os pares de itens sejam avaliados por todos os avaliadores, embora seja necessária a repetição de cada par avaliado com avaliações por pessoas diferentes. Também não é permitido haver empates entre os itens avaliados, ou seja, os avaliadores sempre deve apontar qual dos itens em cada um dos pares é o preferido. Este procedimento consiste em uma ordenação por preferência aos pares. Os resultados assintóticos para os estimadores das probabilidades de cada um dos itens ser indicado

como o preferido no modelo de Bradley-Terry são estudados por [Bradley \(1955\)](#). [Bradley e El-Helbawy \(1976\)](#) apresentam um método para a estimação de contrastes referentes às proporções que cada item é selecionado como preferido.

O modelo de Bradley-Terry tem como uma de suas suposições que a quantidade de comparações para cada par de itens sejam iguais. Tal restrição, como aponta [Ford Jr. \(1957\)](#), pode ser difícil de se cumprir para amostras grandes em alguns contextos (um exemplo ocorre quanto se compara produtos de marcas diferentes e algumas destas marcas são menos comuns que outras, caso em que pode haver poucas pessoas que conhecem o produto e, assim, poucas pessoas que opinam sobre os mesmos). Neste contexto, [Ford Jr.](#) propõe uma extensão para o modelo acomodar frequências diferentes para os pares de itens quando a amostra é grande o suficiente. Outra restrição que ocorre no modelo original de Bradley-Terry é a impossibilidade de empates entre pares de itens. Ao serem apresentados aos avaliadores, cada par tem que ter obrigatoriamente um dos itens apontados como preferido. [Davidson e Beaver \(1977\)](#) apresentam uma extensão do modelo para permitir que, ao avaliar um par de itens, o avaliador possa apontar igualdade de preferência entre os itens.

[Springall \(1973\)](#) apresenta uma extensão do modelo de Bradley-Terry em que as comparações entre os pares são associadas à variáveis quantitativas chamadas de resposta percebida. Neste caso, um item é considerado preferido se a resposta percebida for maior que a do item concorrente. Neste cenário, covariáveis contínuas são usadas para modelar a média das distribuições de tais respostas.

[Grizzle, Starmer e Koch \(1969\)](#) apresenta um método para a obtenção de funções lineares ou logarítmicas para descrever a proporção de respostas para dados categorizados. Este modelo é aplicado por [Beaver \(1977\)](#) para implementar o modelo de Bradley-Terry através de funções lineares considerando comparações aos pares ou em trios.

[Bockenholt \(1993\)](#) propõe o uso de um modelo de mistura finita da mesma distribuição de probabilidade para modelar dados de postos em processos de avaliação de itens por ordenação. Tal proposta é feita com o objetivo de acomodar a heterogeneidade observada entre avaliadores. [Stern \(1993\)](#) usa tais modelos de mistura finita para analisar o resultado de eleições com base em dados sócio-econômicos. [Murphy e Martin \(2003\)](#) propõem ajustar os modelos de misturas finitas para uma medida de distância entre as respostas dos diferentes avaliadores.

As propostas anteriores construídas a partir de misturas de distribuições consideram apenas componentes da mesma distribuição, diferindo apenas nos parâmetros. Uma nova classe de modelos que surge como variação de tal proposta consiste em definir misturas finitas de distribuições discretas distintas. O modelo CUB (*Combined Uniform and shifted Binomial distribution*), apresentado em 2005, é um exemplo de modelo de tal classe.

Considere que um rol de itens é apresentado para um grupo de avaliadores. Os avaliadores ordenam os itens por ordem de preferência, atribuindo a posição 1 para o item preferido, a posição 2 para o segundo item preferido, sucessivamente, até o último item (o menos preferido). Neste cenário, considerando as posições (postos) atribuídas à um dos objetos sob avaliação, [D’Elia \(2000\)](#) propõe modelar a distribuição dos postos como uma amostra aleatória da distribuição binomial deslocada (em 1 unidade). No entanto, esta proposta considera apenas avaliadores homogêneos. Suponha que cada avaliador ordene os itens considerando um nível de preferência (ou percepção de qualidade) predefinida ou, caso contrário, ao acaso sorteando os itens sem reposição de uma urna. Fixando um dos itens sob avaliação, se a ordenação ocorrer de acordo com um nível de preferência predefinido, os postos atribuídos ao item é uma amostra da distribuição binomial deslocada. Por outro lado, se a ordenação ocorre ao acaso é razoável que a distribuição dos postos seja uma uniforme discreta – situação na qual cada possível posto atribuído por um avaliador é equiprovável. O modelo de mistura destas duas distribuições, binomial deslocada com uniforme discreta, é conhecido como modelo CUB e foi proposto por [D’Elia e Piccolo \(2005\)](#).

O modelo CUB é um modelo que foi bastante estudado e desenvolvido pela mesma equipe que o propôs. [Piccolo \(2006\)](#) apresenta o modelo CUB com a incorporação de covariáveis e apresenta resultados referentes à matriz de informação observada. A identificabilidade dos parâmetros do modelo é estudada por [Iannario \(2010\)](#). [Iannario \(2012a\)](#) incorpora uma estrutura hierárquica ao modelo CUB para incorporar efeitos aleatórios ao modelo. [Grilli et al. \(2014\)](#) propõem a incorporação de variáveis latentes ao modelo CUB para acomodar heterogeneidade relacionadas à grupos de avaliadores. [Iannario \(2014\)](#) propõe uma nova classe de modelos chamada CUBE, que consiste em um modelo CUB com o componente binomial deslocado substituído por uma distribuição beta-binomial deslocada, com o objetivo de acomodar sobre-dispersão da variável resposta. [Gottard, Iannario e Piccolo \(2016\)](#) apresentam uma variação do modelo CUB em que o componente uniforme discreto é substituído por outra distribuição discreta definida no mesmo suporte, com o objetivo de acomodar diferentes níveis de incerteza quanto à resposta do avaliador ao acaso.

Supondo que o processo de avaliação de um determinado item ocorre pela seleção de uma resposta em uma escala ou lista de opções cujas opções são ordinais. A escala de Likert é um exemplo de tal escala. Neste caso, o modelo CUB pode ser utilizado supondo que as avaliações por preferência ocorrem como se fosse uma amostra aleatória da distribuição binomial deslocada (cada valor possível na escala tem uma relação um-para-um com os valores correspondentes de uma variável aleatória com distribuição binomial deslocada). Neste cenário, [Iannario \(2012b\)](#) propõe um modelo para a ocorrência de inflação em uma das possíveis respostas da escala. Ainda no mesmo cenário, [Manisera e Zuccolotto \(2014\)](#) apresentam uma proposta de modelo CUB não linear para acomodar probabilidades não

constantes para a ‘transição’ entre valores distintos da escala (por exemplo, em uma escala com opções ótimo, bom, regular ou ruim, a probabilidade de mudar a opinião de ‘ruim’ para ‘regular’ no modelo CUB usual é a mesma probabilidade de mudar a opinião de ‘bom’ para ‘ótimo’; o modelo CUB não linear permite que as probabilidades sejam diferentes).

O modelo CUB ajusta a distribuição dos postos atribuídos à um dos itens sob avaliação. A modelagem da distribuição dos postos para todos os itens requer o ajuste de um modelo CUB para cada item. No entanto, os parâmetros destes modelos são inter-relacionados. [Iannario e Piccolo \(2014\)](#) apresenta um teorema que relaciona os parâmetros das distribuições dos postos para cada item sob avaliação.

Contribuições deste trabalho

A mistura finita de distribuições é uma opção simples e prática para a modelagem de dados cujo mecanismo gerador pode ser diferentes distribuições e que, após observado os dados, não é possível apontar com certeza qual das distribuições foi utilizada. No cenário de avaliação de produtos (ou de serviços) através da ordenação por ordem de preferência ou por categorização com base em uma escala, considerando a heterogeneidade dos avaliadores, o modelo se mostra uma excelente opção. É neste caso que o modelo CUB e todas as variantes se encaixam. No entanto, esta classe de modelos pode ser estendida se for considerado que a ordem atribuída ao item pelo avaliador não precisa obrigatoriamente ser um valor da sequência iniciada em 1. Um caso particular que pode ser de interesse prático é a sequência iniciada em 0. Considere que o avaliador observa um item (produto ou serviço) e julga se um conjunto fixado de características (ou quesitos) são consideradas satisfatórias. Se os avaliadores são homogêneos e supondo equiprobabilidade para os quesitos serem julgados como satisfatórios e independência entre os quesitos, a atribuição do resultado satisfatório/insatisfatório aos quesitos são ensaios de Bernoulli e uma nota formada pela contagem dos quesitos satisfatórios segue uma distribuição binomial. Imaginando que os avaliadores podem observar os quesitos ou simplesmente chutar uma nota, obtém-se um modelo muito parecido com o CUB porém com a escala iniciada em 0 ao invés de 1. A generalização para escalas iniciadas em qualquer valor inteiro é relativamente simples e está apresentada no capítulo 2. A seção 2.1 apresenta a generalização da classe de modelos CUB linear para escalas iniciadas em qualquer valor inteiro (positivo ou negativo), escrevendo o componente referente às avaliações por preferência definida como um modelo linear generalizado, com as extensões para acomodar covariáveis e diferentes níveis de incerteza para o avaliador que atribui valores ao acaso. Os modelos supõem que os avaliadores são provenientes de 2 grupos distintos: (1) avaliadores que atribuem as respostas, seja um posto ou uma nota, considerando a qualidade do item sob avaliação ou a preferência correspondente; e (2) avaliadores que atribuem as respostas ao acaso. A possibilidade de modelar diferentes níveis de preferência quando os avaliadores fazem as avaliações

de acordo com alguma preferência predefinida faz com que seja necessária a adoção de misturas de distribuições com maior número de elementos de distribuições do tipo binomial (binomial usual ou binomial deslocada). Tal extensão é apresentada na referida seção.

No capítulo 3 são apresentados modelos para o cenário em que a avaliação ocorre por atribuição de uma nota discreta (ou a contagem de quesitos satisfatórios) porém no cenário em que a variância da resposta não é uma função direta da média – ocorrendo sobredispersão – desta forma, o uso de distribuições baseadas em reparametrização da binomial pode ser inadequada. Neste capítulo apresenta-se os modelos de avaliação de produtos para os modelos binomial duplo e binomial multiplicativo. Aplicações dos modelos para dados simulados estão apresentadas no capítulo 5.1. As considerações finais são apresentadas no capítulo 6.

2 Extensão do modelo CUB

A avaliação de produtos (ou serviços) pode ser realizada com a atribuição de um valor em uma sequência de números inteiros. Tal valor atribuído ao item sob avaliação pode ser uma nota (positiva ou negativa, com saltos de tamanho fixado) ou um posto (em uma escala iniciada em 1 com salto de tamanho 1, indicando a posição (posto) do objeto em uma ordenação).

D’Elia (2000) propõe, para o cenário de avaliação de $m \geq 1$ produtos por ordenação, modelar os postos atribuídos à um dos itens com uma distribuição *binomial deslocada* em 1 unidade. Tal cenário considera que todos os avaliadores ordenam os objetos de acordo com um nível único de preferência. O modelo CUB (*Combined Uniform and shifted Binomial*), proposto por D’Elia e Piccolo (2005), estende tal proposta para a situação em que existem 2 tipos distintos de avaliadores: (1) aqueles que avaliam os itens considerando um nível de preferência definido ou (2) aqueles que atribuem os postos ao acaso (selecionando os itens sem reposição de uma urna). Algumas extensões para o modelo CUB são propostas por Piccolo (2006) com a incorporação de covariáveis ao modelo, Iannario (2012b) apresenta o modelo com inflação em uma nota específica (ou posto específico) e Gottard, Iannario e Piccolo (2016) substitui o componente que modela a avaliação ao acaso (o componente uniforme discreto) para acomodar diferentes níveis de incerteza.

Neste capítulo, a distribuição *binomial deslocada* é estendida para acomodar um parâmetro (fixado) de escala $s > 0$ e um parâmetro (também fixado) de deslocamento $0 \leq d < m$ de tal modo que as notas atribuídas aos produtos sejam valores no conjunto $\{z \in \mathbb{N} : z = s \cdot (y + d), y = 0, 1, \dots, m - d, s \in \mathbb{R} > 0, 0 \leq d \in \mathbb{N} < m\}$. Esta extensão é nomeada distribuição *binomial redimensionada e deslocada* (*stretched-and-shifted binomial*) e está apresentada na seção 2.1 em conjunto com o modelo linear correspondente. De uso desta distribuição, os modelos citados são estendidos para acomodar notas entre $s \times d$ e $s \times m$ com saltos de tamanho $s > 0$ e também a possibilidade de diferentes níveis de preferência entre os avaliadores (seções 2.2 e 2.3). É conveniente observar que os modelos originais são casos particulares dos modelos propostos nas seções 2.1, 2.2 e 2.3, tomando $s = d = 1$.

2.1 Os modelos lineares

Suponha que um produto (ou serviço) é avaliado com a atribuição de um valor da sequência $\{z \in \mathbb{N} : z = s \cdot (y + d), y = 0, 1, \dots, m - d, s \in \mathbb{R} > 0, 0 \leq d \in \mathbb{N} < m\}$. Como exemplo, imagine que um produto é apresentado para um grupo de avaliadores que devem precificá-lo em múltiplos de \$5,00, com preço mínimo igual a \$10,00 e preço

máximo igual a \$40,00. Neste cenário, o modelo é configurado com $m = 6$, $s = 5$ e $d = 2$. Observe que m é o número de valores distintos na escala.

Se Y é uma variável aleatória com distribuição binomial de parâmetros $m - d$ e $0 \leq \pi \leq 1$, e $Z = s \cdot (Y + d)$, $s \in \mathbb{R} > 0$, $0 \leq d \in \mathbb{N} < m$, $m \geq 1$, então Z segue uma distribuição cuja função de probabilidade, parametrizada em função de $\mu = (m - d)\pi$, é

$$\mathbb{P}(Z|\mu) = \binom{m-d}{Z/s-d} \left(\frac{\mu}{m-d}\right)^{Z/s-d} \left(1 - \frac{\mu}{m-d}\right)^{m-Z/s}$$

A esperança e a variância de tal distribuição são dadas, respectivamente, por

$$\mathbb{E}(Z|\mu) = s(\mu + d), \quad \mathbb{V}(Z|\mu) = s^2\mu \left(1 - \frac{\mu}{m-d}\right).$$

Tal distribuição é nomeada *binomial redimensionada e deslocada* (*stretched-and-shifted binomial*) e pode ser escrita na forma

$$\mathbb{P}(Z|\mu) = \binom{m-d}{Z/s-d} \exp \left\{ (Z/s) \ln \left(\frac{\mu}{m-d-\mu} \right) - d \ln \left(\frac{\mu}{m-d} \right) + m \ln \left(1 - \frac{\mu}{m-d} \right) \right\}.$$

Desta forma, os modelos definidos com base nesta distribuição podem ser escritos com a notação de modelo linear generalizado (Nelder; Wedderburn, 1972). Tomando a relação (o preditor linear)

$$\eta = \ln \left(\frac{\mu}{m-d-\mu} \right) \rightarrow \mu = \frac{(m-d)e^\eta}{1+e^\eta}$$

é possível escrever tal distribuição em função do preditor η como

$$\mathbb{P}(Z|\eta) = \binom{m-d}{Z/s-d} \exp \left\{ \left(\frac{Z}{s} - d \right) \eta - (m-d) \ln(e^\eta + 1) \right\}.$$

A esperança e a variância da variável aleatória Z são, respectivamente,

$$\mathbb{E}(Z|\eta) = s \frac{(m-d)e^\eta}{1+e^\eta} + sd, \quad \mathbb{V}(Z|\eta) = s^2(m-d) \frac{e^\eta}{(1+e^\eta)^2}.$$

Observada uma amostra de tamanho $n \geq 1$ de avaliações de um determinado produto, com o vetor \mathbf{Z} (de dimensão n) de notas e a matriz \mathbf{X} (de dimensão $n \times p$) de covariáveis associadas aos avaliadores, é possível escrever a função de verossimilhança do modelo como

$$L(\boldsymbol{\eta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \binom{m-d}{z_i/s-d} \exp \left\{ \left(\frac{z_i}{s} - d \right) \eta_i - (m-d) \ln(e^{\eta_i} + 1) \right\},$$

em que $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_n\}$ é vetor dos preditores lineares associados à cada avaliador. A incorporação de covariáveis ao modelo ocorre de forma trivial, como nos modelos lineares generalizados, tomando $\eta_i = \mathbf{X}'_i \boldsymbol{\beta}$, $i = 1, 2, \dots, n$, em que \mathbf{X}'_i é um vetor de dimensão p de covariáveis associadas a cada avaliador e $\boldsymbol{\beta}$ um vetor de dimensão p de coeficientes. Os modelos com intercepto são facilmente configuráveis considerando que uma das colunas da matriz \mathbf{X} é tal que todos os valores são iguais a 1.

Proposição 2.1.1 *Os parâmetros (de regressão) do modelo são estimados de forma iterativa, tomando na k -ésima iteração,*

$$\widehat{\boldsymbol{\beta}}^{(k+1)} = \widehat{\boldsymbol{\beta}}^{(k)} + \left[\widehat{\mathbf{K}}_{\beta}^{-1} \right]^{(k)} \widehat{\mathbf{U}}_{\beta}^{(k)}$$

com $\widehat{\mathbf{U}}_{\beta}^{(k)}$ a estimativa na k -ésima iteração do vetor escore,

$$\mathbf{U}_{\beta} = \mathbf{X}'(\mathbf{Z}/s - d - \boldsymbol{\mu}),$$

e $\widehat{\mathbf{K}}_{\beta}$ a estimativa na k -ésima iteração da matriz de informação (de Fisher),

$$\mathbf{K}_{\beta} = s^{-2} \mathbf{X}' \mathbf{V} \mathbf{X},$$

em que \mathbf{X} é uma matriz $n \times p$ de posto completo com cada linha representando o vetor das covariáveis associadas a cada avaliador, \mathbf{Z} é um vetor de dimensão n das respectivas notas atribuídas ao produto sob avaliação, $\boldsymbol{\mu}$ é o vetor de esperança das notas atribuídas ao produto e \mathbf{V} é uma matriz diagonal, $n \times n$, das respectivas variâncias.

Prova: O estimador de máxima verossimilhança do vetor $\boldsymbol{\beta}$ é obtido derivando e igualando a zero a função de log-verossimilhança. Assim, calculando a derivada de primeira ordem da função de log-verossimilhança com relação ao vetor $\boldsymbol{\beta}$ obtém-se

$$\begin{aligned} \mathbf{U}_{\beta} &= \frac{\partial}{\partial \boldsymbol{\beta}} \sum_{i=1}^n \left\{ \left(\frac{z_i}{s} - d \right) \eta_i - (m-d) \ln(e^{\eta_i} + 1) \right\} \\ &= \sum_{i=1}^n \left\{ \left(\frac{z_i}{s} - d \right) \frac{\partial}{\partial \boldsymbol{\beta}} \eta_i - (m-d) \frac{\partial}{\partial \boldsymbol{\beta}} \ln(e^{\eta_i} + 1) \right\} \\ &= \sum_{i=1}^n \left(\frac{\partial}{\partial \boldsymbol{\beta}} \eta_i \right) \left\{ \left(\frac{z_i}{s} - d \right) - (m-d) \frac{e^{\eta_i}}{e^{\eta_i} + 1} \right\} \\ &= \sum_{i=1}^n \mathbf{x}'_i \left\{ \left(\frac{z_i}{s} - d \right) - (m-d) \frac{e^{\eta_i}}{e^{\eta_i} + 1} \right\}. \end{aligned}$$

Substituindo

$$\eta_i = \ln \left(\frac{\mu_i}{m-d-\mu_i} \right)$$

se obtém

$$\begin{aligned} \mathbf{U}_{\beta} &= \sum_{i=1}^n \mathbf{x}'_i \left\{ \left(\frac{z_i}{s} - d \right) - (m-d) \frac{\frac{\mu_i}{m-d-\mu_i}}{\frac{\mu_i}{m-d-\mu_i} + 1} \right\} \\ &= \sum_{i=1}^n \mathbf{x}'_i \left\{ \left(\frac{z_i}{s} - d \right) - (m-d) \frac{\frac{\mu_i}{m-d-\mu_i}}{\frac{\mu_i+m-d-\mu_i}{m-d-\mu_i}} \right\} \\ &= \sum_{i=1}^n \mathbf{x}'_i \left\{ \left(\frac{z_i}{s} - d \right) - (m-d) \frac{\mu_i}{m-d} \right\} \\ &= \sum_{i=1}^n \mathbf{x}'_i \left\{ \frac{z_i}{s} - d - \mu_i \right\}. \end{aligned}$$

Escrevendo esta expressão em forma matricial, segue a expressão do vetor escore apresentada na proposição.

Para igualar a derivada acima à zero (para obter o ponto de máximo da função) é possível aplicar o método de Newton-Raphson. Para isto, é necessário calcular a derivada de segunda ordem da função de log-verossimilhança com relação ao vetor β ,

$$\begin{aligned}
\mathbf{K}_\beta &= -\frac{\partial}{\partial \beta'} \sum_{i=1}^n \mathbf{x}'_i \left\{ \left(\frac{z_i}{s} - d \right) - (m-d) \frac{e^{\eta_i}}{e^{\eta_i} + 1} \right\} \\
&= -\sum_{i=1}^n \mathbf{x}'_i \left\{ (m-d) \frac{\partial}{\partial \beta'} \left(\frac{e^{\eta_i}}{e^{\eta_i} + 1} \right) \right\} \\
&= \sum_{i=1}^n \mathbf{x}'_i \left\{ \frac{(m-d)}{(e^{\eta_i} + 1)^2} \left[(e^{\eta_i} + 1) \left(\frac{\partial e^{\eta_i}}{\partial \beta'} \right) - e^{\eta_i} \left(\frac{\partial (e^{\eta_i} + 1)}{\partial \beta'} \right) \right] \right\} \\
&= \sum_{i=1}^n \mathbf{x}'_i \left\{ \frac{(m-d)}{(e^{\eta_i} + 1)^2} \left[e^{\eta_i} (e^{\eta_i} + 1) \left(\frac{\partial \eta_i}{\partial \beta'} \right) - e^{2\eta_i} \left(\frac{\partial \eta_i}{\partial \beta'} \right) \right] \right\} \\
&= \sum_{i=1}^n \mathbf{x}'_i \left\{ \frac{(m-d)}{(e^{\eta_i} + 1)^2} [e^{\eta_i} (e^{\eta_i} + 1) - e^{2\eta_i}] \right\} \left(\frac{\partial \eta_i}{\partial \beta'} \right) \\
&= \sum_{i=1}^n \mathbf{x}'_i \left\{ \frac{(m-d)}{(e^{\eta_i} + 1)^2} [e^{\eta_i} (e^{\eta_i} + 1) - e^{2\eta_i}] \right\} \mathbf{x}_i \\
&= \sum_{i=1}^n \mathbf{x}'_i \left\{ \frac{(m-d)}{(e^{\eta_i} + 1)^2} [e^{2\eta_i} + e^{\eta_i} - e^{2\eta_i}] \right\} \mathbf{x}_i \\
&= \sum_{i=1}^n \mathbf{x}'_i \left\{ \frac{(m-d)e^{\eta_i}}{(e^{\eta_i} + 1)^2} \right\} \mathbf{x}_i \\
&= \frac{1}{s^2} \sum_{i=1}^n \mathbf{x}'_i \left\{ s^2 \frac{(m-d)e^{\eta_i}}{(e^{\eta_i} + 1)^2} \right\} \mathbf{x}_i.
\end{aligned}$$

Como a expressão entre chaves é a variância associada ao i -ésimo elemento amostral, segue a expressão para a matriz de informação (de Fisher) observada. O estimador para o vetor de parâmetros β segue da aplicação do método de Newton-Raphson. ■

2.2 Modelos para diferentes tipos de avaliadores

No modelo descrito na seção anterior, todos os avaliadores atribuem uma nota ao produto com a mesma distribuição. Este cenário supõe uma homogeneidade quanto aos avaliadores que pode não existir na prática. Uma diferença possível entre avaliadores, apontada por D'Elia e Piccolo (2005), é a existência de alguns avaliadores que atribuem notas ao acaso, sem verificar o quanto preferem o produto sob avaliação. O modelo CUB (*Combined Uniform and shifted Binomial*), definido pelos autores citados, acomoda estes diferentes tipos de avaliadores construindo um modelo de mistura para o caso em que um rol de produtos é ordenado quanto à preferência (ou quando as notas são atribuídas na sequência $1, 2, \dots, m$). Os avaliadores que ordenam os produtos considerando alguma preferência são modelados de acordo com um componente *binomial deslocado* enquanto

que os avaliadores que ordenam os produtos ao acaso (retirando-os de uma urna) são modelados de acordo com um componente *uniforme discreto*. [Gottard, Iannario e Piccolo \(2016\)](#) generaliza o modelo CUB, representando a ordenação de produtos ao acaso por diferentes distribuições que serão descritas nesta seção.

A adoção de um único componente da mistura para representar os avaliadores que atribuem nota (ou posições) aos produtos considerando a preferência que possuem pelos itens também pode não ser coerente com algumas situações práticas. No caso em que os avaliadores possuem grande conhecimento a respeito do item sob avaliação é razoável supor que, aqueles que apresentam alguma preferência, serão homogêneos quanto à preferência apresentada pelo produto. Por outro lado, se os avaliadores são ‘leigos’ quanto ao item sob avaliação é razoável supor a existência de diferentes populações de avaliadores. Por exemplo, um grupo de enólogos que ordenam vinhos de uma determinada safra quanto à qualidade tenderão a classificar estes vinhos com base em uma única distribuição de probabilidade para os postos atribuídos. Por outro lado, um grupo de compradores de automóveis de uma determinada marca podem preferir um modelo ou outro por conta de características diversas (mulheres podem dar mais peso à estética de que à potência do motor, homens jovens podem ponderar bastante a potência do motor enquanto homens mais maduros podem preferir o conforto, etc.).

Nesta seção, os modelos são estendidos para acomodar diferentes *níveis de preferência* para os avaliadores que atribuem notas (ou postos) aos produtos de acordo com uma preferência predefinida. O modelo resultante possui função de probabilidade dada por

$$\mathbb{P}(Z|\boldsymbol{\eta}) = \sum_{r=1}^t \alpha_r \mathbb{P}_r(Z|\eta_r) + \alpha_0 \mathbb{P}_0(Z)$$

com $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_t\}$, $\eta_r = \mathbf{x}'\boldsymbol{\beta}_r$, $r = 1, 2, \dots, t$, $t \in \mathbb{N} \geq 1$, e

$$\mathbb{P}_r(Z|\eta_r) = \binom{m-d}{Z/s-d} \exp \left\{ \left(\frac{Z}{s} - d \right) \eta_r - (m-d) \ln(e^{\eta_r} + 1) \right\}$$

a probabilidade associada à avaliação de acordo com o r -ésimo, $r = 1, 2, \dots, t$, $t \in \mathbb{N} \geq 1$, nível de preferência, $\mathbb{P}_0(Z = z|m, s, d)$ a probabilidade associada à avaliação do item ao acaso, $m \in \mathbb{N} \geq 1$, $s \in \mathbb{R} > 0$, $0 \leq d \in \mathbb{N} < m$, $\boldsymbol{\beta} = \{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_t\}$ o conjunto dos vetores p -dimensionais $\boldsymbol{\beta}_r$, associados ao parâmetros de regressão de cada componente binomial da mistura, os parâmetros da mistura dados por $0 < \alpha_r < 1$, que indicam as proporções de avaliadores que atribuem os valores de acordo com o r -ésimo *nível de preferência* e $0 \leq \alpha_0 = 1 - \alpha_1 - \alpha_2 - \dots - \alpha_t \leq 1$ a proporção de avaliadores que atribuem os valores ao acaso.

A probabilidade associada às avaliações ao acaso são apresentadas por [Gottard, Iannario e Piccolo \(2016\)](#) para acomodar diferentes cenários de avaliação:

- *Uniforme discreta*: Supõe equiprobabilidade para os valores possíveis de avaliação dos itens. A avaliação sistemática é representada por uma distribuição que atribui equiprobabilidade para os $m - d + 1$ valores do conjunto de notas possíveis. A função de probabilidade é

$$\mathbb{P}_0(Z) = \frac{1}{m - d + 1}, \quad (2.1)$$

a esperança e a variância são, respectivamente,

$$\mathbb{E}_0(Z) = s \left(\frac{m - d}{2} \right), \quad \mathbb{V}_0(Z) = s^2 \left[\frac{(m - d + 1)^2 - 1}{12} \right].$$

- *Uniforme truncada*: Atribui probabilidade 0 para os $0 \leq T < (m - d)/2$ valores em cada extremo do conjunto de notas possíveis e equiprobabilidade para os $m - d - 2T + 1$ valores restantes. A função de probabilidade é

$$\mathbb{P}_0(Z) = \frac{1}{m - d - 2T + 1}, \quad (2.2)$$

a esperança e a variância são, respectivamente,

$$\mathbb{E}_0(Z) = s \left(\frac{m - d}{2} \right), \quad \mathbb{V}_0(Z) = s^2 \left[\frac{(m - d - 2T + 1)^2 - 1}{12} \right].$$

- *Uniforme limitada à esquerda*: Atribui probabilidade 0 para os $0 \leq T < m - d$ menores valores do conjunto de notas possíveis e equiprobabilidade para os $m - d - T + 1$ valores restantes. A função de probabilidade é

$$\mathbb{P}_0(Z) = \frac{1}{m - d - T + 1}, \quad (2.3)$$

a esperança e a variância são, respectivamente,

$$\mathbb{E}_0(Z) = s \left(\frac{m - d - T}{2} \right), \quad \mathbb{V}_0(Z) = s^2 \left[\frac{(m - d - T + 1)^2 - 1}{12} \right].$$

- *Uniforme limitada à direita*: Atribui probabilidade 0 para os $0 \leq T < m - d$ maiores valores do conjunto de notas possíveis e equiprobabilidade para os $m - d - T + 1$ valores restantes. A função de probabilidade é

$$\mathbb{P}_0(Z) = \frac{1}{m - d - T + 1}, \quad (2.4)$$

a esperança e a variância são, respectivamente,

$$\mathbb{E}_0(Z) = s \left(\frac{m - d - T}{2} \right), \quad \mathbb{V}_0(Z) = s^2 \left[\frac{(m - d - T + 1)^2 - 1}{12} \right].$$

Com relação ao modelo de mistura para as notas (ou postos) atribuídos aos produtos (serviços), a esperança e a variância dos valores atribuídos aos itens são, respectivamente,

$$\begin{aligned}\mathbb{E}(Z|\boldsymbol{\eta}) &= \sum_{r=1}^t \alpha_r s \left(\frac{(m-d)e^{\eta_r}}{1+e^{\eta_r}} + d \right) + \alpha_0 \mathbb{E}_0(Z), \\ \mathbb{V}(Z|\boldsymbol{\eta}) &= \sum_{r=1}^t \alpha_r s^2 \left[\left(\frac{(m-d)e^{\eta_r}}{1+e^{\eta_r}} + d \right)^2 + \frac{(m-d)e^{\eta_r}}{(1+e^{\eta_r})^2} \right] + \alpha_0 (\mathbb{E}_0(Z) + \mathbb{V}_0(Z)) \\ &\quad - \left[\sum_{r=1}^t \alpha_r s \left(\frac{(m-d)e^{\eta_r}}{1+e^{\eta_r}} + d \right) + \alpha_0 \mathbb{E}_0(Z) \right]^2.\end{aligned}$$

em que $\eta_r = \mathbf{x}'\boldsymbol{\beta}_r$.

A função de log-verossimilhança do modelo é

$$\ell(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{z}, \mathbf{x}) = \sum_{i=1}^n \ln \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z_i | \eta_{ri}) + \alpha_0 \mathbb{P}_0(z_i) \right]$$

e a matriz de variância dos estimadores dos parâmetros é dada por

$$\widehat{\mathbf{V}} = \begin{bmatrix} \widehat{\mathbf{A}} & \widehat{\mathbf{B}}_{(1)} & \widehat{\mathbf{B}}_{(2)} & \cdots & \widehat{\mathbf{B}}_{(t)} \\ \widehat{\mathbf{B}}'_{(1)} & \widehat{\mathbf{C}}_{(11)} & \widehat{\mathbf{C}}_{(12)} & \cdots & \widehat{\mathbf{C}}_{(1t)} \\ \widehat{\mathbf{B}}'_{(2)} & \widehat{\mathbf{C}}_{(21)} & \widehat{\mathbf{C}}_{(22)} & \cdots & \widehat{\mathbf{C}}_{(2t)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{B}}'_{(t)} & \widehat{\mathbf{C}}_{(t1)} & \widehat{\mathbf{C}}_{(t2)} & \cdots & \widehat{\mathbf{C}}_{(tt)} \end{bmatrix}^{-1} \quad (2.5)$$

em que $\widehat{\mathbf{A}}$ é uma matriz $t \times t$ cujo elemento \widehat{A}_{uv} , $u, v = 1, 2, \dots, t$, associado aos parâmetros (α_u, α_v) , é

$$\widehat{A}_{uv} = \sum_{i=1}^n \prod_{r \in \{u,v\}} \left(\frac{\mathbb{P}_r(z_i | \eta_{ri}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right),$$

$\widehat{\mathbf{B}}_{(u)}$, $u = 1, 2, \dots, t$, é uma matriz $t \times p$, cuja v -ésima linha, $\widehat{\mathbf{B}}_{(uv)}$, $u = 1, 2, \dots, t$, associada aos parâmetros $(\alpha_u, \boldsymbol{\beta}_v)$, é

$$\widehat{\mathbf{B}}_{(uv)} = \sum_{i=1}^n (\mathbf{x}'_i) \left(\frac{\alpha_v \mathbb{P}_v(z_i | \eta_{vi})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{\mathbb{P}_u(z_i | \eta_{ui}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{vi}}}{1+e^{\eta_{vi}}} \right)$$

e $\widehat{\mathbf{C}}_{(uv)}$ é uma matriz $p \times p$, associada aos parâmetros $(\boldsymbol{\beta}_u, \boldsymbol{\beta}_v)$, dada por

$$\widehat{\mathbf{C}}_{(uv)} = \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}'_i) \prod_{r \in \{u,v\}} \left(\frac{\alpha_r \mathbb{P}_r(z_i | \eta_{ri})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ri}}}{1+e^{\eta_{ri}}} \right).$$

com $\mathbb{I}_{(u=v)}$ a função indicadora que vale 1 se $u = v$ e 0 caso contrário, $u, v = 1, 2, \dots, t$.

A função de log-verossimilhança aumentada do modelo, usada para obter os estimadores dos parâmetros, é dada por

$$\ell^*(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{z}, \mathbf{x}) = \sum_{i=1}^n \sum_{r=1}^t g_{ri} \left[\left(\frac{z_i}{s} - d \right) \eta_{ri} - (m-d) \ln(1 + e^{\eta_{ri}}) \right] + C,$$

em que $G_{ri} = g_{ri} \in \{0, 1\}$, $r = 1, 2, \dots, t$, $i = 1, 2, \dots, n$, $1 - g_{0i} = g_{1i} + g_{2i} + \dots + g_{ti}$, são variáveis aleatórias (latentes) indicadoras de que a avaliação do i -ésimo item foi realizada com base no r -ésimo nível predefinido de preferência (de acordo com o componente da família binomial) e C é uma constante que independe dos vetores $\beta_1, \beta_2, \dots, \beta_t$. Tomando a esperança $\mathbb{E}(G_{ri}) = \omega_{ri}$ obtém-se a função

$$Q(\beta) = \sum_{i=1}^n \sum_{r=1}^t \omega_{ri} \left[\left(\frac{z_i}{s} - d \right) \eta_{ri} - (m - d) \ln(1 + e^{\eta_{ri}}) \right] + C,$$

com $0 \leq \omega_{ri} \leq 1$ e $\omega_{0i} + \omega_{1i} + \dots + \omega_{ti} = 1$.

O parâmetro da mistura é estimado, aplicando o algoritmo EM (Dempster; Laird; Rubin, 1977), tomando

$$\hat{\omega}_{ri}^{(k+1)} = \frac{\hat{\alpha}_r^{(k)} \mathbb{P}_r(z_i | \hat{\eta}_{ri}^{(k)})}{\mathbb{P}(z_i | \hat{\eta}_i^{(k)})}, \quad \hat{\alpha}_r^{(k+1)} = \frac{1}{n} \sum_{i=1}^n \hat{\omega}_{ri}^{(k+1)}$$

e os parâmetros da regressão são estimados de acordo com o Corolário 2.2.1.

Corolário 2.2.1 *No r -ésimo componente binomial do modelo, os parâmetros de regressão são estimados iterativamente por*

$$\hat{\beta}_r^{(k+1)} = \hat{\beta}_r^{(k)} + \left[\hat{\mathbf{K}}_{\beta_r}^{-1} \right]^{(k)} \hat{\mathbf{U}}_{\beta_r}^{(k)}$$

em que

$$\mathbf{U}_{\beta_r} = (\mathbf{W}\mathbf{X})'(\mathbf{Z}/s - d - \boldsymbol{\mu}), \quad \mathbf{K}_{\beta_r} = s^{-2}(\mathbf{W}^{1/2}\mathbf{X})'\mathbf{V}(\mathbf{W}^{1/2}\mathbf{X})$$

com \mathbf{W} é uma matriz diagonal $n \times n$ com componentes ω_{ri} , $i = 1, 2, \dots, n$, \mathbf{X} é uma matriz $n \times p$ de covariáveis associadas às $n \geq 1$ observações, \mathbf{Z} é um vetor de dimensão n das respostas de cada avaliador e $\boldsymbol{\mu}$ é um vetor de dimensão n das esperanças associadas aos n avaliadores.

Prova: O resultado segue da Proposição 2.1.1, observando que na função $Q(\beta)$ os termos associados à cada componente binomial do modelo estão multiplicados pelos pesos ω_{ri} , $r = 1, 2, \dots, t$, $i = 1, 2, \dots, n$. ■

2.3 Inflação em um valor da escala

Um avaliador que atribui notas aos produtos ao acaso pode, para simplificar a atribuição destas notas, fixar um valor dentre aqueles possíveis e atribuir sempre este valor à todos os itens avaliados. É comum que isto ocorra, por exemplo, com a atribuição da nota mínima, da nota mediana ou da nota máxima. Com o objetivo de acomodar a possibilidade de um grupo de avaliadores atribuírem uma nota fixada, Iannario (2012b)

propõe uma extensão ao modelo CUB com a possibilidade dos avaliadores atribuírem sistematicamente um valor fixado. Em tal modelo, as avaliações podem ser realizadas (1) com base em uma preferência predefinida, (2) ao acaso de acordo com uma distribuição de probabilidade que represente um comportamento sistemático ou (3) atribuindo um valor fixado da escala (distribuição degenerada).

De forma análoga, é possível inserir no modelo apresentado na seção 2.2 um componente para acomodar atribuições de um valor $c \in \{z \in \mathbb{N} : z = s \cdot (y + d), y = 0, 1, \dots, m - d, s \in \mathbb{R} > 0, 0 \leq d \in \mathbb{N} < m\}$ fixado. Assim, o modelo fica definido por

$$\mathbb{P}(z|\boldsymbol{\eta}) = \varphi \mathbb{I}_{(z=c)} + (1 - \varphi) \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z|\boldsymbol{\eta}_r) + \alpha_0 \mathbb{P}_0(z) \right]$$

com $0 \leq \varphi \leq 1$, o parâmetro da mistura entre o componente degenerado do modelo e os demais termos (também um modelo de mistura), $0 \leq \alpha_r \leq 1$, $r = 0, 1, \dots, t$, $\alpha_0 + \alpha_1 + \dots + \alpha_t = 1$, parâmetros da mistura que modela as avaliações de não inflação do modelo, $\mathbb{I}_{(z=c)}$ é a função indicadora que vale 1 se $z = c$ ou 0 caso contrário, $\mathbb{P}_r(z|\boldsymbol{\eta}_r)$ é a componente *binomial redimensionada e deslocada* que representa a avaliação de acordo com o r -ésimo nível de preferência e $\mathbb{P}_0(z)$ é uma distribuição que representa uma avaliação ao acaso (uma das distribuições apresentadas na seção anterior). Observe que o modelo proposto é a combinação de 2 misturas finitas de distribuições. A primeira mistura, modulada pelo parâmetro φ , possui um componente degenerado em c e um componente que é uma mistura de distribuições. A segunda mistura, modulada pelos parâmetros $\alpha_0, \alpha_1, \dots, \alpha_t$, possui t componentes binomiais redimensionados e deslocados e 1 componente referente às avaliações ao acaso.

A esperança e a variância dos valores atribuídos aos itens (notas ou postos) são, respectivamente,

$$\begin{aligned} \mathbb{E}(Z|\boldsymbol{\eta}) &= \varphi c + s(1 - \varphi) \sum_{r=1}^t \alpha_r \left(\frac{(m - d)e^{\eta_r}}{1 + e^{\eta_r}} + d \right) + (1 - \varphi)\alpha_0 \mathbb{E}_0(Z), \\ \mathbb{V}(Z|\boldsymbol{\eta}) &= s^2(1 - \varphi) \sum_{r=1}^t \alpha_r \left[\frac{(m - d)e^{\eta_r}}{(1 + e^{\eta_r})^2} + \left(\frac{(m - d)e^{\eta_r}}{1 + e^{\eta_r}} + d \right)^2 \right] \\ &\quad - \left(\varphi c + s(1 - \varphi) \sum_{r=1}^t \alpha_r \left(\frac{(m - d)e^{\eta_r}}{1 + e^{\eta_r}} + d \right) + (1 - \varphi)\alpha_0 \mathbb{E}_0(Z) \right)^2 \\ &\quad + (1 - \varphi)\alpha_0 (V_0(Z) + \mathbb{E}_0(Z)) + \varphi c^2 \end{aligned}$$

em que $\mathbb{E}_0(Z)$ e $\mathbb{V}_0(Z)$ são, respectivamente, a esperança e a variância da resposta supondo uma avaliação ao acaso (conforme descritas na seção anterior) e $0 \leq \varphi \leq 1$ é o parâmetro da mistura do componente degenerado e (a mistura de) os demais componentes do modelo.

Para simplificar a estimação do modelo, é possível reescrever o modelo como

$$\mathbb{P}(Z|\boldsymbol{\eta}) = \tau_0 \mathbb{P}_0(Z) + \sum_{r=1}^t \tau_r \mathbb{P}_r(Z|\boldsymbol{\eta}_r) + \left(1 - \sum_{r=0}^t \tau_r \right) \mathbb{I}_{(z=c)}$$

com $0 \leq \tau_r \leq 1$, $r = 0, 1, \dots, t$, $\tau_0 + \tau_1 + \dots + \tau_t \leq 1$, em que

$$\varphi = 1 - \sum_{r=0}^t \tau_r, \quad \alpha_r = \frac{\tau_r}{1 - \varphi}, \quad r = 0, 1, \dots, t.$$

A função de log-verossimilhança do modelo é

$$\ell(\boldsymbol{\tau}, \boldsymbol{\beta} | \mathbf{z}, \mathbf{x}) = \sum_{i=1}^n \ln \left[\tau_0 \mathbb{P}_0(z_i) + \sum_{r=1}^t \tau_r \mathbb{P}_r(z_i | \boldsymbol{\eta}_{ri}) + \left(1 - \sum_{r=0}^t \tau_r \right) \mathbb{I}_{(z_i=c)} \right]$$

e a matriz de variância dos estimadores dos parâmetros é dada por

$$\widehat{\mathbf{V}} = \frac{1}{n} \begin{bmatrix} \widehat{\mathbf{A}} & \widehat{\mathbf{B}}_{(1)} & \widehat{\mathbf{B}}_{(2)} & \cdots & \widehat{\mathbf{B}}_{(t)} \\ \widehat{\mathbf{B}}'_{(1)} & \widehat{\mathbf{C}}_{(11)} & \widehat{\mathbf{C}}_{(12)} & \cdots & \widehat{\mathbf{C}}_{(1t)} \\ \widehat{\mathbf{B}}'_{(2)} & \widehat{\mathbf{C}}_{(21)} & \widehat{\mathbf{C}}_{(22)} & \cdots & \widehat{\mathbf{C}}_{(2t)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{B}}'_{(t)} & \widehat{\mathbf{C}}_{(t1)} & \widehat{\mathbf{C}}_{(t2)} & \cdots & \widehat{\mathbf{C}}_{(tt)} \end{bmatrix}^{-1} \quad (2.6)$$

$$U_{\tau_0} = \sum_{i=1}^n \frac{\mathbb{P}_0(z_i) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)}, \quad (2.7)$$

a r -ésima posição, $r = 1, 2, \dots, t$, dada por

$$U_{\alpha_r} = \sum_{i=1}^n \frac{\mathbb{P}_r(z_i | \boldsymbol{\eta}_{ri}) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \quad (2.8)$$

e

$$U_{\boldsymbol{\beta}_r(\boldsymbol{\theta})'} = \sum_{i=1}^n (\mathbf{x}'_i) \left(\frac{\tau_r \mathbb{P}_r(z_i | \boldsymbol{\eta}_{ri})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} \right). \quad (2.9)$$

em que $\widehat{\mathbf{A}}$ é uma matriz de dimensão $(t+1) \times (t+1)$ cujo elemento \widehat{A}_{uv} , $u, v = 0, 1, 2, \dots, t$, associado aos parâmetros (τ_u, τ_v) , é

$$\widehat{A}_{uv} = \begin{cases} \sum \left(\frac{\mathbb{P}_0(z_i) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right)^2 & u = 0 \text{ e } v = 0 \\ \sum \left(\frac{\mathbb{P}_0(z_i) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{\mathbb{P}_v(z_i | \boldsymbol{\eta}_{vi}) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) & u = 0 \text{ e } v \neq 0 \\ \sum \left(\frac{\mathbb{P}_u(z_i | \boldsymbol{\eta}_{ui}) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{\mathbb{P}_0(z_i) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) & u \neq 0 \text{ e } v = 0 \\ \sum \left(\frac{\mathbb{P}_u(z_i | \boldsymbol{\eta}_{ui}) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{\mathbb{P}_v(z_i | \boldsymbol{\eta}_{vi}) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) & u \neq 0 \text{ e } v \neq 0 \end{cases}$$

$\widehat{\mathbf{B}}_{(u)}$, $u = 1, 2, \dots, t$, é uma matriz $(t+1) \times p$, cujo elemento $\widehat{\mathbf{B}}_{(uv)}$, $v = 1, 2, \dots, p$, associado aos parâmetros $(\tau_u, \boldsymbol{\beta}_v)$, é

$$\widehat{\mathbf{B}}_{(uv)} = \begin{cases} \sum (\mathbf{x}'_i) \left(\frac{\tau_u \mathbb{P}_u(z_i | \boldsymbol{\eta}_{ui})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ui}}}{1 + e^{\eta_{ui}}} \right) \left(\frac{\mathbb{P}_0(z_i) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) & u = 0 \\ \sum (\mathbf{x}'_i) \left(\frac{\tau_u \mathbb{P}_u(z_i | \boldsymbol{\eta}_{ui})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ui}}}{1 + e^{\eta_{ui}}} \right) \left(\frac{\mathbb{P}_v(z_i | \boldsymbol{\eta}_{vi}) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) & u \neq 0 \end{cases}$$

e $\widehat{\mathbf{C}}_{(uv)}$ é uma matriz $p \times p$, associada aos parâmetros $(\boldsymbol{\beta}_u, \boldsymbol{\beta}_v)$, dada por

$$\widehat{\mathbf{C}}_{(uv)} \prod_{r \in \{u,v\}} = \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \left(\frac{\tau_r \mathbb{P}_r(z_i | \eta_{ri})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} \right).$$

A função de log-verossimilhança aumentada do modelo, usada para obter os estimadores dos parâmetros, é dada por

$$\ell^*(\boldsymbol{\tau}, \boldsymbol{\beta} | \mathbf{z}, \mathbf{x}) = \sum_{i=1}^n \sum_{r=1}^t g_{ri} \left[\left(\frac{z_i}{s} - d \right) \eta_{ri} - (m-d) \ln(1 + e^{\eta_{ri}}) \right] + C$$

em que, o termo correspondente às observações fixas em $Z = c$ é igual a 0, $G_{0i} = g_{0i}$ é a variável latente indicadora de que a i -ésima observação foi gerada por uma avaliação ao acaso, $G_{ri} = g_{ri}$ é a variável latente indicadora de que a i -ésima observação foi gerada de acordo com o r -ésimo ($r = 1, 2, \dots, t$) nível predefinido de preferência e C é uma constante que independe dos vetores de parâmetros. Tomando as esperanças $\mathbb{E}(G_{ri}) = \omega_{ri}$ obtém-se a função

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^n \sum_{r=1}^t \omega_{ri} \left[\left(\frac{z_i}{s} - d \right) \eta_{ri} - (m-d) \ln(1 + e^{\eta_{ri}}) \right] + C.$$

Os parâmetros da mistura são estimados pelo algoritmo EM (Dempster; Laird; Rubin, 1977), tomando

$$\begin{aligned} \widehat{\omega}_{0i}^{(k+1)} &= \frac{\widehat{\tau}_0^{(k)} \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \widehat{\boldsymbol{\eta}}^{(k)}),} & \widehat{\omega}_{ri}^{(k+1)} &= \frac{\widehat{\tau}_r^{(k)} \mathbb{P}_r(z_i | \widehat{\eta}_{ri}^{(k)})}{\mathbb{P}(z_i | \widehat{\boldsymbol{\eta}}^{(k)}),} \\ \widehat{\tau}_0^{(k+1)} &= \frac{1}{n} \sum_{i=1}^n \widehat{\omega}_{0i}^{(k+1)}, & \widehat{\tau}_r^{(k+1)} &= \frac{1}{n} \sum_{i=1}^n \widehat{\omega}_{ri}^{(k+1)}, \end{aligned}$$

e os parâmetros da regressão são estimados de acordo com o Corolário 2.1.1.

3 Modelos para sobredispersão

Neste capítulo será abordada uma característica dos modelos baseados em reparametrizações da distribuição binomial (como aqueles apresentados no capítulo anterior). Em tal cenário, se μ é a esperança da nota Z atribuída ao produto, então a variância de tal nota, σ^2 é uma função da esperança, $\sigma^2 = g(\mu)$. No entanto, existem cenários em que tal relação entre o valor esperado e a variância não é desejável. Para tais cenários é necessário que a variância da nota, σ^2 , não dependa exclusivamente da esperança, μ ; ou seja, é interessante que exista um parâmetro adicional θ tal que $\sigma^2 = g(\mu, \theta)$.

Este capítulo define modelos para a avaliação de produtos (serviços) nos quais uma nota na escala $[0, 1, \dots, m]$, $m \in \mathbb{N} \geq 1$, é atribuída aos itens sob avaliação. O primeiro modelo, apresentado na seção 3.1, é construído a partir da distribuição *binomial dupla*, uma distribuição da classe *exponencial dupla* proposta por Efron (1986). O segundo modelo, apresentado na seção 3.2, é construído a partir da distribuição *binomial multiplicativa* proposta por Altham (1978). Esquemas generalizados de atribuição de notas aos produtos e serviços para os modelos aqui apresentados podem ser facilmente obtidos por reparametrização da variável resposta original.

3.1 O modelo binomial duplo

A família de distribuições *exponencial dupla* é apresentada por Efron (1986). A versão correspondente à distribuição binomial nesta família é chamada de distribuição *binomial dupla*. A definição desta distribuição é bastante simples. Se Z é uma variável aleatória com distribuição binomial (usual) com parâmetros $m \geq 1$ e $0 < \pi < 1$, a função de probabilidade associada à esta distribuição é dada por

$$\mathbb{P}(Z|\pi) = \binom{m}{Z} \pi^Z (1 - \pi)^{m-Z} = \binom{m}{Z} \left(\frac{\pi}{1 - \pi} \right)^Z (1 - \pi)^m.$$

Tomando a transformação $Y = Z/m$ (de modo que $Y \in \{0, 1/m, 2/m, \dots, 1\}$) obtém-se a função de probabilidade correspondente como (Efron, 1986, equação (2.3))

$$\mathbb{P}(Y|\pi) = \binom{m}{mY} \left(\frac{\pi}{1 - \pi} \right)^{mY} (1 - \pi)^m.$$

Tomando $\eta = \ln(\pi/(1 - \pi))$ e $\psi(\pi) = -\ln(2(1 - \pi))$ é possível escrever a função de probabilidade como

$$\begin{aligned} g_{\pi,m}(Y) &= \binom{m}{mY} 2^{-m} \exp \left[m \left(Y \ln \left(\frac{\pi}{1 - \pi} \right) + \ln(2(1 - \pi)) \right) \right] \\ &= \binom{m}{mY} 2^{-m} \exp [m (\eta Y - \psi(\pi))]. \end{aligned}$$

A distribuição binomial dupla é definida por [Efron \(1986, equação \(2.4\)\)](#), acrescentando um novo parâmetro $\theta > 0$, que modela a variância da distribuição, de tal forma que $\mathbb{P}(Y|\theta, \pi) = c(\pi, \theta, m)\theta^{1/2}g_{\pi,m}(Y)^\theta g_{Y,m}(Y)^{1-\theta}$. Assim,

$$\mathbb{P}(Y|\theta, \pi) = c(\pi, \theta, m)\theta^{1/2} \binom{m}{mY} \left[\left(\frac{\pi}{1-\pi} \right)^{mY} (1-\pi)^m \right]^\theta \left(\left(\frac{mY}{m} \right)^{mY} \left(1 - \frac{mY}{m} \right)^{m-mY} \right)^{1-\theta}$$

Esta função de probabilidade pode ser reescrita, retornando à variável original ($Z = mY$), como

$$\mathbb{P}(Z|\theta, \pi) = c(\pi, \theta, m)\theta^{1/2} \binom{m}{Z} \left[\left(\frac{\pi}{1-\pi} \right)^Z (1-\pi)^m \right]^\theta \left(\left(\frac{Z}{m} \right)^Z \left(1 - \frac{Z}{m} \right)^{m-Z} \right)^{1-\theta}$$

em que $0 < \pi < 1$, $0 < \theta \leq 1$, $m \geq 1$ e $Z \in \{0, 1, \dots, m\}$. De acordo com [Efron \(1986, fato 1, pág. 711\)](#), a função $c(\pi, \theta, m) \approx 1$, de tal modo que

$$\mathbb{P}(Z|\theta, \pi) \approx \theta^{1/2} \binom{m}{Z} \left[\left(\frac{\pi}{1-\pi} \right)^Z (1-\pi)^m \right]^\theta \left(\left(\frac{Z}{m} \right)^Z \left(1 - \frac{Z}{m} \right)^{m-Z} \right)^{1-\theta}$$

e de acordo com [Efron \(1986, fato 2, pág. 711\)](#), a esperança da variável aleatória Z é $\mathbb{E}(Z) \approx m\pi$ e a variância é $\mathbb{V}(Z) \approx m\pi(1-\pi)/\theta$. O modelo binomial usual é obtido tomando $\theta = 1$. Observe que, tomando

$$\eta = \ln \left(\frac{\pi}{1-\pi} \right)$$

a função de probabilidade pode ser reescrita como

$$\mathbb{P}(Z|\theta, \eta) \approx \theta^{1/2} g_m(Z) \exp \left\{ m\theta \left[\left(\frac{Z}{m} \right) \eta - \ln(1 + e^\eta) - \ln \left[\left(\frac{Z}{m} \right)^{Z/m} \left(1 - \frac{Z}{m} \right)^{1-Z/m} \right] \right] \right\}$$

com

$$g_m(Z) = \binom{m}{Z} \left(\frac{Z}{m} \right)^Z \left(1 - \frac{Z}{m} \right)^{m-Z}.$$

Fixando o parâmetro θ a distribuição de Z condicional ao parâmetro η é uma distribuição da família exponencial com um único parâmetro $\eta \in \mathbb{R}$, de acordo com [Efron \(1986, fato 3, pág. 711\)](#),

$$\mathbb{P}(Z|\eta) \approx \theta^{1/2} g_m(Z) \exp \left\{ m\theta \left[\left(\frac{Z}{m} \right) \eta - \ln(1 + e^\eta) - \ln \left(\left(\frac{Z}{m} \right)^{Z/m} \left(1 - \frac{Z}{m} \right)^{1-Z/m} \right) \right] \right\}.$$

De forma análoga, fixando o parâmetro η a distribuição de Z condicional ao parâmetro θ é uma distribuição da família exponencial com um único parâmetro $0 \leq \theta \leq 1$, de acordo com [Efron \(1986, fato 5, pág. 712\)](#),

$$\mathbb{P}(Z|\theta) \approx \theta^{1/2} g_m(Z) \exp \left\{ -m\theta \left[\ln \left(\left(\frac{Z}{m} \right)^{Z/m} \left(1 - \frac{Z}{m} \right)^{1-Z/m} \right) - \left(\frac{Z}{m} \right) \eta + \ln(1 + e^\eta) \right] \right\}.$$

O estimador de máxima verossimilhança para o parâmetro θ , de acordo com [Efron \(1986, fato 7, pág. 712\)](#), é dado por

$$\hat{\theta} = \frac{n}{2} \left[\sum_{i=1}^n \left[\ln \left(\left(\frac{z_i}{m} \right)^{z_i} \left(1 - \frac{z_i}{m} \right)^{m-z_i} \right) - z_i \eta_i + m \ln(1 + e^{\eta_i}) \right] \right]^{-1}$$

3.1.1 Incorporando covariáveis ao modelo

Para incorporar covariáveis ao modelo *binomial duplo* basta tomar as funções

$$\eta = \ln\left(\frac{\pi}{1-\pi}\right) = \mathbf{x}'\boldsymbol{\beta}.$$

3.1.2 Modelo para avaliação de produtos

Considerando que os produtos sob avaliação recebem uma nota inteira entre 0 e $m \geq 1$. Neste cenário, a função de probabilidade considerando que as avaliações com base em preferências predefinidas seguem uma distribuição binomial dupla é dada por

$$\begin{aligned} \mathbb{P}(Z|\boldsymbol{\theta}, \boldsymbol{\eta}) &= \sum_{r=1}^t \alpha_r g_m(Z) \theta_r^{1/2} \\ &\quad \exp\left\{-m\theta_r \left[\ln\left(\left(\frac{Z}{m}\right)^{Z/m} \left(1 - \frac{Z}{m}\right)^{1-Z/m}\right) - \left(\frac{Z}{m}\right) \eta_r + \ln(1 + e^{\eta_r}) \right]\right\} \\ &\quad + \alpha_0 \mathbb{P}_0(Z) \end{aligned}$$

em que $\boldsymbol{\eta}$ é o conjunto dos valores $\eta_r = \mathbf{x}'\boldsymbol{\beta}_r$, $r = 1, 2, \dots, t$, com \mathbf{x} um vetor de dimensão p de covariáveis, e $\boldsymbol{\theta}$ é o conjunto dos valores θ_r , $\boldsymbol{\beta}_r$ o vetor de dimensão p de coeficientes de regressão associados a cada resposta e $\alpha_1, \alpha_2, \dots, \alpha_t$ os parâmetros da mistura associados com cada componente binomial duplo tais que $\alpha_r > 0$, $\alpha_1 + \alpha_2 + \dots + \alpha_t \leq 1$, e $\alpha_0 = 1 - \alpha_1 - \alpha_2 - \dots - \alpha_t$ é o parâmetro da mistura associado à avaliação ao acaso (sem considerar a preferência do avaliador) e $\mathbb{P}_0(Z)$ é uma das distribuições apresentadas nas equações 2.1 a 2.4 com $s = 1$ e $d = 0$.

A esperança e a variância da variável resposta Z condicional às covariáveis são, respectivamente,

$$\begin{aligned} \mathbb{E}(Z|\boldsymbol{\theta}, \boldsymbol{\eta}) &\approx \sum_{r=1}^t \alpha_r \frac{m e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} + \alpha_0 \mathbb{E}_0(Z), \\ \mathbb{V}(Z|\boldsymbol{\theta}, \boldsymbol{\eta}) &\approx \sum_{r=1}^t \frac{m \alpha_r e^{\eta_{ri}}}{\theta_r (1 + e^{\eta_{ri}})^2} + \alpha_0 \left(\mathbb{E}_0(Z)^2 + \mathbb{V}_0(Z) \right) \\ &\quad - \left(\sum_{r=1}^t \alpha_r \frac{m e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} + \alpha_0 \mathbb{E}_0(Z) \right)^2 \end{aligned}$$

em que $\mathbb{E}_0(Z)$ e $\mathbb{V}_0(Z)$ são, respectivamente, a esperança e a variância associadas à avaliação ao acaso.

A função de log-verossimilhança do modelo é

$$\begin{aligned} \ell(\boldsymbol{\theta}, \boldsymbol{\eta}|\mathbf{Z}, \mathbf{X}) &= \sum_{i=1}^n \ln \left[\sum_{r=1}^t \alpha_r g_m(z_i) \theta_r^{1/2} \right. \\ &\quad \exp\left\{-m\theta_r \left[\ln\left(\left(\frac{z_i}{m}\right)^{z_i/m} \left(1 - \frac{z_i}{m}\right)^{1-z_i/m}\right) - \left(\frac{z_i}{m}\right) \eta_{ri} + \ln(1 + e^{\eta_{ri}}) \right]\right\} \\ &\quad \left. + \alpha_0 \mathbb{P}_0(z_i) \right] \end{aligned}$$

e, de acordo com Efron (1986, fato 6, pág. 712), os parâmetros $\boldsymbol{\eta}$ e $\boldsymbol{\theta}$ são ortogonais. Para efeitos práticos (no cálculo da variância dos parâmetros) é possível considerar também a ortogonalidade entre $\boldsymbol{\alpha}$ e $\boldsymbol{\eta}$ e entre $\boldsymbol{\alpha}$ e $\boldsymbol{\theta}$, de modo que a matriz de variância dos parâmetros é dada por

$$\widehat{\mathbf{V}} = \begin{bmatrix} \widehat{\mathbf{A}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{B}} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{C}}_{(11)} & \widehat{\mathbf{C}}_{(12)} & \cdots & \widehat{\mathbf{C}}_{(1t)} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{C}}_{(21)} & \widehat{\mathbf{C}}_{(22)} & \cdots & \widehat{\mathbf{C}}_{(2t)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{C}}_{(t1)} & \widehat{\mathbf{C}}_{(t2)} & \cdots & \widehat{\mathbf{C}}_{(tt)} \end{bmatrix}^{-1} \quad (3.1)$$

em que $\widehat{\mathbf{A}}$ é uma matriz $t \times t$ cujo elemento \widehat{A}_{uv} , $u, v = 1, 2, \dots, t$, relacionado com os parâmetros (α_u, α_v) , é

$$\widehat{A}_{uv} = \sum_{i=1}^n \prod_{r \in \{u,v\}} \left(\frac{\mathbb{P}_r(z_i | \widehat{\phi}_r, \widehat{\eta}_{ri}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right),$$

$\widehat{\mathbf{B}}$ é uma matriz $t \times t$, cujo elemento \widehat{B}_{uv} , $u, v = 1, 2, \dots, t$, relacionado aos parâmetros (θ_u, θ_v) , de acordo com Efron (1986, fato 7, pág. 712), é

$$\widehat{B}_{uv} = \sum_{i=1}^n \prod_{r \in \{u,v\}} \left(\frac{\widehat{\alpha}_r \mathbb{P}_r(z_i | \widehat{\theta}_r, \widehat{\eta}_{ri})}{\mathbb{P}(z_i | \widehat{\theta}, \widehat{\eta}_i)} \right) \times \left[\frac{1}{2\widehat{\theta}_r} - \left[\ln \left(\left(\frac{z_i}{m} \right)^{z_i} \left(1 - \frac{z_i}{m} \right)^{m-z_i} \right) - z_i \widehat{\eta}_{ri} + m \ln(1 + e^{\widehat{\eta}_{ri}}) \right] \right],$$

e $\widehat{\mathbf{C}}_{(uv)}$ é uma matriz $p \times p$, associada aos parâmetros (β_u, β_v) , é

$$\widehat{\mathbf{C}}_{(uv)} = \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \prod_{r \in \{u,v\}} \widehat{\theta}_r \left(\frac{\alpha_r \mathbb{P}_r(z_i | \widehat{\phi}_r, \widehat{\eta}_{ri})}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right) \left(z_i - \frac{m e^{\widehat{\eta}_{ri}}}{(1 + e^{\widehat{\eta}_{ri}})} \right).$$

A função de log-verossimilhança aumentada do modelo, usada para obter os estimadores dos parâmetros, é, a menos de uma constante, dada por

$$\ell^*(\boldsymbol{\theta}, \boldsymbol{\eta} | \mathbf{Z}, \mathbf{X}) = \sum_{i=1}^n \sum_{r=1}^t g_{ri} \times \left[\frac{\ln \theta_r}{2} - m \theta_r \left[\ln \left(\left(\frac{z_i}{m} \right)^{z_i/m} \left(1 - \frac{z_i}{m} \right)^{1-z_i/m} \right) - \left(\frac{z_i}{m} \right) \eta_{ri} + \ln(1 + e^{\eta_{ri}}) \right] \right],$$

em que $G_{ri} = g_{ri} \in \{0, 1\}$, $r = 1, 2, \dots, t$, $i = 1, 2, \dots, n$, $1 - g_{0i} = g_{1i} + g_{2i} + \cdots + g_{ti}$, são variáveis aleatórias (latentes) indicadoras de que a avaliação do i -ésimo item foi realizada com base no r -ésimo nível predefinido de preferência. Tomando a esperança $\mathbb{E}(G_{ri}) = \omega_{ri}$ obtém-se a função

$$Q(\boldsymbol{\theta}, \boldsymbol{\eta} | \mathbf{Z}, \mathbf{X}) = \sum_{i=1}^n \sum_{r=1}^t \omega_{ri} \times \left[\frac{\ln \theta_r}{2} - m \theta_r \left[\ln \left(\left(\frac{z_i}{m} \right)^{z_i/m} \left(1 - \frac{z_i}{m} \right)^{1-z_i/m} \right) - \left(\frac{z_i}{m} \right) \eta_{ri} + \ln(1 + e^{\eta_{ri}}) \right] \right],$$

com $0 \leq \omega_{ri} \leq 1$ e $\omega_{0i} + \omega_{1i} + \dots + \omega_{ti} = 1$.

O parâmetro da mistura é estimado iterativamente tomando

$$\hat{\omega}_{ri}^{(k+1)} = \frac{\hat{\alpha}_r^{(k)} \mathbb{P}_r(z_i | \hat{\theta}_r^{(k)}, \hat{\eta}_{ri}^{(k)})}{\mathbb{P}(z_i | \hat{\theta}^{(k)}, \hat{\eta}_i^{(k)})}, \quad \hat{\alpha}_r^{(k+1)} = \frac{1}{n} \sum_{i=1}^n \hat{\omega}_{ri}^{(k+1)},$$

os vetores de parâmetros β_u são estimados iterativamente por

$$\hat{\beta}_u^{(k+1)} = \hat{\beta}_u^{(k)} - Q'_{\beta_u}(\hat{\theta}^{(k)}, \hat{\eta}^{(k)} | \mathbf{Z}, \mathbf{X}) \left[Q''_{\beta_u}(\hat{\theta}^{(k)}, \hat{\eta}^{(k)} | \mathbf{Z}, \mathbf{X}) \right]^{-1},$$

com

$$Q'_{\beta_u}(\theta, \eta | \mathbf{Z}, \mathbf{X}) = \sum_{i=1}^n (\mathbf{x}'_i) \omega_{ui} \left(z_i - \frac{m e^{\eta_{ui}}}{1 + e^{\eta_{ui}}} \right)$$

$$Q''_{\beta_u}(\theta, \eta | \mathbf{Z}, \mathbf{X}) = - \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}'_i) \omega_{ui} \left(\frac{m e^{\eta_{ui}}}{(1 + e^{\eta_{ui}})^2} \right)$$

e os parâmetros θ_u são estimados por

$$\hat{\theta}_u^{(k+1)} = \frac{1}{2} \left(\sum_{i=1}^n \omega_{ui} \right) \left[\sum_{i=1}^n \omega_{ui} \left[\ln \left(\left(\frac{z_i}{m} \right)^{z_i} \left(1 - \frac{z_i}{m} \right)^{m-z_i} \right) - z_i \hat{\eta}_{ui}^{(k+1)} + m \ln \left(1 + e^{\hat{\eta}_{ui}^{(k+1)}} \right) \right] \right]^{-1}.$$

em que $\beta_r^{(k)}$ é a estimativa do vetor β_r na k -ésima iteração e $\theta_r^{(k)}$ é a estimativa do parâmetro θ_r na k -ésima iteração.

3.2 O modelo binomial multiplicativo

O modelo *binomial multiplicativo* foi proposto por [Altham \(1978\)](#) e consiste em uma transformação da distribuição binomial para incorporar um parâmetro $\theta > 0$ para modelar a variância. A função de probabilidade é dada por

$$\mathbb{P}(Z | \theta, \pi) = \binom{m}{Z} \pi^Z (1 - \pi)^{m-Z} \theta^{Z(m-Z)} \left[\sum_{j=0}^m \binom{m}{j} \pi^j (1 - \pi)^{m-j} \theta^{j(m-j)} \right]^{-1}$$

Observe que se $\theta = 0$ a distribuição se degenera, em uma função que atribui probabilidade $(1 - \pi)^m / (\pi^m + (1 - \pi)^m)$ em $Z = 0$ e probabilidade $\pi^m / (\pi^m + (1 - \pi)^m)$ em $Z = m$, e se $\theta = 1$ a distribuição é a binomial usual. Se Z segue uma distribuição binomial multiplicativa, sua esperança é dada por

$$\mathbb{E}(Z | \theta, \pi) = m \left(\frac{\pi}{1 - \pi} \right) \theta^{m-1} \times$$

$$\left[\sum_{j=0}^{m-1} \binom{m-1}{j} \left(\frac{\pi}{1 - \pi} \right)^j \theta^{j(m-j-2)} \right] \left[\sum_{j=0}^m \binom{m}{j} \left(\frac{\pi}{1 - \pi} \right)^j \theta^{j(m-j)} \right]^{-1}.$$

Além disto,

$$\mathbb{E}(Z(Z-1) | \theta, \pi) = m(m-1) \left(\frac{\pi}{1 - \pi} \right)^2 \theta^{2(m-2)} \times$$

$$\left[\sum_{j=0}^{m-2} \binom{m-2}{j} \left(\frac{\pi}{1 - \pi} \right)^j \theta^{j(m-j-4)} \right] \left[\sum_{j=0}^m \binom{m}{j} \left(\frac{\pi}{1 - \pi} \right)^j \theta^{j(m-j)} \right]^{-1},$$

do qual segue que a variância é dada por

$$\mathbb{V}(Z|\theta, \pi) = \mathbb{E}(Z(Z-1)|\theta, \pi) + \mathbb{E}(Z|\theta, \pi) - [\mathbb{E}(Z|\theta, \pi)]^2.$$

3.2.1 Incorporando covariáveis ao modelo

Para incorporar covariáveis na distribuição binomial multiplicativa, considere as funções

$$\eta = \ln\left(\frac{\pi}{1-\pi}\right) = \mathbf{x}'\boldsymbol{\beta}, \quad \phi = \ln(\theta).$$

Desta forma, a função de probabilidade fica dada por

$$\mathbb{P}(Z|\phi, \eta) = \binom{m}{Z} \exp\{Z\eta - m \ln(1 + e^\eta) + Z(m-Z)\phi\} \times \left[\sum_{j=0}^m \binom{m}{j} \exp\{j\eta - m \ln(1 + e^\eta) + j(m-j)\phi\} \right]^{-1}.$$

3.2.2 Modelo para avaliação de produtos

Considerando, como no modelo anterior, que os produtos recebem notas inteiras entre 0 e $m \geq 1$, a função de probabilidade das notas, considerando que as avaliações baseadas na preferência do avaliador seguem uma distribuição binomial multiplicativa, é dada por

$$\mathbb{P}(Z|\boldsymbol{\phi}, \boldsymbol{\eta}) = \sum_{r=1}^t \alpha_r \binom{m}{Z} \exp\{Z\eta_r - m \ln(1 + e^{\eta_r}) + Z(m-Z)\phi_r\} \times \left[\sum_{j=0}^m \binom{m}{j} \exp\{j\eta_r - m \ln(1 + e^{\eta_r}) + j(m-j)\phi_r\} \right]^{-1} + \alpha_0 \mathbb{P}_0(Z)$$

em que $\eta_r = \mathbf{x}'\boldsymbol{\beta}_r$, \mathbf{x}_i é o vetor de dimensão p de covariáveis associado ao i -ésimo elemento amostral, $i = 1, 2, \dots, n$, ϕ_r é o parâmetro associado à variância da resposta, $\boldsymbol{\beta}_r$ é o vetor de dimensão p de coeficientes associados à média (esperança) da resposta e $\alpha_1, \alpha_2, \dots, \alpha_t$ são os parâmetros da mistura associados com cada componente binomial multiplicativos tais que $\alpha_r > 0$, $r = 1, 2, \dots, t$, $\alpha_1 + \alpha_2 + \dots + \alpha_t \leq 1$, $\alpha_0 = 1 - \alpha_1 - \alpha_2 - \dots - \alpha_t$ é o parâmetro da mistura associado à avaliação ao acaso (sem considerar a preferência do avaliador) e $\mathbb{P}_0(Z)$ é uma das distribuições apresentadas nas equações 2.1 a 2.4 com $s = 1$ e $d = 0$.

A esperança e a variância das notas condicionais às covariáveis são, respectivamente,

$$\mathbb{E}(Z|\boldsymbol{\phi}, \boldsymbol{\eta}) = \sum_{r=1}^t \alpha_r \mu_r(\mathbf{x}) + \alpha_0 \mathbb{E}_0(Z),$$

$$\mathbb{V}(Z|\boldsymbol{\phi}, \boldsymbol{\eta}) = \sum_{r=1}^t \alpha_r (\mu_r(\mathbf{x})^2 + \sigma_r^2(\mathbf{x})) + \alpha_0 (\mathbb{E}_0(Z)^2 + \mathbb{V}_0(Z)) - \left(\sum_{r=1}^t \alpha_r \mu_r(\mathbf{x}) + \alpha_0 \mathbb{E}_0(Z) \right)^2.$$

em que $\mathbb{E}_0(Z)$ e $V_0(Z)$ são, respectivamente, a esperança e a variância associadas à avaliação ao acaso e,

$$\begin{aligned}\mu_r(\mathbf{x}) &= \sum_{i=1}^n m e^{\eta_{ri}} \left(\frac{e^{\phi_{ri}}}{1 + e^{\phi_{ri}}} \right)^{m-1} \times \\ &\quad \left[\sum_{j=0}^{m-1} \binom{m-1}{j} e^{j\eta_{ri}} \left(\frac{e^{\phi_{ri}}}{e^{\phi_{ri}} + 1} \right)^{j(m-j-2)} \right] \left[\sum_{j=0}^m \binom{m}{j} e^{j\eta_{ri}} \left(\frac{e^{\phi_{ri}}}{e^{\phi_{ri}} + 1} \right)^{j(m-j)} \right]^{-1} \\ \mu_r^{(2)}(\mathbf{x}) &= \sum_{i=1}^n m(m-1) (e^{2\eta_{ri}}) \left(\frac{e^{\phi_{ri}}}{e^{\phi_{ri}} + 1} \right)^{2(m-2)} \times \\ &\quad \left[\sum_{j=0}^{m-2} \binom{m-2}{j} e^{j\eta_{ri}} \left(\frac{e^{\phi_{ri}}}{e^{\phi_{ri}} + 1} \right)^{j(m-j-4)} \right] \left[\sum_{j=0}^m \binom{m}{j} e^{j\eta_{ri}} \left(\frac{e^{\phi_{ri}}}{e^{\phi_{ri}} + 1} \right)^{j(m-j)} \right]^{-1} \\ \sigma_r^2(\mathbf{x}) &= \mu_r^{(2)}(\mathbf{x}) + \mu_r(\mathbf{x}) - [\mu_r(\mathbf{x})]^2\end{aligned}$$

com $\mu_r(\mathbf{x}) = \mathbb{E}_r(Z|\boldsymbol{\phi}, \boldsymbol{\eta})$, $\mu_r^{(2)}(\mathbf{x}) = \mathbb{E}_r(Z(Z-1)|\boldsymbol{\phi}, \boldsymbol{\eta})$ e $\sigma_r^2(\mathbf{x}) = \mathbb{V}_r(Z|\boldsymbol{\phi}, \boldsymbol{\eta})$.

A função de log-verossimilhança do modelo é

$$\begin{aligned}\ell(\boldsymbol{\phi}, \boldsymbol{\eta}|\mathbf{Z}, \mathbf{X}) &= \sum_{i=1}^n \ln \left[\sum_{r=1}^t \alpha_r \binom{m}{z_i} \exp \{z_i \eta_{ri} - m \ln(1 + e^{\eta_{ri}}) + z_i(m - z_i)\phi_r\} \times \right. \\ &\quad \left. \left[\sum_{j=0}^m \binom{m}{j} \exp \{j\eta_{ri} - m \ln(1 + e^{\eta_{ri}}) + j(m - j)\phi_r\} \right]^{-1} \right. \\ &\quad \left. + \alpha_0 \mathbb{P}_0(z_i) \right]\end{aligned}$$

e a matriz de variância dos parâmetros é dada por

$$\widehat{\mathbf{V}} = \begin{bmatrix} \widehat{\mathbf{A}} & \widehat{\mathbf{B}} & \widehat{\mathbf{C}}_{(1)} & \widehat{\mathbf{C}}_{(2)} & \cdots & \widehat{\mathbf{C}}_{(t)} \\ \widehat{\mathbf{B}}' & \widehat{\mathbf{D}} & \widehat{\mathbf{E}}_{(1)} & \widehat{\mathbf{E}}_{(2)} & \cdots & \widehat{\mathbf{E}}_{(t)} \\ \widehat{\mathbf{C}}_{(1)}' & \widehat{\mathbf{E}}_{(1)}' & \widehat{\mathbf{F}}_{(11)} & \widehat{\mathbf{F}}_{(12)} & \cdots & \widehat{\mathbf{F}}_{(1t)} \\ \widehat{\mathbf{C}}_{(2)}' & \widehat{\mathbf{E}}_{(2)}' & \widehat{\mathbf{F}}_{(21)} & \widehat{\mathbf{F}}_{(22)} & \cdots & \widehat{\mathbf{F}}_{(2t)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{C}}_{(t)}' & \widehat{\mathbf{E}}_{(t)}' & \widehat{\mathbf{F}}_{(t1)} & \widehat{\mathbf{F}}_{(t2)} & \cdots & \widehat{\mathbf{F}}_{(tt)} \end{bmatrix}^{-1} \quad (3.2)$$

em que $\widehat{\mathbf{A}}$ é uma matriz $t \times t$ cujo elemento \widehat{A}_{uv} , $u, v = 1, 2, \dots, t$, associado aos parâmetros (α_u, α_v) , é

$$\widehat{A}_{uv} = \sum_{i=1}^n \prod_{r \in \{u, v\}} \left(\frac{\mathbb{P}_r(z_i|\widehat{\phi}_r, \widehat{\eta}_{ri}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i|\widehat{\boldsymbol{\phi}}, \widehat{\boldsymbol{\eta}}_i)} \right),$$

$\widehat{\mathbf{B}}$, é uma matriz $t \times t$, cujo elemento \widehat{B}_{uv} , $u, v = 1, 2, \dots, t$, associada aos parâmetros (α_u, ϕ_v) , é

$$\begin{aligned}\widehat{B}_{uv} &= \sum_{i=1}^n \left(\frac{\mathbb{P}_u(z_i|\widehat{\phi}_u, \widehat{\eta}_{ui}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i|\widehat{\boldsymbol{\phi}}, \widehat{\boldsymbol{\eta}}_i)} \right) \times \\ &\quad \left(\frac{\widehat{\alpha}_v \mathbb{P}_v(z_i|\widehat{\phi}_v, \widehat{\eta}_{vi})}{\mathbb{P}(z_i|\widehat{\boldsymbol{\phi}}, \widehat{\boldsymbol{\eta}}_i)} \right) (z_i(m - z_i) - \mathbb{E}(Z(m - Z)|\widehat{\phi}_v, \widehat{\eta}_{vi}))\end{aligned}$$

$\widehat{\mathbf{C}}_{(v)}$ é uma matriz $t \times p$, cuja u -ésima linha, $u = 1, 2, \dots, t$, associada aos parâmetros (α_u, β_v) , é

$$\widehat{\mathbf{C}}_{uv} = \sum_{i=1}^n (\mathbf{x}'_i) \left(\frac{\mathbb{P}_v(z_i | \widehat{\phi}_v, \widehat{\eta}_{vi}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right) \left(\frac{\widehat{\alpha}_v \mathbb{P}_v(z_i | \widehat{\phi}_v, \widehat{\eta}_{vi})}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right) (z_i - \mathbb{E}(Z | \widehat{\phi}_v, \widehat{\eta}_{vi})),$$

$\widehat{\mathbf{D}}_{(uv)}$ é uma matriz $t \times t$, $u, v = 1, 2, \dots, t$, cujo elemento \widehat{D}_{uv} , associado aos parâmetros (ϕ_u, ϕ_v) , é

$$\widehat{D}_{uv} = \sum_{i=1}^n \prod_{r \in \{u, v\}} \left(\frac{\widehat{\alpha}_r \mathbb{P}_r(z_i | \widehat{\phi}_r, \widehat{\eta}_{ri})}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right) (z_i(m - z_i) - \mathbb{E}(Z(m - Z) | \widehat{\phi}_r, \widehat{\eta}_{ri})),$$

$\widehat{\mathbf{E}}_{(v)}$ é uma matriz $t \times p$, cuja u -ésima linha, $u = 1, 2, \dots, t$, associada aos parâmetros (ϕ_u, β_v) , é

$$\widehat{\mathbf{E}}_{(uv)} = \sum_{i=1}^n (\mathbf{x}'_i) \left(\frac{\widehat{\alpha}_u \mathbb{P}_u(z_i | \widehat{\phi}_u, \widehat{\eta}_{ui})}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right) \left(\frac{\widehat{\alpha}_v \mathbb{P}_v(z_i | \widehat{\phi}_v, \widehat{\eta}_{vi})}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right) \times \\ (z_i(m - z_i) - \mathbb{E}(Z(m - Z) | \widehat{\phi}_u, \widehat{\eta}_{ui})) \cdot (z_i - \mathbb{E}(Z | \widehat{\phi}_v, \widehat{\eta}_{vi})),$$

e $\widehat{\mathbf{F}}_{(uv)}$ são matrizes $p \times p$, $u = 1, 2, \dots, t$, $v = 1, 2, \dots, t$, associadas aos parâmetros (β_u, β_v) , dadas por

$$\widehat{\mathbf{F}}_{(uv)} = \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}'_i) \prod_{r \in \{u, v\}} \left(\frac{\widehat{\alpha}_r \mathbb{P}_r(z_i | \widehat{\phi}_r, \widehat{\eta}_{ri})}{\mathbb{P}(z_i | \widehat{\phi}, \widehat{\eta}_i)} \right) (z_i - \mathbb{E}(Z | \widehat{\phi}_r, \widehat{\eta}_{ri})).$$

A função de log-verossimilhança aumentada do modelo, usada para obter os estimadores dos parâmetros, a menos de uma constante, é dada por

$$\ell^*(\phi, \boldsymbol{\eta} | \mathbf{Z}, \mathbf{X}) = \sum_{i=1}^n \sum_{r=1}^t g_{ri} \left[z_i \eta_{ri} - m \ln(1 + e^{\eta_{ri}}) + z_i(m - z_i) \phi_r - \right. \\ \left. \ln \left[\sum_{j=0}^m \binom{m}{j} \exp [j \eta_{ri} - m \ln(1 + e^{\eta_{ri}}) + j(m - j) \phi_r] \right] \right]$$

em que $G_{ri} = g_{ri} \in \{0, 1\}$, $r = 1, 2, \dots, t$, $i = 1, 2, \dots, n$, $1 - g_{0i} = g_{1i} + g_{2i} + \dots + g_{ti}$, são variáveis aleatórias (latentes) indicadoras de que a avaliação do i -ésimo item foi realizada com base no r -ésimo nível predefinido de preferência. Tomando a esperança $\mathbb{E}(G_{ri}) = \omega_{ri}$ obtém-se a função

$$Q(\phi, \boldsymbol{\eta} | \mathbf{Z}, \mathbf{X}) = \sum_{i=1}^n \sum_{r=1}^t \omega_{ri} \left[z_i \eta_{ri} - m \ln(1 + e^{\eta_{ri}}) + z_i(m - z_i) \phi_r - \right. \\ \left. \ln \left[\sum_{j=0}^m \binom{m}{j} \exp [j \eta_{ri} - m \ln(1 + e^{\eta_{ri}}) + j(m - j) \phi_r] \right] \right],$$

com $0 \leq \omega_{ri} \leq 1$ e $\omega_{0i} + \omega_{1i} + \dots + \omega_{ti} = 1$.

O parâmetro da mistura é estimado iterativamente tomando

$$\widehat{\omega}_{ri}^{(k+1)} = \frac{\widehat{\alpha}_r^{(k)} \mathbb{P}_r(z_i | \widehat{\phi}_r^{(k)}, \widehat{\eta}_{ri}^{(k)})}{\mathbb{P}(z_i | \widehat{\phi}^{(k)}, \widehat{\eta}_i^{(k)}), \quad \widehat{\alpha}_r^{(k+1)} = \frac{1}{n} \sum_{i=1}^n \widehat{\omega}_{ri}^{(k+1)},$$

o parâmetro ϕ_u , $u = 1, 2, \dots, t$, são estimados iterativamente por

$$\hat{\phi}_u^{(k+1)} = \hat{\phi}_u^{(k)} - \frac{Q'_{\phi_u}(\hat{\phi}^{(k)}, \hat{\eta}^{(k)} | \mathbf{Z}, \mathbf{X})}{Q''_{\phi_u}(\hat{\phi}^{(k)}, \hat{\eta}^{(k)} | \mathbf{Z}, \mathbf{X})},$$

com

$$\begin{aligned} Q'_{\phi_u}(\phi, \eta | \mathbf{Z}, \mathbf{X}) &= \sum_{i=1}^n \omega_{ui} [z_i(m - z_i) - \mathbb{E}(Z(m - Z) | \phi_u, \eta_{ui})], \\ Q''_{\phi_u}(\phi, \eta | \mathbf{Z}, \mathbf{X}) &= - \sum_{i=1}^n \omega_{ui} \mathbb{V}(Z(m - Z) | \phi_u, \eta_{ui}) \end{aligned}$$

e os parâmetros β_u , $u = 1, 2, \dots, t$, são estimados iterativamente por

$$\hat{\beta}_u^{(k+1)} = \hat{\beta}_u^{(k)} - Q'_{\beta_u}(\hat{\phi}^{(k)}, \hat{\eta}^{(k)} | \mathbf{Z}, \mathbf{X}) \left[Q''_{\beta_u}(\hat{\phi}^{(k)}, \hat{\eta}^{(k)} | \mathbf{Z}, \mathbf{X}) \right]^{-1},$$

com

$$\begin{aligned} Q'_{\beta_u}(\phi, \eta | \mathbf{Z}, \mathbf{X}) &= \sum_{i=1}^n \omega_{ui}(\mathbf{x}'_i) [z_i - \mathbb{E}(Z | \phi_u, \eta_{ui})], \\ Q''_{\beta_u}(\phi, \eta | \mathbf{Z}, \mathbf{X}) &= - \sum_{i=1}^n \omega_{ui}(\mathbf{x}_i \mathbf{x}'_i) \mathbb{V}(Z | \phi_u, \eta_{ui}). \end{aligned}$$

4 Inferência para os modelos

4.1 Qualidade de ajuste para os modelos

Um modelo de regressão, como os apresentados neste trabalho para avaliação de produtos e serviços, após ajustados devem ser avaliados quanto à qualidade deste ajuste. Na literatura estatística estão apresentadas diversas medidas de qualidade de ajuste, dentre as quais as avaliações de *pontos de alavanca*, *pontos de influência* e *pontos de influência local* são bastante comuns no cenário de modelos lineares generalizados. Estas medidas são apresentadas, para os modelos desenvolvidos neste trabalho, nas seções 4.1.1, 4.1.2 e 4.1.3.

O *ponto de alavanca* avalia o efeito de flutuações na variável resposta sobre o ajuste do modelo. Tal medida consiste em observar o quanto o modelo ajustado é perturbado por pequenas flutuações na resposta de cada um dos elementos amostrais. Os modelos apresentados neste capítulo apresentam respostas discretas, assim, o ponto de alavanca será definido como o efeito no ajuste do modelo de flutuações de 1 unidade para mais ou para menos na resposta de cada avaliador (o efeito que teria no ajuste do modelo se o avaliador tiver dado uma nota acima ou abaixo daquela realmente dada).

O *ponto de influência* avalia o peso que cada observação amostral (a resposta de cada avaliador) tem na função de verossimilhança do modelo. Uma avaliação que tenha um peso excessivo na função de verossimilhança pode afetar o ajuste do modelo, fazendo com que o modelo ajustado tenha a tendência de refletir a resposta desta observação amostral em particular.

Por fim, o *ponto de influência local* avalia o efeito de flutuações infinitesimais na função de verossimilhança do modelo, associada à cada observação amostral.

4.1.1 Ponto de alavanca

A avaliação do *ponto de alavanca* consiste em observar o efeito de flutuações na resposta Z_i sobre o respectivo valor ajustado \hat{Z}_i , $i = 1, 2, \dots, n$ em que

$$\hat{Z}_i = \mathbb{E}(Z|\boldsymbol{\theta}_i)$$

com $\boldsymbol{\theta}_i$ o vetor de parâmetros relativos a i -ésima observação.

O ponto de alavanca, considerando que a variável resposta é discreta, avalia o efeito na resposta ajustada para o i -ésimo elemento amostral, $i = 1, 2, \dots, n$, caso a resposta observada seja perturbada em 1 unidade para mais ou para menos (um deslocamento de tamanho $\pm s$ na nota). Para isso, considere as variáveis Z_i^+ e Z_i^- , $i = 1, 2, \dots, n$, de tal

forma que

$$Z_i^+ = \begin{cases} Z_i & Z_i \neq z_i \\ Z_i + \delta_i & Z_i = z_i \end{cases}, \quad Z_i^- = \begin{cases} Z_i & Z_i \neq z_i \\ Z_i - \delta_i & Z_i = z_i \end{cases},$$

com $\delta_i = s$ na i -ésima observação amostral, com $s > 0$ a distância entre as notas na escala, e $\delta_i = 0$ nas demais observações, de tal forma que as funções de probabilidade de Z_i^+ e Z_i^- são

$$\mathbb{P}^+(Z_i^+|\boldsymbol{\theta}_i) = \begin{cases} \mathbb{P}(Z_i|\boldsymbol{\theta}_i) & Z_i \neq z_i \\ \mathbb{P}(Z_i|\boldsymbol{\theta}_i) + \mathbb{P}(Z_i + \delta_i|\boldsymbol{\theta}_i) & Z_i = z_i \end{cases},$$

$$\mathbb{P}^-(Z_i^-|\boldsymbol{\theta}_i) = \begin{cases} \mathbb{P}(Z_i|\boldsymbol{\theta}_i) & Z_i \neq z_i \\ \mathbb{P}(Z_i - \delta_i|\boldsymbol{\theta}_i) + \mathbb{P}(Z_i|\boldsymbol{\theta}_i) & Z_i = z_i \end{cases}.$$

Observe que para os modelos do capítulo 3 a constante $s = 1$.

A medida de alavanca para uma variação positiva de tamanho s na nota para a i -ésima unidade amostral é

$$\nabla_i^+ = \mathbb{E}(Z^+|\boldsymbol{\theta}_i) - \mathbb{E}(Z|\boldsymbol{\theta}_i).$$

Desta forma, para $i = 1, 2, \dots, n$, e tomando os somatórios para todas as notas possíveis,

$$\begin{aligned} \nabla_i^+ &= \sum_z z\mathbb{P}^+(z|\boldsymbol{\theta}_i) - \sum_z z\mathbb{P}(z|\boldsymbol{\theta}_i) \\ &= (z_i + \delta_i)(\mathbb{P}(z_i + \delta_i|\boldsymbol{\theta}_i) + \mathbb{P}(z_i|\boldsymbol{\theta}_i)) - z_i\mathbb{P}(z_i|\boldsymbol{\theta}_i) \\ &= (z_i + \delta_i)\mathbb{P}(z_i + \delta_i|\boldsymbol{\theta}_i) + \delta_i\mathbb{P}(z_i|\boldsymbol{\theta}_i). \end{aligned}$$

De forma análoga, a medida de alavanca para uma variação negativa de tamanho s na nota para a i -ésima unidade amostral é

$$\nabla_i^- = \mathbb{E}(Z^-|\boldsymbol{\theta}_i) - \mathbb{E}(Z|\boldsymbol{\theta}_i).$$

Como antes, para $i = 1, 2, \dots, n$, e tomando os somatórios para todas as notas possíveis,

$$\begin{aligned} \nabla_i^- &= \sum_z z\mathbb{P}^-(z|\boldsymbol{\theta}_i) - \sum_z z\mathbb{P}(z|\boldsymbol{\theta}_i) \\ &= (z_i - \delta_i)(\mathbb{P}(z_i - \delta_i|\boldsymbol{\theta}_i) + \mathbb{P}(z_i|\boldsymbol{\theta}_i)) - z_i\mathbb{P}(z_i|\boldsymbol{\theta}_i) \\ &= (z_i - \delta_i)\mathbb{P}(z_i - \delta_i|\boldsymbol{\theta}_i) - \delta_i\mathbb{P}(z_i|\boldsymbol{\theta}_i). \end{aligned}$$

Combinando as duas medidas (positiva e negativa), a medida de alavanca pode ser definida por

$$\nabla_i = \max\left(\frac{|\nabla_i^+|}{z_i\mathbb{P}(z_i|\boldsymbol{\theta}_i)}, \frac{|\nabla_i^-|}{z_i\mathbb{P}(z_i|\boldsymbol{\theta}_i)}\right).$$

Modelo binomial redimensionado e deslocado

Considerando os modelos binomial redimensionado e deslocado sem inflação (seção 2.2) e com inflação (seção 2.3), a medida de alavanca para uma variação positiva de tamanho s na nota é dada por

$$\nabla_i^+ = (z_i + s)\mathbb{P}(z_i + s|\boldsymbol{\theta}_i)\mathbb{I}_{(z_i < ms)} + s\mathbb{P}(z_i|\boldsymbol{\theta}_i)$$

com $\mathbb{I}_{(z_i < ms)}$ a função indicadora que vale 1 se $z_i < ms$, $i = 1, 2, \dots, n$ ou 0 caso contrário.

A medida de alavanca para uma variação negativa de tamanho s na nota é dada por

$$\nabla_i^- = (z_i - s)\mathbb{P}(z_i - s|\boldsymbol{\theta}_i)\mathbb{I}_{(z_i > ds)} - s\mathbb{P}(z_i|\boldsymbol{\theta}_i)$$

com $\mathbb{I}_{(z_i > ds)}$ a função indicadora que vale 1 se $z_i > ds$, $i = 1, 2, \dots, n$ ou 0 caso contrário.

Modelo binomial duplo e modelo binomial multiplicativo

Considerando o modelo binomial duplo (seção 3.1) e o modelo binomial multiplicativo (seção 3.2), a medida de alavanca para uma variação positiva de tamanho 1 na nota é dada por

$$\nabla_i^+ = (z_i + 1)\mathbb{P}(z_i + 1|\boldsymbol{\theta}_i)\mathbb{I}_{(z_i < m)} + \mathbb{P}(z_i|\boldsymbol{\theta}_i)$$

com $\mathbb{I}_{(z_i < m)}$ a função indicadora que vale 1 se $z_i < m$, $i = 1, 2, \dots, n$ ou 0 caso contrário.

A medida de alavanca para uma variação negativa de tamanho 1 na nota é dada por

$$\nabla_i^- = (z_i - 1)\mathbb{P}(z_i - 1|\boldsymbol{\theta}_i)\mathbb{I}_{(z_i > 0)} - \mathbb{P}(z_i|\boldsymbol{\theta}_i)$$

com $\mathbb{I}_{(z_i > 0)}$ a função indicadora que vale 1 se $z_i > 0$, $i = 1, 2, \dots, n$ ou 0 caso contrário.

4.1.2 Influência

A influência de uma observação amostral no modelo é avaliada através do impacto da remoção desta observação na função de log-verossimilhança do modelo. Assim, a avaliação do peso de cada observação amostral, $i = 1, 2, \dots, n$, é dada por

$$LD_i = 2\{\ell(\hat{\boldsymbol{\beta}}) - \ell(\hat{\boldsymbol{\beta}}_{(i)})\}$$

em que $\ell(\cdot)$ é a *função de log-verossimilhança* do modelo, $\hat{\boldsymbol{\beta}}' = [\hat{\boldsymbol{\beta}}'_1, \hat{\boldsymbol{\beta}}'_2, \dots, \hat{\boldsymbol{\beta}}'_t]'$ é o vetor conjunto (com as estimativas para cada componente do modelo) de estimativas dos parâmetros de regressão do modelo e $\hat{\boldsymbol{\beta}}'_{(i)} = [\hat{\boldsymbol{\beta}}'_{1;(i)}, \hat{\boldsymbol{\beta}}'_{2;(i)}, \dots, \hat{\boldsymbol{\beta}}'_{t;(i)}]'$ é o vetor de estimativas dos parâmetros de regressão do modelo excluindo a i -ésima observação amostral. É possível

escrever esta expressão em função da segunda aproximação de Taylor para a verossimilhança em torno de $\widehat{\boldsymbol{\beta}}$, dado por

$$LD_i \approx (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})' \left\{ -\ddot{L}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) \right\} (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})$$

em que $\ddot{L}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})$ é uma matriz $tp \times tp$ dada por

$$\ddot{L}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) = \begin{bmatrix} \ddot{L}_{\boldsymbol{\beta}_{(11)}}(\widehat{\boldsymbol{\beta}}) & \ddot{L}_{\boldsymbol{\beta}_{(12)}}(\widehat{\boldsymbol{\beta}}) & \cdots & \ddot{L}_{\boldsymbol{\beta}_{(1t)}}(\widehat{\boldsymbol{\beta}}) \\ \ddot{L}_{\boldsymbol{\beta}_{(21)}}(\widehat{\boldsymbol{\beta}}) & \ddot{L}_{\boldsymbol{\beta}_{(22)}}(\widehat{\boldsymbol{\beta}}) & \cdots & \ddot{L}_{\boldsymbol{\beta}_{(2t)}}(\widehat{\boldsymbol{\beta}}) \\ \vdots & \vdots & \ddots & \vdots \\ \ddot{L}_{\boldsymbol{\beta}_{(t1)}}(\widehat{\boldsymbol{\beta}}) & \ddot{L}_{\boldsymbol{\beta}_{(t2)}}(\widehat{\boldsymbol{\beta}}) & \cdots & \ddot{L}_{\boldsymbol{\beta}_{(tt)}}(\widehat{\boldsymbol{\beta}}) \end{bmatrix}$$

com

$$\ddot{L}_{\boldsymbol{\beta}_{(uv)}}(\widehat{\boldsymbol{\beta}}) = \left. \frac{\partial}{\partial \boldsymbol{\beta}'_u} \frac{\partial}{\partial \boldsymbol{\beta}_v} \ell(\boldsymbol{\beta}) \right|_{\boldsymbol{\beta}_u = \widehat{\boldsymbol{\beta}}_u, \boldsymbol{\beta}_v = \widehat{\boldsymbol{\beta}}_v}$$

e, tomando, $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}_{(i)}$ obtém-se a expressão da medida de influência

$$LD_i \approx (\widehat{\boldsymbol{\beta}}_{(i)} - \widehat{\boldsymbol{\beta}})' \left\{ -\ddot{L}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) \right\} (\widehat{\boldsymbol{\beta}}_{(i)} - \widehat{\boldsymbol{\beta}}).$$

Os vetores $\widehat{\boldsymbol{\beta}}_{(i)}$, $i = 1, 2, \dots, n$, podem ser obtidos estimando n modelos, um para cada observação amostral, ou, de acordo com [Pregibon \(1981\)](#), pode ser aproximado pelo vetor

$$\widehat{\boldsymbol{\beta}}_{(i)} \approx \widehat{\boldsymbol{\beta}} + \left\{ -\ddot{L}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}}) \right\}^{-1} \ell_{(i)}(\widehat{\boldsymbol{\beta}})$$

com

$$\ell_{(i)}(\widehat{\boldsymbol{\beta}}) = \sum_{i=1}^n \delta_i \ell_i(\widehat{\boldsymbol{\beta}})$$

a componente da função de log-verossimilhança associada à i -ésima observação amostral e $\delta_i = 0$ na i -ésima observação e $\delta_i = 1$ caso contrário.

Modelo binomial redimensionado e deslocado sem inflação

No modelo binomial redimensionado e deslocado sem inflação (seção 2.2) a matriz $\ddot{L}_{\boldsymbol{\beta}}(\widehat{\boldsymbol{\beta}})$ é tal que o elemento na posição (u, v) é dado por

$$\begin{aligned} \ddot{L}_{\boldsymbol{\beta}_{(u,v)}}(\widehat{\boldsymbol{\beta}}) = & - \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \left[\prod_{r \in \{u,v\}} \left(\frac{\alpha_r \mathbb{P}_r(z_i | \boldsymbol{\eta}_{ri})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left[\left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} \right) \right] \right. \\ & \left. \mathbb{I}_{(u=v)} \left(\frac{\alpha_u \mathbb{P}_u(z_i | \boldsymbol{\eta}_{ui})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left[\left(\left(\frac{z_i}{s} - d \right) - \frac{(m-d)e^{\eta_{ui}}}{1 + e^{\eta_{ui}}} \right)^2 - \frac{(m-d)e^{\eta_{ui}}}{(1 + e^{\eta_{ui}})^2} \right] \right]. \end{aligned} \quad (4.1)$$

Modelo binomial redimensionado e deslocado com inflação

No modelo binomial redimensionado e deslocado com inflação (seção 2.3) a matriz $\ddot{L}_\beta(\hat{\beta})$ é tal que o elemento na posição (u, v) é dado por

$$\begin{aligned} \ddot{L}_{\beta(u,v)}(\hat{\beta}) = & -\sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \left[\prod_{r \in \{u,v\}} \left(\frac{\tau_r \mathbb{P}_r(z_i | \eta_{ri})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left[\left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ri}}}{1+e^{\eta_{ri}}} \right) \right. \right. \\ & \left. \left. \mathbb{I}_{(u=v)} \left(\frac{\tau_u \mathbb{P}_u(z_i | \eta_{ui})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left[\left(\left(\frac{z_i}{s} - d \right) - \frac{(m-d)e^{\eta_{ui}}}{1+e^{\eta_{ui}}} \right)^2 - \frac{(m-d)e^{\eta_{ui}}}{(1+e^{\eta_{ui}})^2} \right] \right] \right]. \end{aligned} \quad (4.2)$$

Modelo binomial duplo

No modelo binomial duplo (seção 3.1) a matriz $\ddot{L}_\beta(\hat{\beta})$ é tal que o elemento na posição (u, v) é dado por

$$\begin{aligned} \ddot{L}_{\beta(u,v)}(\hat{\beta}) = & -\sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \left[\prod_{r \in \{u,v\}} \theta_r \left(\frac{\alpha_r \mathbb{P}_r(z_i | \phi_r, \eta_{ri})}{\mathbb{P}(z_i | \boldsymbol{\phi}, \boldsymbol{\eta}_i)} \right) \left(z_i - \frac{me^{\eta_{ri}}}{1+e^{\eta_{ri}}} \right) \right. \\ & \left. - \mathbb{I}_{(u=v)} \theta_u \left(\frac{\alpha_u \mathbb{P}_u(z_i | \phi_u, \eta_{ui})}{\mathbb{P}(z_i | \boldsymbol{\phi}, \boldsymbol{\eta}_i)} \right) \left[\theta_u \left(z_i - \frac{me^{\eta_{ui}}}{1+e^{\eta_{ui}}} \right)^2 - \frac{me^{\eta_{ui}}}{(1+e^{\eta_{ui}})^2} \right] \right]. \end{aligned} \quad (4.3)$$

Modelo binomial multiplicativo

No modelo binomial multiplicativo (seção 3.2) a matriz $\ddot{L}_\beta(\hat{\beta})$ é tal que o elemento na posição (u, v) é dado por

$$\begin{aligned} \ddot{L}_{\beta(u,v)}(\hat{\beta}) = & -\sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \left[\prod_{r \in \{u,v\}} \left(\frac{\alpha_r \mathbb{P}_r(z_i | \phi_r, \eta_{ri})}{\mathbb{P}(z_i | \boldsymbol{\phi}, \boldsymbol{\eta}_i)} \right) [z_i - \mathbb{E}(Z | \phi_r, \boldsymbol{\eta}_i)] \right. \\ & - \left(\frac{\alpha_u \mathbb{P}_u(z_i | \phi_u, \eta_{ui})}{\mathbb{P}(z_i | \boldsymbol{\phi}, \boldsymbol{\eta}_i)} \right) \times \\ & \left. \left[\mathbb{V}(Z | \phi_u, \eta_{ui}) - 2\mathbb{E}(Z | \phi_u, \eta_{ui}) \left(z_i - \frac{me^{\eta_{ui}}}{1+e^{\eta_{ui}}} \right) + z_i^2 - \left(\frac{me^{\eta_{ui}}}{(1+e^{\eta_{ui}})} \right)^2 \right] \right]. \end{aligned} \quad (4.4)$$

4.1.3 Influência local

A influência local avalia o efeito de perturbações infinitesimais na componente da função de log-verossimilhança associada à i -ésima, $i = 1, 2, \dots, n$, observação amostral. A influência local avalia a função

$$LD(\boldsymbol{\delta}) = 2\{\ell(\boldsymbol{\beta}) - \ell(\boldsymbol{\beta} | \boldsymbol{\delta})\}$$

em torno do vetor $\boldsymbol{\delta}_0$ definido de tal forma que $LD(\boldsymbol{\delta}_0) = 0$ (ou seja, $\ell(\boldsymbol{\beta} | \boldsymbol{\delta}_0) = \ell(\boldsymbol{\beta})$). Selecionando um vetor unitário $\boldsymbol{\zeta}$, define-se o gráfico de $LD(\boldsymbol{\delta}_0 + a\boldsymbol{\zeta})$ em função de

$a \in \mathbb{R}$ chamado de *linha projetada*. Cada linha projetada pode ser caracterizada por uma *curvatura normal* $C_\zeta(\beta)$ em torno de $a = 0$. Cook (1986) mostra que esta curvatura assume a forma

$$C_\zeta(\beta) = 2|\zeta' \Delta' \ddot{L}_\beta^{-1}(\beta) \Delta \zeta|$$

em que $\ddot{L}_\beta(\hat{\beta})$ é a matriz de informação observada de Fisher, Δ é uma matriz $p \times n$ cujos elementos são dados por $\Delta_{jk} = \partial^2 L(\beta|\delta) / \partial \beta_j \partial \delta_k$, $j = 1, 2, \dots, p$, $k = 1, 2, \dots, n$. A máxima curvatura normal é obtida tomando ζ como o maior autovalor (em valor absoluto) da matriz $n \times n$, $\zeta' \Delta' \ddot{L}_{\beta' \beta}^{-1}(\hat{\beta}) \Delta \zeta$. O gráfico de $|\zeta_{\max}|$, o valor absoluto dos componentes do vetor ζ na direção de máxima curvatura normal, contra a ordem (identificação) das observações, permite identificar os pontos que causam grande influência local na vizinhança de δ_0 .

Tomando a log-verossimilhança escrita em função do vetor δ dos pesos associados à contribuição de cada observação amostral para a log-verossimilhança,

$$\ell(\beta|\delta) = \sum_{i=1}^n \delta_i \ell_i(\beta_r)$$

com β_r , $r = 1, 2, \dots, t$, o vetor de parâmetros associados à r -ésima componente do modelo, $\delta = \{\delta_1, \delta_2, \dots, \delta_n\}$ o vetor de pesos (latentes) associados à cada observação amostral e C é uma constante que independe de β_r . A influência local pode ser obtida para cada componente da mistura calculando uma matriz $\Delta^{(r)}$, $r = 1, 2, \dots, t$, de dimensão $p \times n$, cujos elementos são dados por

$$\Delta_{jk}^{(r)} = \frac{\partial}{\partial \beta_{rj}} \frac{\partial}{\partial \delta_k} \sum_{i=1}^n \delta_i \ell_i(\beta_r)$$

em que β_{rj} é o parâmetro associado à j -ésima covariável, $j = 1, 2, \dots, p$, do r -ésimo componente, $r = 1, 2, \dots, t$, do modelo. Fixada uma direção (um vetor) ζ , a curvatura normal, associada ao r -ésimo componente binomial do modelo, nesta direção é dada por

$$C_\zeta^{(r)}(\beta) = 2|\zeta' \Delta^{(r)' \ddot{L}_{\beta_r}(\hat{\beta}) \Delta^{(r)} \zeta|.$$

A direção de máxima curvatura é o vetor ζ_{\max} associado ao maior autovalor (em valor absoluto) de $\Delta^{(r)' \ddot{L}_{\beta_r \beta_r}(\hat{\beta}) \Delta^{(r)}$.

Modelo binomial redimensionado e deslocado sem inflação

A curvatura normal $C_\zeta^{(r)}(\beta)$ associada à r -ésima, $r = 1, 2, \dots, t$ componente do modelo binomial redimensionado e deslocado sem inflação (seção 2.2) é dada por

$$C_\zeta^{(r)}(\beta) = 2|\zeta' \Delta^{(r)' \ddot{L}_{\beta_r}(\hat{\beta}) \Delta^{(r)} \zeta|.$$

em que $\ddot{L}_{\beta_r}(\hat{\beta})$ é dada por (4.1) e $\Delta^{(r)}$ é uma matriz de dimensão $p \times n$ cujos componentes são

$$\Delta_{kj}^{(r)} = x_{jk} \left(\frac{\alpha_r \mathbb{P}_r(z_j | \eta_{rj})}{\mathbb{P}(z_j | \boldsymbol{\eta}_j)} \right) \left[\left(\frac{z_j}{s} - d \right) - \frac{(m-d)e^{\eta_{rj}}}{1 + e^{\eta_{rj}}} \right].$$

Modelo binomial redimensionado e deslocado com inflação

A curvatura normal $C_{\zeta}^{(r)}(\beta)$ associada à r -ésima, $r = 1, 2, \dots, t$ componente do modelo binomial redimensionado e deslocado com inflação (seção 2.3) é dada por

$$C_{\zeta}^{(r)}(\beta) = 2|\zeta' \Delta^{(r)'} \ddot{L}_{\beta_r}(\hat{\beta}) \Delta^{(r)} \zeta|.$$

em que $\ddot{L}_{\beta_r}(\hat{\beta})$ é dada por (4.2) e $\Delta^{(r)}$ é uma matriz de dimensão $p \times n$ cujos componentes são

$$\Delta_{kj}^{(r)} = x_{jk} \left(\frac{\tau_r \mathbb{P}_r(z_j | \eta_{rj})}{\mathbb{P}(z_j | \boldsymbol{\eta}_j)} \right) \left[\left(\frac{z_j}{s} - d \right) - \frac{(m-d)e^{\eta_{rj}}}{1+e^{\eta_{rj}}} \right].$$

Modelo binomial duplo

A curvatura normal $C_{\zeta}^{(r)}(\beta)$ associada à r -ésima, $r = 1, 2, \dots, t$ componente do modelo binomial duplo (seção 3.1) é dada por

$$C_{\zeta}^{(r)}(\beta) = 2|\zeta' \Delta^{(r)'} \ddot{L}_{\beta_r}(\hat{\beta}) \Delta^{(r)} \zeta|.$$

em que $\ddot{L}_{\beta_r}(\hat{\beta})$ é dada por (4.3) e $\Delta^{(r)}$ é uma matriz de dimensão $p \times n$ cujos componentes são

$$\Delta_{kj}^{(r)} = x_{jk} \theta_r \left(\frac{\alpha_r \mathbb{P}_r(z_j | \phi_r, \eta_{rj})}{\mathbb{P}(z_j | \phi_r, \boldsymbol{\eta}_j)} \right) \left(z_j - \frac{me^{\eta_r}}{1+e^{\eta_{rj}}} \right).$$

Modelo binomial multiplicativo

A curvatura normal $C_{\zeta}^{(r)}(\beta)$ associada à r -ésima, $r = 1, 2, \dots, t$ componente do modelo binomial multiplicativo (seção 3.2) é dada por

$$C_{\zeta}^{(r)}(\beta) = 2|\zeta' \Delta^{(r)'} \ddot{L}_{\beta_r}(\hat{\beta}) \Delta^{(r)} \zeta|.$$

em que $\ddot{L}_{\beta_r}(\hat{\beta})$ é dada por (4.4) e $\Delta^{(r)}$ é uma matriz de dimensão $p \times n$ cujos componentes são

$$\Delta_{kj}^{(r)} = x_{jk} \left(\frac{\alpha_r \mathbb{P}_r(z_j | \phi_r, \eta_{rj})}{\mathbb{P}(z_j | \phi_r, \boldsymbol{\eta}_j)} \right) [z_j - \mathbb{E}(Z | \phi_r, \eta_{rj})].$$

4.2 Testes de hipóteses

Nesta seção, serão descritos alguns métodos para a realização de testes de hipóteses para os modelos descritos neste capítulo: os testes de *razão de verossimilhanças*, de Wald, escore e gradiente.

Considere que o vetor de parâmetros $\boldsymbol{\theta}$ possa ser escrito como $\boldsymbol{\theta} = [\boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2]'$, de tal modo que $\dim \boldsymbol{\theta} = p = \dim \boldsymbol{\theta}_1 + \dim \boldsymbol{\theta}_2$ e $\dim \boldsymbol{\theta}_2 = q$, e que as hipóteses do teste podem ser escritas como $H_0 : \boldsymbol{\theta}_2 = \boldsymbol{\theta}_0$ e $H_1 : \boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_0$. Neste cenário, serão definidos abaixo o *teste de razão de verossimilhanças* (Wilks, 1938), o *teste de Wald* (Wald, 1943), o *teste escore* (Rao, 1948) e o *teste gradiente* (Terrell, 2002).

4.2.1 Teste de razão de verossimilhanças

O teste de razão de verossimilhança foi proposto inicialmente por Wilks (1938) e é utilizado para o teste de modelos encaixados – um modelo cujos parâmetros assume valores em um subconjunto do espaço paramétrico original. Suponha que o modelo é descrito pela função de verossimilhança $L(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X})$ em que \mathbf{Z} é um vetor de respostas, \mathbf{X} é uma matriz de covariáveis e $\boldsymbol{\theta}$ é um vetor de parâmetros. Seja Ω o espaço dos valores possíveis para o vetor de parâmetros $\boldsymbol{\theta}$. Se ω for um subespaço de Ω , ou seja, $\omega \subset \Omega$ com $\dim \omega < \dim \Omega$, de modo que Ω é o espaço associado ao modelo não restrito e ω é o espaço paramétrico associado ao modelo restrito (pela hipótese nula – $H_0 : \boldsymbol{\theta}_2 = \boldsymbol{\theta}_0$). A estatística do teste é dada por

$$\lambda = \frac{\max_{\boldsymbol{\theta}} L_{\omega}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X})}{\max_{\boldsymbol{\theta}} L_{\Omega}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X})} \quad (4.5)$$

em que $L_{\omega}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X})$ é a verossimilhança do modelo no espaço ω restrito pela hipótese nula e $L_{\Omega}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X})$ é a verossimilhança do modelo no espaço Ω irrestrito, de modo que $-2 \log \lambda$ segue uma distribuição qui-quadrado com os graus de liberdade dados por $\dim \Omega - \dim \omega$.

Existência de avaliações ao acaso

O teste para a hipótese de que não existem avaliadores que fazem as avaliações ao acaso, no modelo binomial redimensionado e deslocado sem inflação (seção 2.2), consiste em testar $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$. Sob a hipótese nula, o vetor de parâmetros $\boldsymbol{\theta} \subset \omega$ em que

$$\omega = \{\boldsymbol{\theta} \in \Omega : 0 \leq \alpha_u^* \leq 1, \alpha_0^* = 0, \alpha_1^* + \dots + \alpha_t^* = 1, 0 \leq u \in \mathbb{N} \leq t\}$$

e sob o modelo irrestrito $\boldsymbol{\theta} \subset \Omega$ em que

$$\Omega = \{\boldsymbol{\theta} : 0 \leq \alpha_u \leq 1, \alpha_0 + \dots + \alpha_t = 1, 0 \leq u \in \mathbb{N} \leq t, \boldsymbol{\beta}_v \in \mathbb{R}_p, 1 \leq v \in \mathbb{N} \leq t\}.$$

Sejam

$$L_{\Omega}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z_i|\eta_{ri}) + \alpha_0 \mathbb{P}_0(z_i) \right]$$

a função de verossimilhança no modelo irrestrito e

$$L_{\omega}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \sum_{r=1}^t \alpha_r^* \mathbb{P}_r(z_i|\eta_{ri}^*) \quad (4.6)$$

a função de verossimilhança no modelo restrito por $H_0 : \alpha_0 = 0$. A estatística do teste de razão de verossimilhança para as hipóteses $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$ é dada pela equação (4.5), com as respectivas funções de verossimilhança, e segue uma distribuição qui-quadrado com 1 grau de liberdade.

No modelo binomial redimensionado e deslocado com inflação (seção 2.3), a inexistência de avaliações ao acaso implica na inexistência de avaliações com valor fixo (uma vez que tal avaliação é também um tipo de avaliação ao acaso). Desta forma, para testar a hipótese $H_0 : \tau_0 = 0$ contra $H_1 : \tau_0 \neq 0$ é necessário considerar a restrição $\tau_1 + \tau_2 + \dots + \tau_t = 1$ para eliminar da verossimilhança

restrita o termo de inflação. Desta forma, a hipótese nula ($H_0 : \tau_0 = 0, \tau_1 + \tau_2 + \dots + \tau_t = 1$) é representada pelo vetor $\boldsymbol{\theta} \subset \boldsymbol{\omega}$ de tal forma que

$$\boldsymbol{\omega} = \{\boldsymbol{\theta} \in \boldsymbol{\Omega} : 0 \leq \tau_u^* \leq 1, \tau_0^* = 0, \tau_1^* + \dots + \tau_t^* = 1, 0 \leq u \in \mathbb{N} \leq t\}$$

e o modelo irrestrito é tal que $\boldsymbol{\theta} \in \boldsymbol{\Omega}$ com

$$\boldsymbol{\Omega} = \{\boldsymbol{\theta} : 0 \leq \tau_u \leq 1, \tau_0 + \dots + \tau_t \leq 1, 0 \leq u \in \mathbb{N} \leq t, \boldsymbol{\beta}_v \in \mathbb{R}_v, 1 \leq v \in \mathbb{N} \leq t\}.$$

Sejam

$$L_{\boldsymbol{\Omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\tau_0 \mathbb{P}_0(z_i) + \sum_{r=1}^t \tau_r \mathbb{P}_r(z_i|\eta_{ri}) + \left(1 - \sum_{r=0}^t \tau_r\right) \mathbb{I}_{z=c} \right]$$

a função de verossimilhança no modelo irrestrito, e

$$L_{\boldsymbol{\omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \sum_{r=1}^t \tau_r^* \mathbb{P}_r(z_i|\eta_{ri}^*) \quad (4.7)$$

a função de verossimilhança no modelo restrito por $H_0 : \tau_0 = 0$. A estatística do teste de razão de verossimilhança para as hipóteses $H_0 : \tau_0 = 0$ contra $H_1 : \tau_0 \neq 0$ é dada pela equação (4.5) e segue uma distribuição qui-quadrado com 2 graus de liberdade (1 grau de liberdade referente à restrição em τ_0 e 1 grau de liberdade referente à restrição na soma $\tau_1 + \dots + \tau_t = 1$).

O teste para a hipótese de que não existem avaliadores que fazem as avaliações ao acaso, no modelo binomial duplo (seção 3.1) ou no modelo binomial multiplicativo (seção 3.2), consiste em testar $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$. Sob a hipótese nula, o vetor de parâmetros $\boldsymbol{\theta} \subset \boldsymbol{\omega}$ em que

$$\boldsymbol{\omega} = \{\boldsymbol{\theta} \in \boldsymbol{\Omega} : 0 \leq \alpha_u^* \leq 1, \alpha_0^* = 0, \alpha_1^* + \dots + \alpha_t^* = 1, \phi_u^* > 0, 0 \leq u \in \mathbb{N} \leq t\}$$

e sob o modelo irrestrito $\boldsymbol{\theta} \subset \boldsymbol{\Omega}$ em que

$$\boldsymbol{\Omega} = \{\boldsymbol{\theta} : 0 \leq \alpha_u \leq 1, \alpha_0 + \dots + \alpha_t = 1, \phi_u > 0, 0 \leq u \in \mathbb{N} \leq t, \boldsymbol{\beta}_v \in \mathbb{R}_p, 1 \leq v \in \mathbb{N} \leq t\}.$$

Sejam $\mathbb{P}_r(z_i|\phi_r, \eta_{ri})$, $r = 1, 2, \dots, t$, a função de probabilidade do modelo no r -ésimo componente binomial do modelo,

$$L_{\boldsymbol{\Omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z_i|\phi_r, \eta_{ri}) + \alpha_0 \mathbb{P}_0(z_i) \right]$$

a função de verossimilhança no modelo irrestrito e

$$L_{\boldsymbol{\omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \sum_{r=1}^t \alpha_r^* \mathbb{P}_r(z_i|\phi_r^*, \eta_{ri}^*) \quad (4.8)$$

a função de verossimilhança no modelo restrito por $H_0 : \alpha_0 = 0$. A estatística do teste de razão de verossimilhança para as hipóteses $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$ é dada pela equação (4.5), com as respectivas funções de verossimilhança, e segue uma distribuição qui-quadrado com 1 grau de liberdade.

Existência de inflação na avaliação ao acaso

No modelo binomial redimensionado e deslocado com inflação (seção 2.3), o teste para a existência de inflação nas avaliações ao acaso é realizado pela verificação de que os parâmetros $\tau_0 + \tau_1 + \dots + \tau_t = 1$. Desta forma, a hipótese nula ($H_0 : \tau_0 + \tau_1 + \dots + \tau_t = 1$) é representada pelo vetor $\boldsymbol{\theta} \subset \boldsymbol{\omega}$ de tal forma que

$$\boldsymbol{\omega} = \{\boldsymbol{\theta} \in \boldsymbol{\Omega} : 0 \leq \tau_u^* \leq 1, \tau_0^* + \tau_1^* + \dots + \tau_t^* = 1, 0 \leq u \in \mathbb{N} \leq t\}$$

e o modelo irrestrito é representado pelo vetor $\boldsymbol{\theta} \subset \boldsymbol{\Omega}$ tal que

$$\boldsymbol{\Omega} = \{\boldsymbol{\theta} : 0 \leq \tau_u \leq 1, \tau_0 + \dots + \tau_t \leq 1, 0 \leq u \in \mathbb{N} \leq t, \boldsymbol{\beta}_v \in \mathbb{R}_v, 1 \leq v \in \mathbb{N} \leq t\}.$$

Sejam $\boldsymbol{\theta} = \{\tau_0, \tau_1, \dots, \tau_t, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_t\}$ os parâmetros do modelo inflacionado no espaço paramétrico $\boldsymbol{\Omega}$ (irrestrito) e $\boldsymbol{\theta}_0 = \{\tau_1^*, \dots, \tau_t^*, \boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*, \dots, \boldsymbol{\beta}_t^*\}$ os parâmetros do modelo inflacionado no espaço paramétrico $\boldsymbol{\omega}$ (restrito pela hipótese nula $H_0 : \alpha_0 = 0$). Além disto, sejam

$$L_{\boldsymbol{\Omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\tau_0 \mathbb{P}_0(z_i) + \sum_{r=1}^t \tau_r \mathbb{P}_r(z_i|\eta_{ri}) + \left(1 - \sum_{r=0}^t \tau_r\right) \mathbb{I}_{z=c} \right]$$

a função de verossimilhança no modelo irrestrito e

$$L_{\boldsymbol{\omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\tau_0^* \mathbb{P}_0(z_i) + \sum_{r=1}^t \tau_r^* \mathbb{P}_r(z_i|\eta_{ri}^*) \right] \quad (4.9)$$

a função de verossimilhança no modelo restrito. A estatística do teste de razão de verossimilhança é dada pela equação (4.5) e segue uma distribuição qui-quadrado com 1 grau de liberdade.

Testes para os parâmetros de regressão

Considere que $t^* \in \mathbb{N} \leq t$ de modo que, para cada componente do modelo com índice $t^* < v \leq t$, sejam definidas restrições do tipo $\beta_{i_{v,1}} = \beta_{i_{v,2}} = \dots = \beta_{i_{v,k_v}} = 0$, $1 \leq k_v < p$. Sejam $\boldsymbol{\beta}_v^*$, os vetores de parâmetros de regressão associados ao v -ésimo componente do modelo, de tal modo que os vetores $\boldsymbol{\beta}_v^*$ sejam restritos como descrito anteriormente. Desta forma, a hipótese nula (H_0) no modelo redimensionado e deslocado não inflacionado (seção 2.2) é representada pelo vetor $\boldsymbol{\theta} \subset \boldsymbol{\omega}$ de tal forma que

$$\boldsymbol{\omega} = \{\boldsymbol{\theta} \in \boldsymbol{\Omega} : \beta_{v,j}^* = \beta_{v,j} \cdot \mathbb{I}_{(j \notin \{i_{v,1}, i_{v,2}, \dots, i_{v,k_v}\})}, 1 \leq k_v < p, t^* < v \leq t\}$$

e o modelo irrestrito não inflacionado é tal que $\boldsymbol{\theta} \subset \boldsymbol{\Omega}$ com

$$\boldsymbol{\Omega} = \{\boldsymbol{\theta} : 0 \leq \alpha_u \leq 1, \alpha_0 + \dots + \alpha_t = 1, \boldsymbol{\beta}_u \in \mathbb{R}_p, 0 \leq u \in \mathbb{N} \leq t\}.$$

Sejam

$$L_{\boldsymbol{\Omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z_i|\eta_{ri}) + \alpha_0 \mathbb{P}_0(z_i) \right]$$

a função de verossimilhança no modelo irrestrito e

$$L_{\boldsymbol{\omega}}(\boldsymbol{\theta}|\mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r^* \mathbb{P}_r(z_i|\eta_{ri}^*) + \alpha_0^* \mathbb{P}_0(z_i) \right] \quad (4.10)$$

a função de verossimilhança no modelo restrito. A estatística do teste de razão de verossimilhança é dada pela equação (4.5), com as respectivas funções de verossimilhança, e segue uma distribuição qui-quadrado com o número de grau de liberdade dado por

$$gl = \sum_{v=t^*+1}^t k_v.$$

No modelo binomial redimensionado e deslocado com inflação (seção 2.3) o teste é semelhante, tomando $\theta \subset \omega$ com

$$\omega = \{\theta \in \Omega : \beta_{v,j}^* = \beta_{v,j} \cdot \mathbb{I}_{(j \notin \{i_{v,1}, i_{v,2}, \dots, i_{v,k_v}\})}, 1 \leq k_v < p, t^* < v \leq t\}$$

no modelo restrito e, no modelo restrito, tomando $\theta \subset \Omega$ com

$$\Omega = \{\theta : 0 \leq \tau_u \leq 1, \tau_0 + \dots + \tau_t \leq 1, \beta_u \in \mathbb{R}_p, 0 \leq u \in \mathbb{N} \leq t\},$$

a função de verossimilhança para o modelo irrestrito dada por

$$L_{\Omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\tau_0 \mathbb{P}_0(z_i) + \sum_{r=1}^t \tau_r \mathbb{P}_r(z_i | \eta_{ri}) + \left(1 - \sum_{r=0}^t \tau_r\right) \mathbb{I}_{z_i=c} \right]$$

e a função de verossimilhança no modelo restrito dada por

$$L_{\Omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\tau_0^* \mathbb{P}_0(z_i) + \sum_{r=1}^t \tau_r^* \mathbb{P}_r(z_i | \eta_{ri}^*) + \left(1 - \sum_{r=0}^t \tau_r^*\right) \mathbb{I}_{z_i=c} \right] \quad (4.11)$$

a função de verossimilhança no modelo restrito. A estatística do teste de razão de verossimilhança é dada pela equação (4.5), com as respectivas funções de verossimilhança, e segue uma distribuição qui-quadrado com o número de grau de liberdade dado por

$$gl = \sum_{v=t^*+1}^t k_v.$$

A hipótese nula (H_0) no modelo binomial duplo (seção 3.1) ou no modelo binomial multiplicativo (seção 3.2) é representada pelo vetor $\theta \subset \omega$ de tal forma que

$$\omega = \{\theta \in \Omega : \beta_{v,j}^* = \beta_{v,j} \cdot \mathbb{I}_{(j \notin \{i_{v,1}, i_{v,2}, \dots, i_{v,k_v}\})}, 1 \leq k_v < p, t^* < v \leq t\}$$

e o modelo irrestrito não inflacionado é tal que $\theta \subset \Omega$ com

$$\Omega = \{\theta : 0 \leq \alpha_u \leq 1, \alpha_0 + \dots + \alpha_t = 1, \beta_u \in \mathbb{R}_p, 0 \leq u \in \mathbb{N} \leq t\}.$$

Sejam $\mathbb{P}_r(z_i | \phi_r, \eta_{ri})$ a função de probabilidade do r -ésimo, $r = 1, 2, \dots, t$, componente do modelo e

$$L_{\Omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z_i | \phi_r, \eta_{ri}) + \alpha_0 \mathbb{P}_0(z_i) \right]$$

a função de verossimilhança no modelo irrestrito e

$$L_{\omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r^* \mathbb{P}_r(z_i | \phi_r^*, \eta_{ri}^*) + \alpha_0^* \mathbb{P}_0(z_i) \right] \quad (4.12)$$

a função de verossimilhança no modelo restrito. A estatística do teste de razão de verossimilhança é dada pela equação (4.5), com as respectivas funções de verossimilhança, e segue uma distribuição qui-quadrado com o número de grau de liberdade dado por

$$gl = \sum_{v=t^*+1}^t k_v.$$

Homogeneidade entre avaliadores (na avaliação por preferência)

O teste para verificar se os avaliadores que fazem as avaliações de acordo com níveis de preferência são homogêneos é realizado verificando se $\alpha_1 = \alpha_2 = \dots = \alpha_t = \alpha^*/t$ (e, por consequência, $\alpha_0 = 1 - \alpha^*$). No entanto, esta restrição implica que $\beta_1 = \beta_2 = \dots = \beta_t = \beta^*$. Desta forma, a hipótese nula, no modelo binomial redimensionado e deslocado sem inflação (seção 2.2), é representada pelo vetor $\theta \subset \omega$ de tal forma que

$$\omega = \{\theta \in \Omega : \alpha_u^* = \alpha^*/t, 1 \leq u \in \mathbb{N} \leq t; \beta_u^* \in \mathbb{R}^p, 1 \leq u \in \mathbb{N} \leq t\}$$

e o modelo irrestrito é tal que $\theta \subset \Omega$, com

$$\Omega = \{\theta : 0 \leq \alpha_u \leq 1, \alpha_0 + \dots + \alpha_t = 1, \beta_u \in \mathbb{R}, 0 \leq u \in \mathbb{N} \leq t\}.$$

Sejam

$$L_{\Omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z_i | \eta_{ri}) + \alpha_0 \mathbb{P}_0(z_i) \right]$$

a função de verossimilhança no modelo irrestrito, e

$$L_{\omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\alpha^* \mathbb{P}_1(z_i | \eta_i^*) + (1 - \alpha^*) \mathbb{P}_0(z_i) \right] \quad (4.13)$$

a função de verossimilhança no modelo restrito por $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_t = \alpha^*/t, \beta_1 = \beta_2 = \dots = \beta_t = \beta^*$. A estatística do teste de razão de verossimilhança para as hipóteses é dada pela equação (4.5) e segue uma distribuição qui-quadrado com $(t-1)(p+1)$ grau de liberdade.

A hipótese nula, no modelo binomial redimensionado e deslocado com inflação (seção 2.3), é representada pelo vetor $\theta \subset \omega$ de tal forma que

$$\omega = \{\theta \in \Omega : \tau_u^* = \tau^*/t, \tau_0^* \leq 1 - \tau^*, \beta_u^* \in \mathbb{R}^p, 1 \leq u \in \mathbb{N} \leq t\}$$

e o modelo irrestrito é tal que $\theta \subset \Omega$, com

$$\Omega = \{\theta : 0 \leq \tau_u \leq 1, \tau_0 + \dots + \tau_t \leq 1, \beta_u \in \mathbb{R}, 0 \leq u \in \mathbb{N} \leq t\}.$$

Sejam

$$L_{\Omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\tau_0 \mathbb{P}_0(z_i) + \sum_{r=1}^t \tau_r \mathbb{P}_r(z_i | \eta_{ri}) + \sum_{r=1}^t (1 - \tau_r) \mathbb{I}_{(z_i=c)} \right]$$

a função de verossimilhança no modelo irrestrito, e

$$L_{\omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\tau_0^* \mathbb{P}_0(z_i) + \tau^* \mathbb{P}_1(z_i | \eta_i^*) + (1 - \tau_0 - \tau^*) \mathbb{I}_{(z_i=c)} \right] \quad (4.14)$$

a função de verossimilhança no modelo restrito por $H_0 : \tau_1 = \tau_2 = \dots = \tau_t = \tau^*/t, \beta_1 = \beta_2 = \dots = \beta_t = \beta^*$. A estatística do teste de razão de verossimilhança para as hipóteses é dada pela equação (4.5) e segue uma distribuição qui-quadrado com $(t-1)(p+1)$ grau de liberdade.

O teste para verificar se os avaliadores que fazem as avaliações de acordo com níveis de preferência são homogêneos é realizado verificando se $\alpha_1 = \alpha_2 = \dots = \alpha_t = \alpha^*/t$ (e, por consequência, $\alpha_0 = 1 - \alpha^*$). No entanto, esta restrição implica que $\beta_1 = \beta_2 = \dots = \beta_t = \beta^*$. Desta forma, a hipótese nula, no modelo binomial duplo (seção 3.1) ou no modelo binomial multiplicativo (seção 3.2), é representada pelo vetor $\theta \in \omega$ de tal forma que

$$\omega = \{\theta \in \Omega : \alpha_u^* = \alpha^*/t, \phi_u^* = \phi^*, 1 \leq u \in \mathbb{N} \leq t; \beta_u^* \in \mathbb{R}^p, 1 \leq u \in \mathbb{N} \leq t\}$$

e o modelo irrestrito é tal que $\theta \in \Omega$, com

$$\Omega = \{\theta : 0 \leq \alpha_u \leq 1, \alpha_0 + \dots + \alpha_t = 1, \beta_u \in \mathbb{R}, 0 \leq u \in \mathbb{N} \leq t\}.$$

Sejam

$$L_{\Omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\sum_{r=1}^t \alpha_r \mathbb{P}_r(z_i | \phi_r, \eta_{ri}) + \alpha_0 \mathbb{P}_0(z_i) \right]$$

a função de verossimilhança no modelo irrestrito, e

$$L_{\omega}(\theta | \mathbf{Z}, \mathbf{X}) = \prod_{i=1}^n \left[\alpha^* \mathbb{P}_1(z_i | \phi^*, \eta_i^*) + (1 - \alpha^*) \mathbb{P}_0(z_i) \right] \quad (4.15)$$

a função de verossimilhança no modelo restrito por $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_t = \alpha^*/t, \phi_1 = \phi_2 = \dots = \phi_t = \phi^*, \beta_1 = \beta_2 = \dots = \beta_t = \beta^*$. A estatística do teste de razão de verossimilhança para as hipóteses é dada pela equação (4.5) e segue uma distribuição qui-quadrado com $(t-1)(p+2)$ grau de liberdade.

4.2.2 Teste de Wald

O teste de Wald (Wald, 1943) é um teste de hipótese assintótico que, desta forma, depende de haver um número grande o suficiente de observações amostrais. Sejam $\theta^* = [\theta_1', \theta_0']'$ o conjunto dos parâmetros sob a hipótese nula, $\hat{\theta}_n^*$ as estimativas (sob o modelo restrito) dos parâmetros e $\hat{\theta}_n$ as estimativas (sob o modelo irrestrito) dos parâmetros para uma amostra de tamanho $n \geq 1$ e $\mathbf{K}(\hat{\theta}_n)$ a matriz de informação observada de Fisher, então (Lehmann; Romano, 2005, p. 510),

$$n(\hat{\theta}_n - \hat{\theta}_n^*)' \mathbf{K}(\hat{\theta}_n)(\hat{\theta}_n - \hat{\theta}_n^*) \xrightarrow{n \rightarrow \infty} \chi_{p-q}^2$$

em que o número de graus de liberdade $p - q = \dim \theta^* - \dim \theta_0$. A hipótese nula é rejeitada se

$$Q = n(\hat{\theta}_n - \hat{\theta}_n^*)' \mathbf{K}(\hat{\theta}_n)(\hat{\theta}_n - \hat{\theta}_n^*) \geq c_{1-\gamma; p-q} \quad (4.16)$$

em que $c_{1-\gamma; p-q}$ é o percentil de ordem $1 - \gamma$ da distribuição qui-quadrado com $p - q$ graus de liberdade.

Existência de avaliações ao acaso

O teste de Wald para a hipótese de não haver avaliações ao acaso no modelo binomial redimensionado e deslocado sem inflação (seção 2.2), dada pela hipótese $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$ é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (2.5). A estatística Q do teste (equação 4.16) segue uma distribuição qui-quadrado com 1 grau de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.6).

O teste de Wald para a hipótese de não haver avaliações ao acaso no modelo binomial redimensionado e deslocado com inflação (seção 2.3), dada pela hipótese $H_0 : \tau_0 = 0$ contra $H_1 : \tau_0 \neq 0$ é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (2.6). A estatística Q do teste (equação 4.16) segue uma distribuição qui-quadrado com 1 grau de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.7).

No modelo binomial duplo (seção 3.1), o teste de Wald para a hipótese de não haver avaliações ao acaso, dado pela hipótese $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$ é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (3.1). A estatística Q do teste (equação 4.16) segue uma distribuição qui-quadrado com 1 grau de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.8).

Por fim, no modelo binomial multiplicativo (seção 3.2), o teste de Wald para a hipótese de não haver avaliações ao acaso, $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$ é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (3.2). A estatística Q do teste (equação 4.16) segue uma distribuição qui-quadrado com 1 grau de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.8).

Existência de inflação na avaliação ao acaso

No modelo binomial redimensionado e deslocado com inflação (seção 2.3), o teste de Wald para a hipótese de não haver inflação nas respostas ao acaso, $H_0 : \tau_0 + \tau_1 + \tau_2 + \dots + \tau_t = 1$ contra $H_1 : \tau_0 + \tau_1 + \tau_2 + \dots + \tau_t \leq 1$, é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (2.6), e a estatística Q (equação 4.16) do teste segue uma distribuição qui-quadrado com 1 grau de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.9).

Testes para os parâmetros de regressão

Considere que o vetor de parâmetros possa ser escrito, para o modelo binomial redimensionado e deslocado sem inflação (seção 2.2), como $\boldsymbol{\theta} = [\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\beta}^{*'}]'$ em que $\boldsymbol{\beta}$ é o conjunto de parâmetros associados aos termos não restrito pela hipótese nula e $\boldsymbol{\beta}^*$ é o conjunto de parâmetros associados aos termos que sofrem alguma restrição nos parâmetros devido à hipótese nula. Sem perda de generalidade, suponha que os termos $1 \leq r \leq t^* \in \mathbb{N} < t$ não sejam restritos pela

hipótese nula e que os termos $t^* < r \leq t$ o sejam. A estatística do teste de Wald será calculada como em (4.16) e o número de graus de liberdade será $p - q$ em que p é a dimensão do vetor β e q é a quantidade de parâmetros restritos em β^* . O vetor de parâmetros sob a hipótese nula é estimado no modelo não inflacionado usando a verossimilhança (4.10).

Para o modelo binomial redimensionado e deslocado inflacionado (seção 2.3) o procedimento é similar. Se $\theta = [\tau', \beta', \beta^{*'}]'$, com β e β^* definidos como antes, então a estatística do teste de Wald será calculada como em (4.16) e o número de graus de liberdade será $p - q$ em que p é a dimensão do vetor β e q é a quantidade de parâmetros restritos em β^* . O vetor de parâmetros sob a hipótese nula é estimado no modelo inflacionado usando a verossimilhança (4.11).

Nos modelos binomial duplo (seção 3.1) e binomial multiplicativo (seção 3.2), tomando $\theta = [\tau', \beta', \beta^{*'}]'$, com β e β^* definidos como antes, então a estatística do teste de Wald será calculada como em (4.16) e o número de graus de liberdade será $p - q$ em que p é a dimensão do vetor β e q é a quantidade de parâmetros restritos em β^* . O vetor de parâmetros sob a hipótese nula é estimado no modelo inflacionado usando a verossimilhança (4.12) com as respectivas distribuições de probabilidade.

Homogeneidade entre avaliadores (na avaliação por preferência)

No teste de homogeneidade entre avaliadores o interesse é verificar se, quando as avaliações são feitas por preferência, existem mais que um nível de preferência. Neste caso, sob a hipótese nula, o vetor de parâmetros para o modelo binomial redimensionado e deslocado sem inflação (seção 2.2) pode ser escrito como $\theta^* = [\alpha, \beta']$ em que α é o parâmetro da mistura e β é um vetor de dimensão p de parâmetros de regressão associado ao termo da avaliação sob preferência. O vetor de parâmetros no modelo irrestrito é dado por $\theta = [\alpha', \beta'_1, \beta'_2, \dots, \beta'_t]$ em que α é um vetor de dimensão t com os parâmetros da mistura e os vetores dos parâmetros de regressão β_r , $r = 1, 2, \dots, t$ possuem dimensão p . Desta forma, a estatística do teste é calculada como em (4.16) e se distribui de acordo com uma qui-quadrado com $(t - 1)(p + 1)$ graus de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.13).

De forma análoga, para o modelo binomial redimensionado e deslocado inflacionado (seção 2.3), o vetor de parâmetros no modelo irrestrito é dado por $\theta = [\tau', \beta'_1, \beta'_2, \dots, \beta'_t]$ em que τ é um vetor de dimensão t com os parâmetros da mistura e os vetores dos parâmetros de regressão β_r , $r = 1, 2, \dots, t$ possuem dimensão p . Sob a hipótese nula o vetor de parâmetros é dada por $\theta = [\tau, \beta']$. A estatística do teste é calculada como em (4.16) e é ditribuída de acordo com uma qui-quadrado com $(t - 1)(p + 1)$ graus de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.14).

Finalmente, para os modelos binomial duplo (seção 3.1) e binomial multiplicativo (seção 3.2), o vetor de parâmetros no modelo irrestrito é dado por $\theta = [\alpha', \phi', \beta'_1, \beta'_2, \dots, \beta'_t]$ em que α é um vetor de dimensão t com os parâmetros da mistura, ϕ é um vetor de dimensão t com os parâmetros de variância e os vetores dos parâmetros de regressão β_r , $r = 1, 2, \dots, t$ possuem dimensão p . Sob a hipótese nula o vetor de parâmetros é dada por $\theta = [\alpha, \phi, \beta']$. A

estatística do teste é calculada como em (4.16) e é ditribuída de acordo com uma qui-quadrado com $(t-1)(p+2)$ graus de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.15) calculada com as respectivas distribuições de probabilidade.

4.2.3 Teste escore

O teste *escore* de Rao (1948) é uma alternativa ao teste de Wald. Seja $\mathbf{U}(\boldsymbol{\theta})$ o vetor escore definido por

$$\mathbf{U}(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log L(\boldsymbol{\theta} | \mathbf{Z}, \mathbf{X})$$

e $\mathbf{K}(\hat{\boldsymbol{\theta}})$ a matriz de informação observada de Fisher. Desta forma,

$$\mathbf{U}(\boldsymbol{\theta}^*)' \mathbf{K}(\boldsymbol{\theta}^*)^{-1} \mathbf{U}(\boldsymbol{\theta}^*) \xrightarrow{n \rightarrow \infty} \chi_{p-q}^2$$

em que $\boldsymbol{\theta}^*$ é o vetor de parâmetros restrito à hipótese nula. O número de graus de liberdade é, como antes, dado por $p - q = \dim \boldsymbol{\theta} - \dim \boldsymbol{\theta}_0$. A hipótese nula do teste é rejeitada se

$$Q = \mathbf{U}(\hat{\boldsymbol{\theta}}_n^*)' \mathbf{K}(\hat{\boldsymbol{\theta}}_n^*)^{-1} \mathbf{U}(\hat{\boldsymbol{\theta}}_n^*) \geq c_{1-\gamma; p-q} \quad (4.17)$$

em que $\hat{\boldsymbol{\theta}}_n^*$ é a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$ sob a hipótese nula considerando uma amostra de tamanho $n \geq 1$ e $c_{1-\gamma; p-q}$ é o percentil de ordem $1 - \gamma$ da distribuição qui-quadrado com $p - q$ graus de liberdade.

O vetor escore $\mathbf{U}(\boldsymbol{\theta})$ para o modelo binomial redimensionado e deslocado sem inflação (seção 2.2) é na forma $\mathbf{U}(\boldsymbol{\theta}) = [\mathbf{U}_{\alpha}(\boldsymbol{\theta})', \mathbf{U}_{\beta_1}(\boldsymbol{\theta})', \mathbf{U}_{\beta_2}(\boldsymbol{\theta})', \dots, \mathbf{U}_{\beta_t}(\boldsymbol{\theta})']'$ em que $\boldsymbol{\theta} = [\alpha_1, \alpha_2, \dots, \alpha_t, \beta'_1, \beta'_2, \dots, \beta'_t]'$, e $\mathbf{U}_{\alpha}(\boldsymbol{\theta})$ é um vetor de dimensão t com a r -ésima posição, $r = 1, 2, \dots, t$, dada por

$$U_{\alpha_r} = \sum_{i=1}^n \frac{\mathbb{P}_r(z_i | \eta_{ri}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \quad (4.18)$$

e

$$\mathbf{U}_{\beta_r}(\boldsymbol{\theta})' = \sum_{i=1}^n (\mathbf{x}'_i) \left(\frac{\alpha_r \mathbb{P}_r(z_i | \eta_{ri})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} \right). \quad (4.19)$$

O vetor escore $\mathbf{U}(\boldsymbol{\theta})$ para o modelo binomial redimensionado e deslocado com inflação (seção 2.3) é na forma $\mathbf{U}(\boldsymbol{\theta}) = [\mathbf{U}_{\tau}(\boldsymbol{\theta})', \mathbf{U}_{\beta_1}(\boldsymbol{\theta})', \mathbf{U}_{\beta_2}(\boldsymbol{\theta})', \dots, \mathbf{U}_{\beta_t}(\boldsymbol{\theta})']'$ em que $\boldsymbol{\theta} = [\tau_0, \tau_1, \tau_2, \dots, \tau_t, \beta'_1, \beta'_2, \dots, \beta'_t]'$, e $\mathbf{U}_{\tau}(\boldsymbol{\theta})$ é um vetor de dimensão $t+1$ com a posição 0 dada por

$$U_{\tau_0} = \sum_{i=1}^n \frac{\mathbb{P}_0(z_i) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)}, \quad (4.20)$$

a r -ésima posição, $r = 1, 2, \dots, t$, dada por

$$U_{\alpha_r} = \sum_{i=1}^n \frac{\mathbb{P}_r(z_i | \eta_{ri}) - \mathbb{I}_{(z_i=c)}}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \quad (4.21)$$

e

$$\mathbf{U}_{\beta_r}(\boldsymbol{\theta})' = \sum_{i=1}^n (\mathbf{x}'_i) \left(\frac{\tau_r \mathbb{P}_r(z_i | \eta_{ri})}{\mathbb{P}(z_i | \boldsymbol{\eta}_i)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} \right). \quad (4.22)$$

O vetor escore $\mathbf{U}(\boldsymbol{\theta})$ para o modelo binomial duplo (seção 3.1) é na forma $\mathbf{U}(\boldsymbol{\theta}) = [\mathbf{U}_\alpha(\boldsymbol{\theta})', \mathbf{U}_\phi(\boldsymbol{\theta})', \mathbf{U}_{\beta_1}(\boldsymbol{\theta})', \mathbf{U}_{\beta_2}(\boldsymbol{\theta})', \dots, \mathbf{U}_{\beta_t}(\boldsymbol{\theta})']'$ em que $\boldsymbol{\theta} = [\alpha_1, \dots, \alpha_t, \phi_1, \dots, \phi_t, \beta'_1, \dots, \beta'_t]'$, $\mathbf{U}_\alpha(\boldsymbol{\theta})$ é um vetor de dimensão t com a r -ésima posição dada por

$$U_{\alpha_r} = \sum_{i=1}^n \frac{\mathbb{P}_r(z_i|\eta_{ri}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i|\boldsymbol{\eta}_i)}, \quad (4.23)$$

$\mathbf{U}_\phi(\boldsymbol{\theta})$ é um vetor de dimensão t com a r -ésima posição dada por

$$U_{\phi_r} = \sum_{i=1}^n \left(\frac{\alpha_r \mathbb{P}_r(z_i|\theta_r, \eta_{ri})}{\mathbb{P}(z_i|\boldsymbol{\theta}, \boldsymbol{\eta}_i)} \right) \times \left[\frac{1}{2\theta_r} - \left[\ln \left(\left(\frac{z_i}{m} \right)^{z_i} \left(1 - \frac{z_i}{m} \right)^{m-z_i} \right) - z_i \eta_{ri} + m \ln(1 + e^{\eta_{ri}}) \right] \right], \quad (4.24)$$

e

$$\mathbf{U}_{\beta_r}(\boldsymbol{\theta})' = \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \theta_r \left(\frac{\alpha_r \mathbb{P}_r(z_i|\phi_r, \eta_{ri})}{\mathbb{P}(z_i|\boldsymbol{\phi}, \boldsymbol{\eta}_i)} \right) \left(z_i - \frac{m e^{\eta_{ri}}}{1 + e^{\eta_{ri}}} \right). \quad (4.25)$$

O vetor escore $\mathbf{U}(\boldsymbol{\theta})$ para o modelo binomial multiplicativo (seção 3.2) é na forma $\mathbf{U}(\boldsymbol{\theta}) = [\mathbf{U}_\alpha(\boldsymbol{\theta})', \mathbf{U}_\phi(\boldsymbol{\theta})', \mathbf{U}_{\beta_1}(\boldsymbol{\theta})', \mathbf{U}_{\beta_2}(\boldsymbol{\theta})', \dots, \mathbf{U}_{\beta_t}(\boldsymbol{\theta})']'$ em que $\boldsymbol{\theta} = [\alpha_1, \dots, \alpha_t, \phi_1, \dots, \phi_t, \beta'_1, \dots, \beta'_t]'$, $\mathbf{U}_\alpha(\boldsymbol{\theta})$ é um vetor de dimensão t com a r -ésima posição dada por

$$U_{\alpha_r} = \sum_{i=1}^n \frac{\mathbb{P}_r(z_i|\eta_{ri}) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i|\boldsymbol{\eta}_i)}, \quad (4.26)$$

$\mathbf{U}_\phi(\boldsymbol{\theta})$ é um vetor de dimensão t com a r -ésima posição dada por

$$U_{\phi_r} = \sum_{i=1}^n \left(\frac{\alpha_r \mathbb{P}_r(z_i|\phi_r, \eta_{ri})}{\mathbb{P}(z_i|\boldsymbol{\phi}, \boldsymbol{\eta}_i)} \right) (z_i(m - z_i) - \mathbb{E}(Z(m - Z)|\phi_r, \eta_{ri})) \quad (4.27)$$

e

$$\mathbf{U}_{\beta_r}(\boldsymbol{\theta})' = \sum_{i=1}^n (\mathbf{x}_i \mathbf{x}_i') \left(\frac{\alpha_r \mathbb{P}_r(z_i|\phi_r, \eta_{ri})}{\mathbb{P}(z_i|\boldsymbol{\phi}, \boldsymbol{\eta}_i)} \right) (z_i - \mathbb{E}(Z|\phi_r, \eta_{ri})). \quad (4.28)$$

A matriz $\mathbf{K}(\boldsymbol{\theta})$ é a informação de Fisher

$$\mathbf{K}(\hat{\boldsymbol{\theta}}) = \mathbf{V}(\hat{\boldsymbol{\theta}})^{-1}$$

em que $\mathbf{V}(\hat{\boldsymbol{\theta}})$ é a matriz de variância dadas, respectivamente, pelas expressões (2.5) para o modelo binomial redimensionado e deslocado sem inflação, (2.6) para o modelo binomial redimensionado e deslocado inflacionado, (3.1) para o modelo binomial duplo e (3.1) para o modelo binomial multiplicativo.

Existência de avaliações ao acaso

O teste escore para a hipótese de não haver avaliações ao acaso no modelo binomial redimensionado e deslocado sem inflação (seção 2.2), $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$, é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (2.5) e o vetor

escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.18) e (4.19). A estatística (equação 4.17) do teste segue uma distribuição qui-quadrado com 1 grau de liberdade e o vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.6).

O teste escore para a hipótese de inexistência de avaliações ao acaso no modelo binomial redimensionado e deslocado com inflação (seção 2.3), $H_0 : \tau_0 = 0, \tau_1 + \tau_2 + \dots + \tau_t = 1$ contra $H_1 : \tau_0 \neq 0$, é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (2.6) e o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.20), (4.21) e (4.22). A estatística Q do teste (equação 4.17) segue uma distribuição qui-quadrado com 1 graus de liberdade e vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.7).

O teste escore para a hipótese de não haver avaliações ao acaso no modelo binomial duplo (seção 3.1), $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$, é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (3.1) e o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.23), (4.24) e (4.25). A estatística (equação 4.17) do teste segue uma distribuição qui-quadrado com 1 grau de liberdade e o vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.8).

O teste escore para a hipótese de não haver avaliações ao acaso no modelo binomial multiplicativo (seção 3.2), $H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$, é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (3.2) e o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.26), (4.27) e (4.28). A estatística (equação 4.17) do teste segue uma distribuição qui-quadrado com 1 grau de liberdade e o vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.8).

Existência de inflação na avaliação ao acaso

No modelo binomial redimensionado e deslocado com inflação, o teste score para a hipótese de não haver inflação nas respostas ao acaso ($H_0 : \tau_0 + \tau_1 + \tau_2 + \dots + \tau_t = 1$ contra $H_1 : \tau_0 + \tau_1 + \tau_2 + \dots + \tau_t \leq 1$) é tal que a matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ é o inverso da matriz de variância dada em (2.6) e o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.20), (4.21) e (4.22). A estatística Q (equação 4.17) do teste segue uma distribuição qui-quadrado com 1 grau de liberdade e o vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.9).

Testes para os parâmetros de regressão

Da mesma forma como no teste de Wald, considere que o vetor de parâmetros possa ser escrito, para o modelo binomial redimensionado e deslocado sem inflação (seção ??), como $\boldsymbol{\theta} = [\boldsymbol{\alpha}', \boldsymbol{\beta}', (\boldsymbol{\beta}^*)']'$ em que $\boldsymbol{\beta}$ é o conjunto de parâmetros associados aos termos não restrito pela hipótese nula e $\boldsymbol{\beta}^*$ é o conjunto de parâmetros associados aos termos que sofrem alguma restrição nos parâmetros devido à hipótese nula. Sem perda de generalidade, suponha que os termos $1 \leq r \leq t^* \in \mathbb{N} < t$ não sejam restritos pela hipótese nula e que os termos $t^* < r \leq t$ sejam. A estatística do teste score será calculada pelas equações (4.17), em que o vetor escore é

dado pelas equações (4.18) e (4.22), e o número de graus de liberdade será $p - q$ tal que p é a dimensão do vetor β e q é a quantidade de parâmetros restritos em β^* . O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.10).

Para o modelo binomial redimensionado e deslocado com inflação (seção 2.3) o procedimento é similar. Se $\theta = [\tau', \beta', (\beta^*)']'$, com β e β^* definidos como antes, então a estatística do teste escore será calculada como em (4.17), em que o vetor escore é dado pelas equações (4.20), (4.21) e (4.22), e o número de graus de liberdade será $p - q$ em que p é a dimensão do vetor β e q é a quantidade de parâmetros restritos em β^* . O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.11).

No modelo binomial duplo (seção 3.1), se $\theta = [\alpha', \phi', \beta', (\beta^*)']'$, com β e β^* definidos como antes, então a estatística do teste escore será calculada como em (4.17), em que o vetor escore é dado pelas equações (4.23), (4.24) e (4.25), e o número de graus de liberdade será $p - q$ em que p é a dimensão do vetor β e q é a quantidade de parâmetros restritos em β^* . O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.12).

Por último, no modelo binomial multiplicativo (seção 3.2), se $\theta = [\alpha', \phi', \beta', (\beta^*)']'$, com β e β^* definidos como antes, então a estatística do teste escore será calculada como em (4.17), em que o vetor escore é dado pelas equações (4.26), (4.27) e (4.28), e o número de graus de liberdade será $p - q$ em que p é a dimensão do vetor β e q é a quantidade de parâmetros restritos em β^* . O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.12).

Homogeneidade entre avaliadores (na avaliação por preferência)

Como no teste de Wald, sob a hipótese nula, o vetor de parâmetros para o modelo binomial redimensionado e deslocado sem inflação (seção 2.2) pode ser escrito como $\theta^* = [\alpha, \beta']$ em que α é o parâmetro da mistura e β é um vetor de dimensão p de parâmetros de regressão associado ao termo da avaliação sob preferência. O vetor de parâmetros no modelo irrestrito é dado por $\theta = [\alpha', \beta'_1, \beta'_2, \dots, \beta'_t]$ em que α é um vetor de dimensão t com os parâmetros da mistura e os vetores dos parâmetros de regressão β_r , $r = 1, 2, \dots, t$ possuem dimensão p . Desta forma, a estatística do teste é calculada como em (4.17) com o vetor escore $U(\theta)$ dado pelas equações (4.18) e (4.19), distribuída de acordo com uma qui-quadrado com $(t - 1)(p + 1)$ graus de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.13).

Para o modelo binomial redimensionado e deslocado inflacionado (seção 2.3), sob a hipótese nula, o vetor de parâmetros para o modelo sem inflação pode ser escrito como $\theta^* = [\tau, \beta']$ em que τ é o parâmetro da mistura e β é um vetor de dimensão p de parâmetros de regressão associado ao termo da avaliação sob preferência e o vetor de parâmetros no modelo irrestrito é dado por $\theta = [\tau', \beta'_1, \beta'_2, \dots, \beta'_t]$ em que τ é um vetor de dimensão $t + 1$ com os parâmetros da mistura e os vetores dos parâmetros de regressão β_r , $r = 1, 2, \dots, t$ possuem dimensão p . Desta forma, a estatística do teste é calculada como em (4.17) com o vetor escore $U(\theta)$ dado pelas equações (4.20), (4.21) e (4.22), é distribuída de acordo com uma qui-quadrado com

$(t-1)(p+1)$ graus de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.14).

No modelo binomial duplo (seção 3.1), o vetor de parâmetros pode ser escrito como $\theta^* = [\alpha, \beta']$ em que α é o parâmetro da mistura e β é um vetor de dimensão p de parâmetros de regressão associado ao termo da avaliação sob preferência. O vetor de parâmetros no modelo irrestrito é dado por $\theta = [\alpha', \beta'_1, \beta'_2, \dots, \beta'_t]$ em que α é um vetor de dimensão t com os parâmetros da mistura e os vetores dos parâmetros de regressão β_r , $r = 1, 2, \dots, t$ possuem dimensão p . Desta forma, a estatística do teste é calculada como em (4.17) com o vetor escore $U(\theta)$ dado pelas equações (4.23), (4.24) e (4.25), distribuída de acordo com uma qui-quadrado com $(t-1)(p+2)$ graus de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.15).

No modelo binomial multiplicativo (seção 3.2) pode ser escrito como $\theta^* = [\alpha, \beta']$ em que α é o parâmetro da mistura e β é um vetor de dimensão p de parâmetros de regressão associado ao termo da avaliação sob preferência. O vetor de parâmetros no modelo irrestrito é dado por $\theta = [\alpha', \beta'_1, \beta'_2, \dots, \beta'_t]$ em que α é um vetor de dimensão t com os parâmetros da mistura e os vetores dos parâmetros de regressão β_r , $r = 1, 2, \dots, t$ possuem dimensão p . Desta forma, a estatística do teste é calculada como em (4.17) com o vetor escore $U(\theta)$ dado pelas equações (4.26), (4.27) e (4.28), distribuída de acordo com uma qui-quadrado com $(t-1)(p+2)$ graus de liberdade. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.15).

4.2.4 Teste gradiente

O teste *gradiente* foi proposto por Terrell (2002). Seja a hipótese nula dada por $H_0 : \theta_2 = \theta_0$ e a hipótese alternativa dada por $H_1 : \theta_2 \neq \theta_0$. A estatística do teste gradiente é definido por

$$S = U(\hat{\theta})' (\hat{\theta} - \hat{\theta}^*) \quad (4.29)$$

em que $\hat{\theta}$ é a estimativa de máxima verossimilhança do vetor de parâmetros θ sob o modelo irrestrito e $\hat{\theta}^*$ é a estimativa de máxima verossimilhança do vetor de parâmetros sob a hipótese nula.

Considere que o vetor de parâmetros θ pode ser escrito como $\theta = [\theta'_1, \theta'_2]'$ com θ_1 o vetor de parâmetros não restritos pela hipótese nula e θ_2 o vetor de parâmetros com restrição. Supondo que, sob a hipótese nula, a estimativa de θ_1 seja aproximadamente igual à estimativa no modelo irrestrito, a estatística do teste gradiente pode ser aproximada por

$$S \approx U(\hat{\theta}_2)' (\hat{\theta}_2 - \theta_0) \quad (4.30)$$

em que $\hat{\theta}_2$ é a estimativa de θ_2 sob a hipótese nula e θ_0 é o vetor de restrições impostas pela mesma hipótese. A estatística do teste segue uma distribuição qui-quadrado com $p - q = \dim \theta - \dim \theta_0$ graus de liberdade.

Existência de avaliações ao acaso

O teste gradiente no modelo binomial redimensionado e deslocado sem inflação (seção 2.2) para a hipótese de não haver avaliações ao acaso ($H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$) é tal que o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.18) e (4.19). A estatística do teste (equação 4.29) segue uma distribuição qui-quadrado com 1 grau de liberdade e, supondo que os parâmetros de regressão do modelo não sofrem influência da hipótese nula, de acordo com (4.30) pode ser aproximada por

$$S \approx \sum_{i=1}^n \sum_{r=1}^t (\hat{\alpha}_r - \hat{\alpha}_r^*) \left(\frac{\mathbb{P}_r(z_i | \hat{\phi}_r^*, \hat{\eta}_{ri}^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \hat{\phi}_r^*, \hat{\eta}_i^*)} \right)$$

com $\hat{\alpha}_r$ e $\hat{\alpha}_r^*$, $r = 1, 2, \dots, t$, respectivamente, as estimativas do parâmetro α_r sob o modelo sem restrição e sob a hipótese nula. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.6).

O teste gradiente para a hipótese de não haver avaliações ao acaso no modelo binomial redimensionado e deslocado com inflação ($H_0 : \tau_0 = 0, \tau_1 + \tau_2 + \dots + \tau_t = 1$ contra $H_1 : \tau_0 \neq 0, \tau_1 + \tau_2 + \dots + \tau_t \leq 1$) é tal que o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.20), (4.21) e (4.22). A estatística S do teste (equação 4.29) segue uma distribuição qui-quadrado com 1 grau de liberdade e, supondo novamente que os parâmetros de regressão do modelo não sofrem influências significativas da hipótese nula, de acordo com (4.30) pode ser aproximada por

$$S \approx \sum_{i=1}^n \hat{\tau}_0 \left(\frac{\mathbb{P}_0(z_i) - \mathbb{I}_{z_i=c}}{\mathbb{P}(z_i | \hat{\eta}_i)} \right) + \sum_{i=1}^n \sum_{r=1}^t (\hat{\tau}_r - \hat{\tau}_r^*) \left(\frac{\mathbb{P}_r(z_i | \hat{\eta}_{ri}^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \hat{\eta}_i^*)} \right)$$

com $\hat{\tau}_0$ a estimativa sob o modelo sem restrição de τ_0 e $\hat{\tau}_r$ e $\hat{\tau}_r^*$, $r = 1, 2, \dots, t$, respectivamente, as estimativas do parâmetro τ_r sob o modelo sem restrição e sob a hipótese nula. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.7).

O teste gradiente no modelo binomial duplo (seção 3.1) para a hipótese de não haver avaliações ao acaso ($H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$) é tal que o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.23), (4.24) e (4.25). A estatística do teste (equação 4.29) segue uma distribuição qui-quadrado com 1 grau de liberdade e, supondo que os parâmetros de regressão do modelo não sofrem influência da hipótese nula, de acordo com (4.30) pode ser aproximada por

$$S \approx \sum_{i=1}^n \sum_{r=1}^t (\hat{\alpha}_r - \hat{\alpha}_r^*) \left(\frac{\mathbb{P}_r(z_i | \hat{\phi}_r^*, \hat{\eta}_{ri}^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \hat{\phi}_r^*, \hat{\eta}_i^*)} \right)$$

com $\hat{\alpha}_r$ e $\hat{\alpha}_r^*$, $r = 1, 2, \dots, t$, respectivamente, as estimativas do parâmetro α_r sob o modelo sem restrição e sob a hipótese nula. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.8) com as funções de probabilidade correspondentes ao modelo binomial duplo.

O teste gradiente no modelo binomial multiplicativo (seção 3.2) para a hipótese de não haver avaliações ao acaso ($H_0 : \alpha_0 = 0$ contra $H_1 : \alpha_0 \neq 0$) é tal que o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.26), (4.27) e (4.28). A estatística do teste (equação 4.29) segue uma

distribuição qui-quadrado com 1 grau de liberdade e, supondo que os parâmetros de regressão do modelo não sofrem influência da hipótese nula, de acordo com (4.30) pode ser aproximada por

$$S \approx \sum_{i=1}^n \sum_{r=1}^t (\hat{\alpha}_r - \hat{\alpha}_r^*) \left(\frac{\mathbb{P}_r(z_i | \hat{\phi}_r^*, \hat{\eta}_{ri}^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \hat{\phi}_r^*, \hat{\eta}_i^*)} \right)$$

com $\hat{\alpha}_r$ e $\hat{\alpha}_r^*$, $r = 1, 2, \dots, t$, respectivamente, as estimativas do parâmetro α_r sob o modelo sem restrição e sob a hipótese nula. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.8) com as funções de probabilidade correspondentes ao modelo binomial multiplicativo.

Existência de inflação na avaliação ao acaso

No modelo binomial redimensionado e deslocado com inflação (seção 2.3), o teste gradiente para a hipótese de não haver inflação nas respostas ao acaso ($H_0 : \tau_0 + \tau_1 + \tau_2 + \dots + \tau_t = 1$ contra $H_1 : \tau_0 + \tau_1 + \tau_2 + \dots + \tau_t \leq 1$) é tal que o vetor escore $\mathbf{U}(\boldsymbol{\theta})$ é dado pelas equações (4.20), (4.21) e (4.22). A estatística S (equação 4.29) do teste segue uma distribuição qui-quadrado com 1 grau de liberdade e, supondo novamente que os parâmetros de regressão do modelo não sofrem influências significativas da hipótese nula, de acordo com (4.30) pode ser aproximada por

$$S \approx \sum_{i=1}^n (\hat{\tau}_0 - \hat{\tau}_0^*) \left(\frac{\mathbb{P}_0(z_i) - \mathbb{I}_{z_i=c}}{\mathbb{P}_r(z_i | \hat{\eta}_i^*)} \right) + \sum_{i=1}^n \sum_{r=1}^t (\hat{\tau}_r - \hat{\tau}_r^*) \left(\frac{\mathbb{P}_r(z_i | \hat{\eta}_{ri}^*) - \mathbb{P}_0(z_i)}{\mathbb{P}_r(z_i | \hat{\eta}_i^*)} \right)$$

com $\hat{\tau}_0$ e $\hat{\tau}_0^*$, respectivamente, as estimativas de τ_0 sob o modelo sem restrição e sob a hipótese nula, e $\hat{\tau}_r$ e $\hat{\tau}_r^*$, $r = 1, 2, \dots, t$, respectivamente, as estimativas do parâmetro τ_r sob o modelo sem restrição e sob a hipótese nula. O vetor de parâmetros sob a hipótese nula é estimado usando a verossimilhança (4.9) com as funções de probabilidade correspondentes ao modelo binomial multiplicativo.

Testes para os parâmetros de regressão

No modelo binomial redimensionado e deslocado sem inflação (seção 2.2), para realizar o teste gradiente para os parâmetros da regressão, considere que o vetor de parâmetros de regressão para o r -ésimo componente binomial seja dado por $\boldsymbol{\beta}_r = \{\boldsymbol{\beta}_{1r}, \boldsymbol{\beta}_{2r}\}$ e que sob a hipótese nula $\boldsymbol{\beta}_{2r} = \boldsymbol{\beta}_{0r}$. Neste cenário, a função escore é dada pelas equações (4.18) e (4.19), calculados a partir da verossimilhança em (4.10), e a estatística do teste (4.29) pode ser escrita como

$$S \approx \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\boldsymbol{\beta}}_{2r}^* - \boldsymbol{\beta}_{0r}) \left(\frac{\hat{\alpha}_r^* \mathbb{P}_r(z_i | \hat{\eta}_{ri}^*)}{\mathbb{P}(z_i | \hat{\eta}_i^*)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\hat{\eta}_{ri}^*}}{1 + e^{\hat{\eta}_{ri}^*}} \right).$$

A estatística do teste segue uma distribuição qui-quadrado com graus de liberdade dados por

$$d = \sum_{r=1}^t (\dim \boldsymbol{\beta}_r - \dim \boldsymbol{\beta}_{2r}).$$

No modelo binomial redimensionado e deslocado com inflação (seção 2.3), para realizar o teste gradiente para os parâmetros da regressão, considere que o vetor de parâmetros de regressão

para o r -ésimo componente binomial seja dado por $\beta_r = \{\beta_{1r}, \beta_{2r}\}$ e que sob a hipótese nula $\beta_{2r} = \beta_{0r}$. Neste cenário, a função escore é dada pelas equações (4.20), (4.21) e (4.22), calculados a partir da verossimilhança em (4.11), e a estatística do teste (4.29) pode ser escrita como

$$S \approx \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\beta}_{2r}^* - \beta_{0r}) \left(\frac{\hat{\tau}_r^* \mathbb{P}_r(z_i | \hat{\eta}_{ri}^*)}{\mathbb{P}(z_i | \hat{\eta}_i^*)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\hat{\eta}_{ri}^*}}{1 + e^{\hat{\eta}_{ri}^*}} \right).$$

A estatística do teste segue uma distribuição qui-quadrado com graus de liberdade dados por

$$d = \sum_{r=1}^t (\dim \beta_r - \dim \beta_{2r}).$$

No modelo binomial duplo (seção 3.1), para realizar o teste gradiente para os parâmetros da regressão, considere que o vetor de parâmetros de regressão para o r -ésimo componente binomial seja dado por $\beta_r = \{\beta_{1r}, \beta_{2r}\}$ e que sob a hipótese nula $\beta_{2r} = \beta_{0r}$. Neste cenário, a função escore é dada pelas equações (4.23), (4.24) e (4.25), calculados a partir da verossimilhança em (4.12) com relação à distribuição de probabilidade do modelo binomial duplo, e a estatística do teste (4.29) pode ser escrita como

$$S \approx \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\beta}_{2r}^* - \beta_{0r}) \left(\frac{\hat{\alpha}_r^* \mathbb{P}_r(z_i | \hat{\eta}_{ri}^*)}{\mathbb{P}(z_i | \hat{\eta}_i^*)} \right) \left(z_i - \frac{me^{\hat{\eta}_{ri}^*}}{1 + e^{\hat{\eta}_{ri}^*}} \right).$$

A estatística do teste segue uma distribuição qui-quadrado com graus de liberdade dados por

$$d = \sum_{r=1}^t (\dim \beta_r - \dim \beta_{2r}).$$

No modelo binomial multiplicativo (seção 3.2), para realizar o teste gradiente para os parâmetros da regressão, considere que o vetor de parâmetros de regressão para o r -ésimo componente binomial seja dado por $\beta_r = \{\beta_{1r}, \beta_{2r}\}$ e que sob a hipótese nula $\beta_{2r} = \beta_{0r}$. Neste cenário, a função escore é dada pelas equações (4.26), (4.27) e (4.28), calculados a partir da verossimilhança em (4.12) com relação à distribuição de probabilidade do modelo binomial multiplicativo, e a estatística do teste (4.29) pode ser escrita como

$$S \approx \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\beta}_{2r}^* - \beta_{0r}) \left(\frac{\hat{\alpha}_r^* \mathbb{P}_r(z_i | \hat{\phi}_r^*, \hat{\eta}_{ri}^*)}{\mathbb{P}(z_i | \hat{\phi}_r^*, \hat{\eta}_i^*)} \right) \left(z_i - \mathbb{E}(Z | \hat{\phi}_r^*, \hat{\eta}_{ri}^*) \right).$$

A estatística do teste segue uma distribuição qui-quadrado com graus de liberdade dados por

$$d = \sum_{r=1}^t (\dim \beta_r - \dim \beta_{2r}).$$

Homogeneidade entre avaliadores (na avaliação por preferência)

No modelo binomial redimensionado e deslocado sem inflação (seção 2.2), o teste gradiente para verificar a homogeneidade entre avaliadores (ou seja, a inexistência de diferentes tipos de avaliadores) consiste em verificar se os vetores de parâmetros de regressão $\beta_r = \beta$, $r = 1, 2, \dots, t$,

e, além disto, $\alpha_1 = \alpha_2 = \dots = \alpha_t = \alpha/t$. A estatística do teste segue uma distribuição qui-quadrado com $(t-1)(p-1)$ graus de liberdade, em que p é o número de parâmetros de regressão para cada componente, e é aproximada por

$$S \approx \sum_{i=1}^n \left(\sum_{r=1}^t \hat{\alpha}_r^* - \hat{\alpha}^* \right) \left(\frac{\mathbb{P}_1(z_i|\hat{\eta}_i^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i|\hat{\eta}_i^*)} \right) \\ + \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\beta}_{2r}^* - \hat{\beta}^*) \left(\frac{(\hat{\alpha}_r/t)\mathbb{P}_1(z_i|\hat{\eta}_i^*)}{\mathbb{P}(z_i|\hat{\eta}_i^*)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\hat{\eta}_i^*}}{1+e^{\hat{\eta}_i^*}} \right)$$

com $\hat{\beta}^*$ e $\hat{\eta}_i^*$ a estimativa do vetor de parâmetros de regressão e do preditor linear para a i -ésima observação amostral sob o modelo homogêneo.

No modelo binomial redimensionado e deslocado com inflação (seção 2.3), o teste gradiente para verificar a homogeneidade entre avaliadores (ou seja, a inexistência de diferentes tipos de avaliadores) consiste em verificar se os vetores de parâmetros de regressão $\beta_r = \beta$, $r = 1, 2, \dots, t$, e se $\tau_1 = \tau_2 = \dots = \tau_t = \tau/t$. A estatística do teste segue uma distribuição qui-quadrado com $(t-1)(p-1)$ graus de liberdade, em que p é o número de parâmetros de regressão para cada componente, e é aproximada por

$$S \approx \sum_{i=1}^n \left(\sum_{r=1}^t \hat{\tau}_r^* - \hat{\tau}^* \right) \left(\frac{\mathbb{P}_1(z_i|\eta_i^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i|\eta_i^*)} \right) \\ + \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\beta}_{2r}^* - \hat{\beta}^*) \left(\frac{(\hat{\tau}_r/t)\mathbb{P}_1(z_i|\hat{\eta}_i^*)}{\mathbb{P}(z_i|\hat{\eta}_i^*)} \right) \left(\frac{z_i}{s} - d - \frac{(m-d)e^{\hat{\eta}_i^*}}{1+e^{\hat{\eta}_i^*}} \right)$$

com $\hat{\beta}^*$ e $\hat{\eta}_i^*$ a estimativa do vetor de parâmetros de regressão e do preditor linear para a i -ésima observação amostral sob o modelo homogêneo.

No modelo binomial duplo (seção 3.1), o teste gradiente para verificar a homogeneidade entre avaliadores (ou seja, a inexistência de diferentes tipos de avaliadores) consiste em verificar se de regressão $\beta_r = \beta$, $r = 1, 2, \dots, t$, $\alpha_1 = \alpha_2 = \dots = \alpha_t = \alpha/t$ e $\phi_1 = \phi_2 = \dots = \phi_t = \phi/t$. A estatística do teste segue uma distribuição qui-quadrado com $(t-1)(p-2)$ graus de liberdade, em que p é o número de parâmetros de regressão para cada componente, e é aproximada por

$$S \approx \sum_{i=1}^n \left(\sum_{r=1}^t \hat{\alpha}_r^* - \hat{\alpha}^* \right) \left(\frac{\mathbb{P}_1(z_i|\hat{\phi}^*, \hat{\eta}_i^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i|\hat{\phi}^*, \hat{\eta}_i^*)} \right) \\ + \sum_{i=1}^n \left(\sum_{r=1}^t \hat{\phi}_r^* - \hat{\phi}^* \right) \left(\frac{\hat{\alpha}^* \mathbb{P}_1(z_i|\hat{\phi}^*, \hat{\eta}_i^*)}{\mathbb{P}(z_i|\hat{\phi}^*, \hat{\eta}_i^*)} \right) \times \\ \left[\frac{1}{2\hat{\theta}_r^*} - \left[\ln \left(\left(\frac{z_i}{m} \right)^{z_i} \left(1 - \frac{z_i}{m} \right)^{m-z_i} \right) - z_i \hat{\eta}_{ri}^* + m \ln \left(1 + e^{\hat{\eta}_{ri}^*} \right) \right] \right] \\ + \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\beta}_{2r}^* - \hat{\beta}^*) \left(\frac{(\hat{\alpha}^*/t)\mathbb{P}_1(z_i|\hat{\phi}^*, \hat{\eta}_i^*)}{\mathbb{P}(z_i|\hat{\phi}^*, \hat{\eta}_i^*)} \right) \left(\frac{e^{\hat{\phi}^*}}{1+e^{\hat{\phi}^*}} \right) \left(z_i - \frac{me^{\hat{\eta}_i^*}}{1+e^{\hat{\eta}_i^*}} \right)$$

com $\hat{\beta}^*$ e $\hat{\eta}_i^*$ a estimativa do vetor de parâmetros de regressão e do preditor linear para a i -ésima observação amostral sob o modelo homogêneo.

No modelo binomial multiplicativo (seção 3.2), o teste gradiente para verificar a homogeneidade entre avaliadores (ou seja, a inexistência de diferentes tipos de avaliadores) consiste em verificar se de regressão $\beta_r = \beta$, $r = 1, 2, \dots, t$, $\alpha_1 = \alpha_2 = \dots = \alpha_t = \alpha/t$ e $\phi_1 = \phi_2 = \dots = \phi_t = \phi/t$. A estatística do teste segue uma distribuição qui-quadrado com $(t-1)(p-1)$ graus de liberdade, em que p é o número de parâmetros de regressão para cada componente, e é aproximada por

$$\begin{aligned}
S \approx & \sum_{i=1}^n \left(\sum_{r=1}^t \hat{\alpha}_r^* - \hat{\alpha}^* \right) \left(\frac{\mathbb{P}_1(z_i | \hat{\phi}^*, \hat{\eta}_i^*) - \mathbb{P}_0(z_i)}{\mathbb{P}(z_i | \hat{\phi}^*, \hat{\eta}_i^*)} \right) \\
& + \sum_{i=1}^n \left(\sum_{r=1}^t \hat{\phi}_r^* - \hat{\phi}^* \right) \left(\frac{\hat{\alpha}^* \mathbb{P}_1(z_i | \hat{\phi}^*, \hat{\eta}_i^*)}{\mathbb{P}(z_i | \hat{\phi}^*, \hat{\eta}_i^*)} \right) \times \\
& \quad \left(z_i(m - z_i) - \mathbb{E}(Z(m - Z) | \hat{\phi}^*, \hat{\eta}_i^*) \right) \\
& + \sum_{i=1}^n \sum_{r=1}^t \mathbf{x}'_i (\hat{\beta}_{2r}^* - \hat{\beta}^*) \left(\frac{(\hat{\alpha}^*/t) \mathbb{P}_1(z_i | \hat{\phi}^*, \hat{\eta}_i^*)}{\mathbb{P}(z_i | \hat{\phi}^*, \hat{\eta}_i^*)} \right) (z_i - \mathbb{E}(Z | \hat{\phi}^*, \hat{\eta}_i^*))
\end{aligned}$$

com $\hat{\beta}^*$ e $\hat{\eta}_i^*$ a estimativa do vetor de parâmetros de regressão e do preditor linear para a i -ésima observação amostral sob o modelo homogêneo.

5 Aplicação dos modelos

5.1 Estudos de simulação

Modelo binomial redimensionado e deslocado sem inflação (sem covariáveis)

O modelo binomial redimensionado e deslocado sem inflação (seção 2.2) foi avaliado quando ao efeito do tamanho amostral, do tamanho da escala de notas (quantidade de notas distintas) e da distância entre os preditores lineares dos componentes do modelo na estimação dos parâmetros de mistura e de regressão do modelo. Em uma simulação inicial, foram consideradas 500 amostras na proporção 40%, 40% e 20% da mistura de um componente binomial com preditor $\eta_1 = -0.5$, um componente com preditor η_2 variando em 0.50 e 0.75, e um componente uniforme discreto. Para todos os modelos foram considerados $d = 1$, $s = 2$ e o tamanho da escala de notas m assumindo os valores 10 ou 20. As amostras simuladas foram de tamanho $n = 100$ ou $n = 500$.

A Figura 1 mostra o efeito do preditor linear dos componente binomiais nas estimativas dos parâmetros de mistura do modelo. Apenas o preditor linear do segundo componente foi alterado, partindo de $\eta_2 = 0.50$ para $\eta_2 = 0.75$. A linha azul que aparece nos gráficos (às vezes sobreposta pela linha vermelha) é o valor do parâmetro fixado para a simulação (0.40), a linha vermelha contínua é a mediana das estimativas e as linhas tracejadas são, respectivamente, os quantis de ordem 2.5% e 97.5%. Observa-se que a mediana das estimativas dos coeficientes de mistura praticamente coincidem com o “valor real” do parâmetro e que a amplitude do intervalo entre os quantis se torna maior com uma maior distância entre os preditores lineares dos componentes. Isto implica que tal distância influencia na amplitude das estimativas do parâmetro de mistura do modelo.

A Figura 2 mostra o efeito do tamanho amostral nas estimativas dos parâmetros de mistura do modelo. Observa-se que a mediana das estimativas dos coeficientes de mistura praticamente coincidem com o “valor real” do parâmetro e que a amplitude do intervalo entre os quantis se torna menor, como é esperado, conforme o tamanho amostral aumenta.

A Figura 3 mostra o efeito do tamanho amostral nas estimativas dos parâmetros de mistura do modelo. Observa-se que, como ocorre com o aumento do tamanho amostral, a mediana das estimativas dos coeficientes de mistura praticamente coincidem com o “valor real” do parâmetro e a amplitude do intervalo entre os quantis se torna menor conforme o tamanho da escala de notas aumenta.

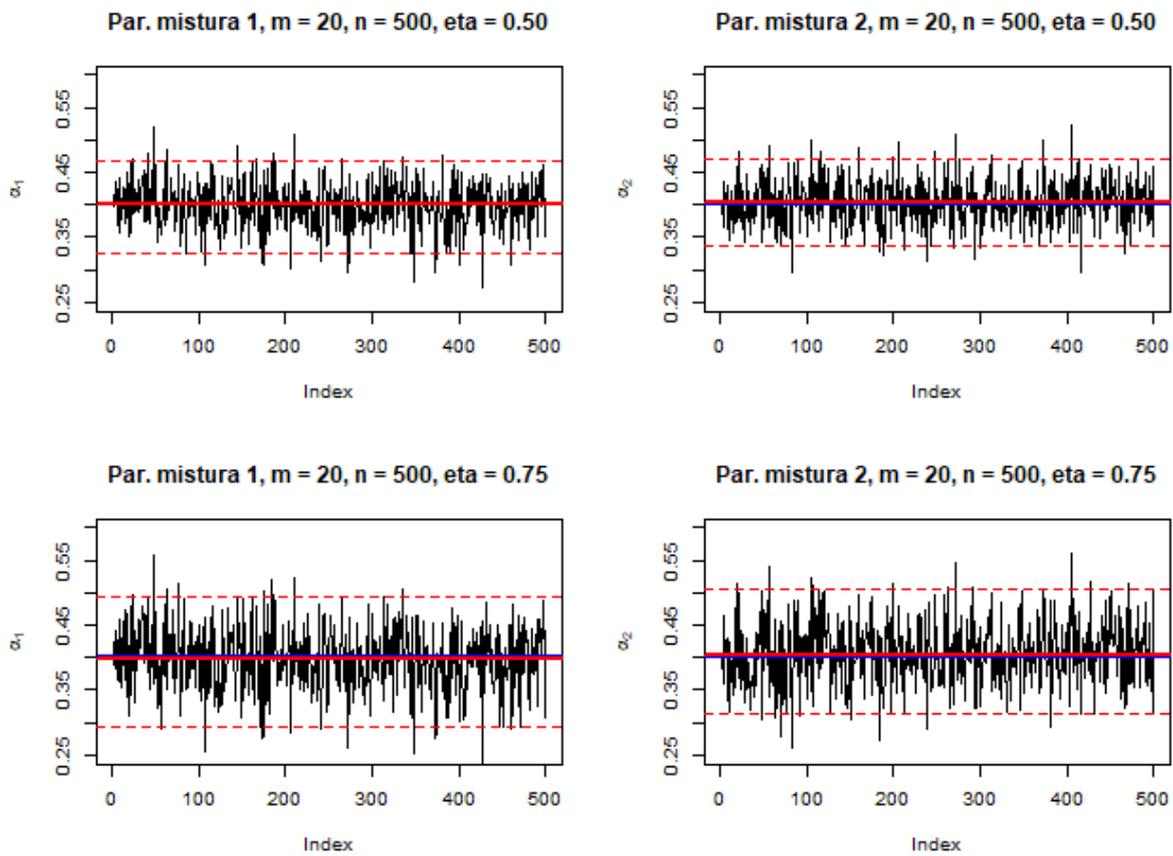


Figura 1 – Estimativas dos parâmetros de mistura para 500 amostras com escala de 20 notas, amostra de tamanho 500 e preditor linear do segundo componente binomial variando entre 0.50 e 0.75 – efeito do preditor linear no modelo binomial redimensionado e deslocado.

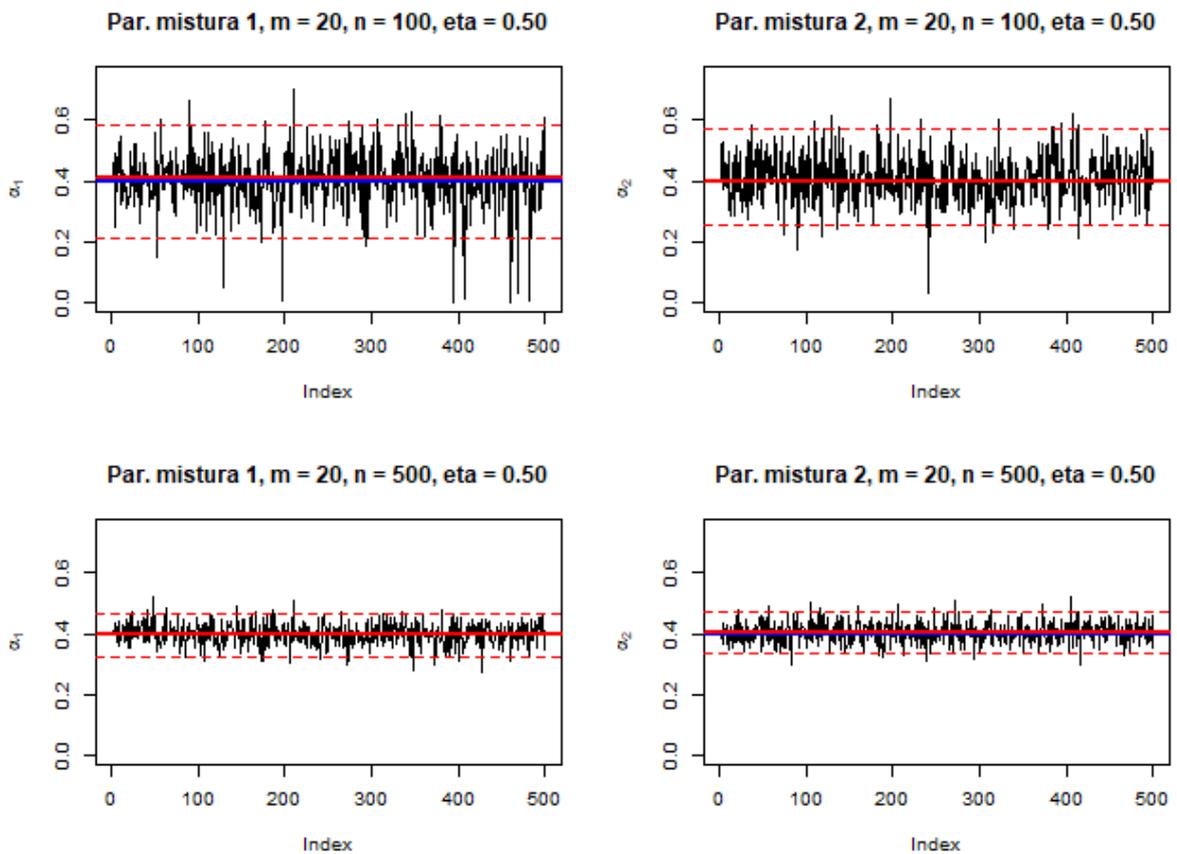


Figura 2 – Estimativas dos parâmetros de mistura para 500 amostras com escala de 20 notas, tamanho de amostra variando entre 100 e 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho amostral no modelo binomial redimensionado e deslocado.

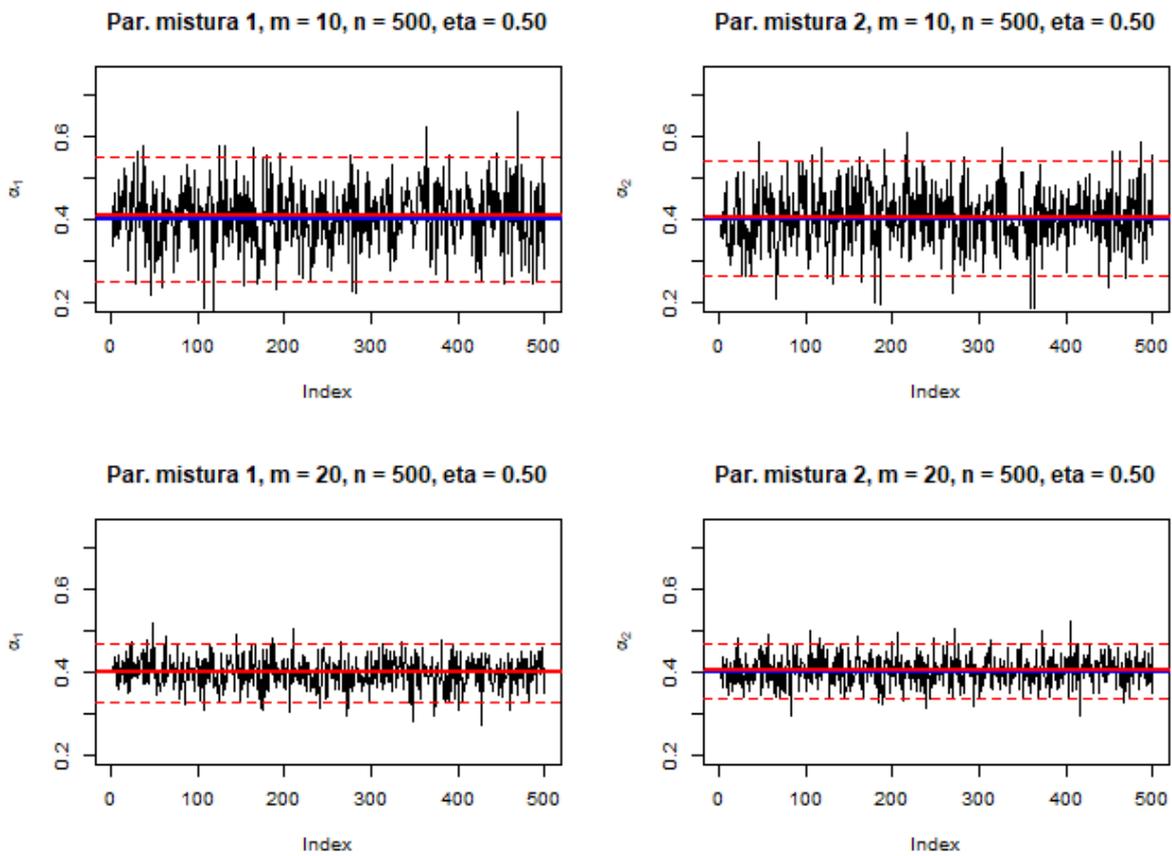


Figura 3 – Estimativas dos parâmetros de mistura para 500 amostras com escala de notas variando entre 10 e 20, tamanho de amostra igual a 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho da escala de notas no modelo binomial redimensionado e deslocado.

A Figura 4 mostra o efeito do preditor linear dos componente binomiais nas estimativas dos parâmetros de regressão do modelo. Apenas o preditor linear do segundo componente foi alterado, partindo de $\eta_2 = 0.50$ para $\eta_2 = 0.75$. Em ambos os cenários a estimativa do coeficiente do intercepto do primeiro componente se manteve centrado no “valor real” do parâmetro, sofrendo um leve aumento de amplitude com o aumento da distância entre os componentes. Por outro lado, a estimativa do coeficiente do intercepto do segundo componente se mostrou bastante superestimado no cenário com os componentes mais próximos.

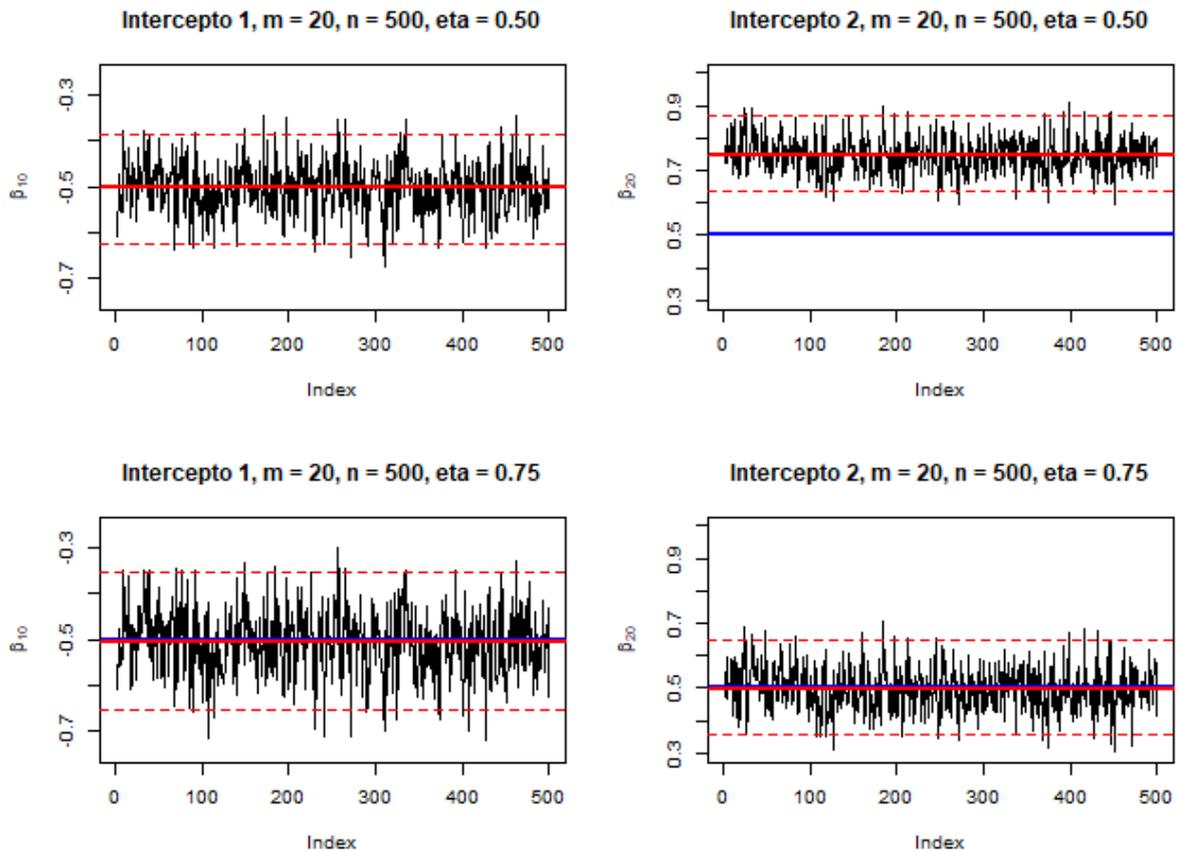


Figura 4 – Estimativas dos parâmetros de regressão para 500 amostras com escala de 20 notas, amostra de tamanho 500 e preditor linear do segundo componente binomial variando entre 0.50 e 0.75 – efeito do preditor linear no modelo binomial redimensionado e deslocado.

A Figura 5 mostra o efeito do tamanho amostral na estimativa do parâmetro de mistura do modelo. É interessante observar que as estimativas do coeficiente do intercepto sofre uma degradação para algumas amostras no cenário com escala de 20 notas e amostra de tamanho 100. Isto mostra que o modelo pode necessitar de amostras grandes em alguns cenários. Observe que, neste caso, se as notas forem distribuídas uniformemente na escala haveria apenas 5 notas para cada valor distinto. A recomendação é utilizar amostras grandes o suficiente para que a proporção entre o tamanho da escala e o tamanho amostral seja razoavelmente grandes.

A Figura 6 mostra o efeito do tamanho amostral na estimativa do parâmetro de regressão do modelo. O parâmetro associado ao intercepto do primeiro componente se manteve centrado

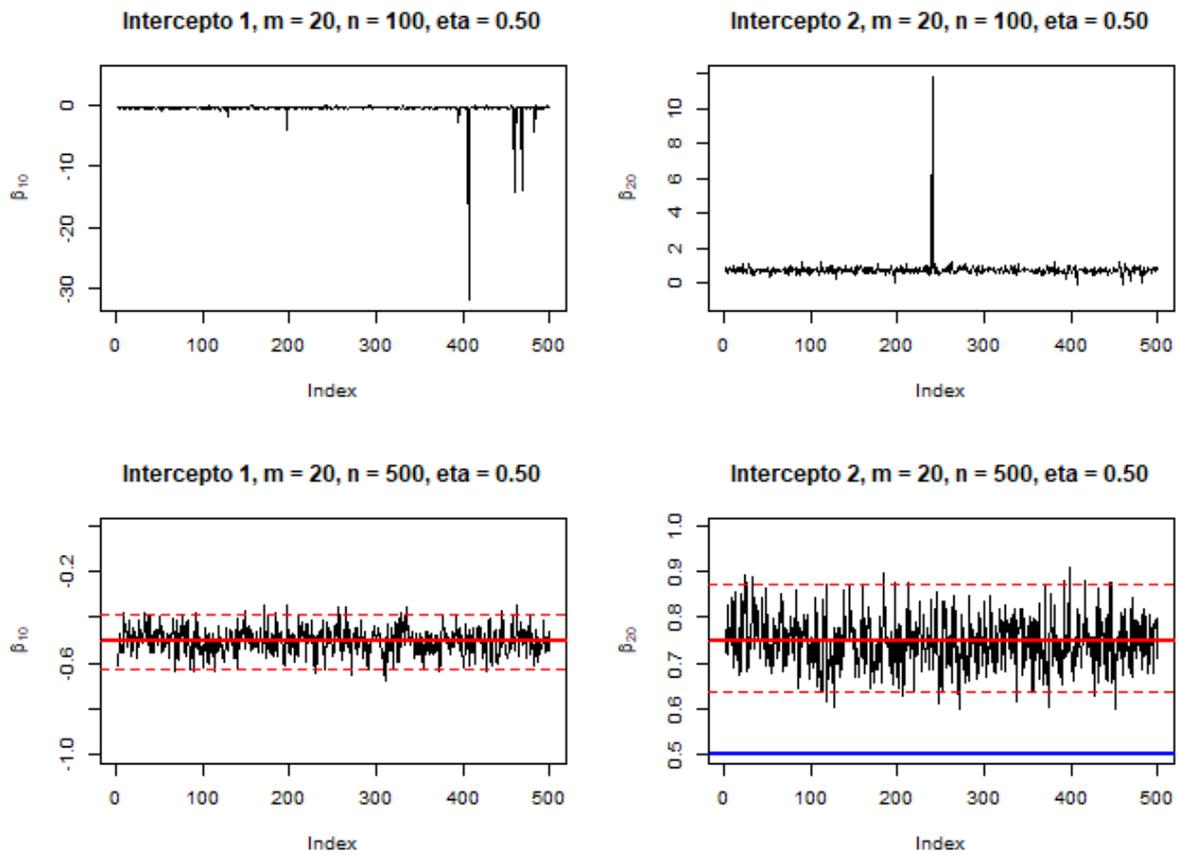


Figura 5 – Estimativas dos parâmetros de regressão para 500 amostras com escala de 20 notas, tamanho de amostra variando entre 100 e 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho amostral no modelo binomial redimensionado e deslocado.

no “valor real” do parâmetro em ambos os cenários e, também, se mostrou superestimado em ambos os cenários. Observa-se que, a amplitude do intervalo entre os quantis se torna menor conforme o tamanho da escala de notas aumenta.

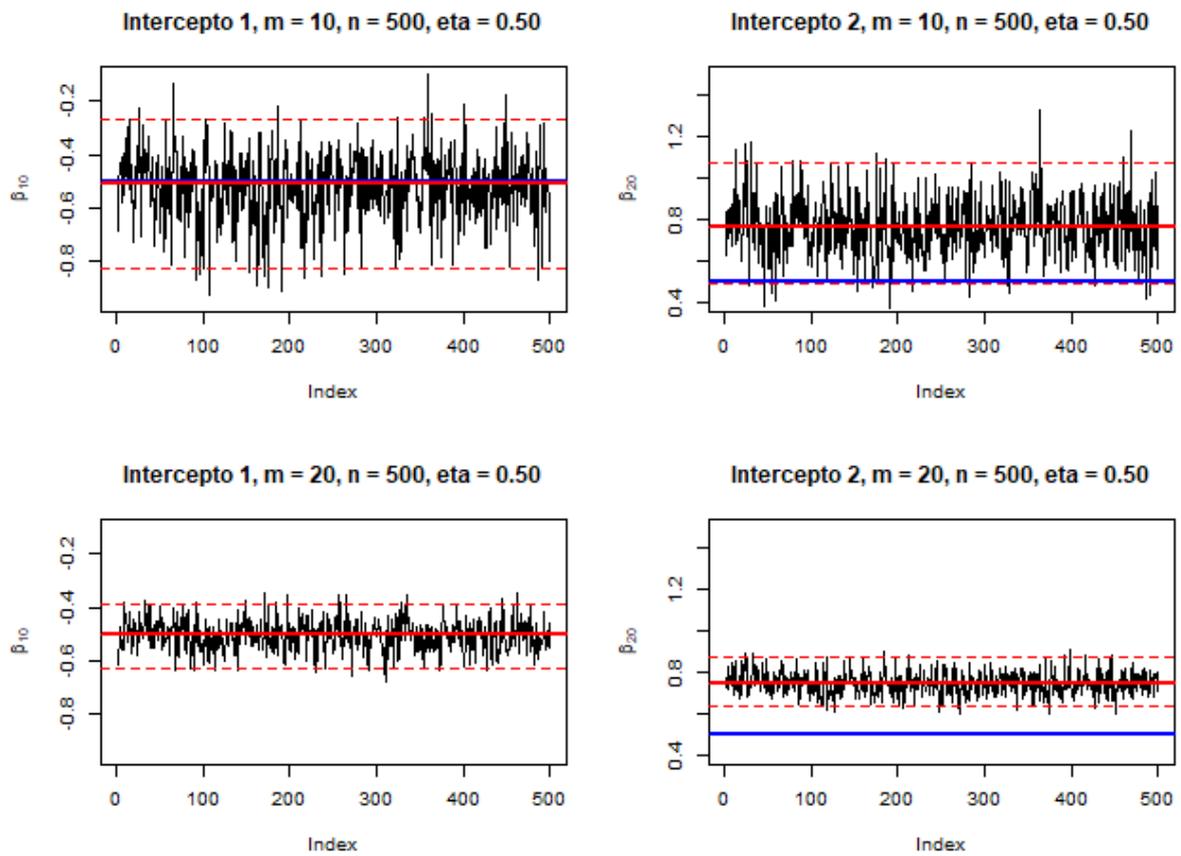


Figura 6 – Estimativas dos parâmetros de mistura para 500 amostras com escala de notas variando entre 10 e 20, tamanho de amostra igual a 500 e preditor linear do segundo componente igual a 0.50 – efeito do tamanho da escala de notas no modelo binomial redimensionado e deslocado.

Modelo binomial redimensionado e deslocado sem inflação (com covariáveis)

Em um segundo estudo, será avaliado o comportamento do modelo com a presença de covariáveis. Para isto, foram consideradas duas covariáveis na construção das amostras. A primeira covariável (X_1) segue uma distribuição de Bernoulli com parâmetro $p = 0.45$ e a segunda covariável (X_2) segue uma distribuição uniforme contínua entre -1.0 e 1.0 .

Os componentes binomiais redimensionados e deslocados foram todos gerados com os parâmetros $d = 1$ e $s = 2$, variando a quantidade de notas distintas (valor de m) e o tamanho amostral (n), construindo-se 4 cenários distintos: (1) amostras de tamanho $n = 100$ e escalas de $m = 10$ notas distintas, (2) amostras de tamanho $n = 500$ e escalas de $m = 10$ notas distintas, (3) amostras de tamanho $n = 100$ e escalas de $m = 20$ notas distintas e (4) amostras de tamanho $n = 500$ e escalas de $m = 20$ notas distintas. Em todos os cenários foram considerados as misturas de distribuições com os preditores lineares dados por $\eta_1 = -1.0 + 0.50X_1$ e $\eta_2 = 1.5 + 0.25X_1 + 0.75X_2$. A Tabela 1 apresenta a relação entre o preditor linear e o parâmetro π (na binomial padrão é a probabilidade dos lançamentos de Bernoulli) correspondente aos componentes da mistura para diferentes valores das covariáveis. Observe que as misturas são compostas de componentes não sobrepostos.

X_1	X_2	η_1	π_1	η_2	π_2
0	-1	-1.00	0.2689	0.75	0.6792
0	0	-1.00	0.2689	1.50	0.8175
0	1	-1.00	0.2689	2.25	0.9047
1	-1	-0.50	0.3775	1.00	0.7311
1	0	-0.50	0.3775	1.75	0.8520
1	1	-0.50	0.3775	2.50	0.9241

Tabela 1 – Preditor linear e probabilidade das distribuições binomiais redimensionadas e deslocadas – simulação do modelo sem inflação.

Foram simuladas 500 amostras para cada cenário, com mistura na seguinte proporção: 40% dos valores gerados pela distribuição binomial redimensionada e deslocada com preditor linear η_1 , 40% dos valores gerados pela distribuição binomial redimensionada e deslocada com preditor linear η_2 e 20% dos valores gerados pela distribuição uniforme discreta.

A Figura 7 mostra o efeito do tamanho amostral na estimativa dos parâmetros de mistura do modelo binomial redimensionado e deslocado. Como antes, a linha azul representa o valor “real” do parâmetro, a linha vermelha contínua representa a mediana das estimativas e as linhas tracejadas representam, respectivamente, os quantis de ordem 2.5% e 97.5% das estimativas.

Para ambos os tamanhos amostrais as estimativas dos parâmetros estão centralizadas no “valor real” do parâmetro, sofrendo as estimativas uma redução de amplitude com o aumento do tamanho amostral.

As Figuras 8 e 9 apresentam o efeito do tamanho amostral para a estimação dos parâmetros de regressão dos dois componentes da mistura. No primeiro componente, observou-se que as

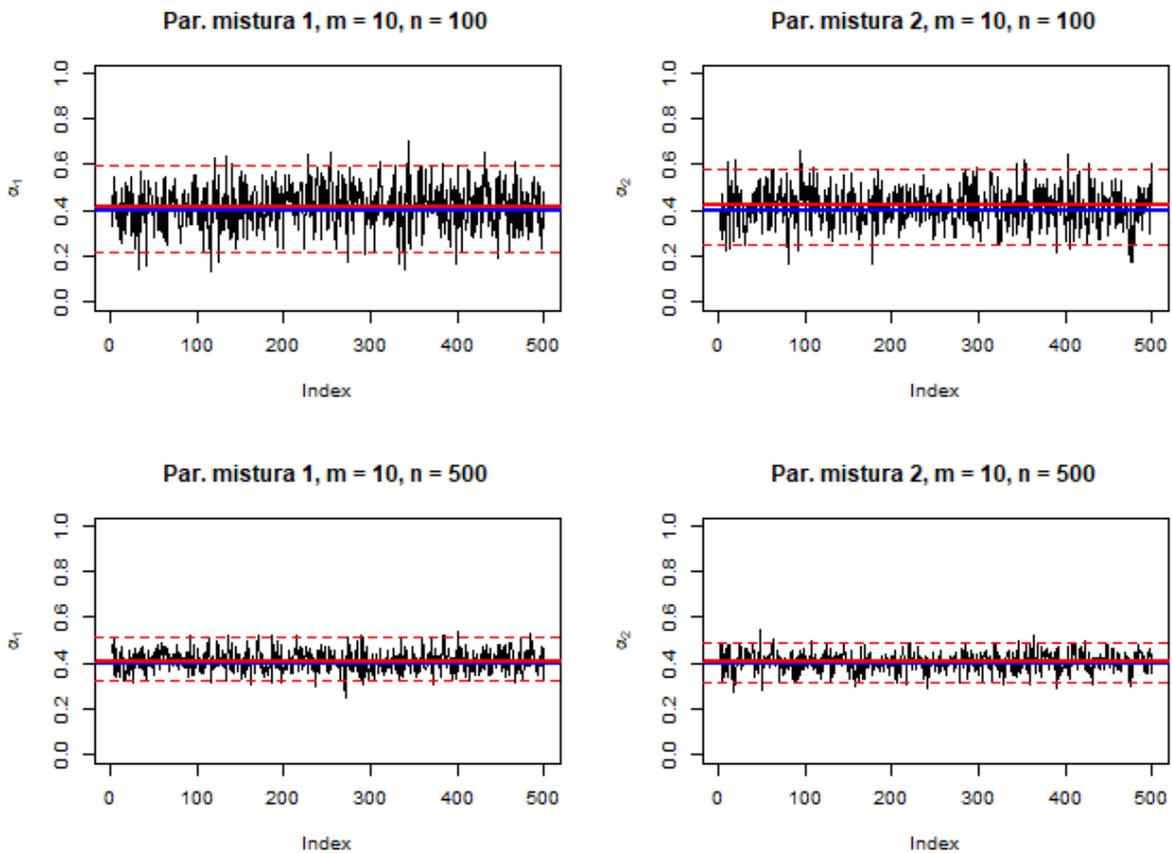


Figura 7 – Estimativas dos parâmetros de mistura para 500 amostras com escala de 10 notas – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.

estimativas dos parâmetros de regressão relacionados com a covariável dicotômica (X_1) e com a covariável contínua (X_2) se concentraram acima do “valor real” mantendo a banda que compreende os quantis de ordem 2.5% e 97.5% acima de tal valor. A banda formada pelos quartis de ordem 2.5% e 97.5% contém, para a amostra de tamanho 100, apenas o coeficiente relacionado com a variável contínua (X_2). No entanto, a superestimação dos coeficientes é corrigida com o aumento do tamanho amostral.

No segundo componente, observa-se alguma instabilidade nas estimativas do intercepto e do coeficiente associado com a variável contínua (X_2) no cenário de amostra tamanho 100. Com a amostra maior, tamanho 500, tal efeito não ocorre. Para ambos os tamanhos amostrais, no segundo componente, o coeficiente relativo à variável dicotômica (X_1) apresenta uma pequena subestimação.

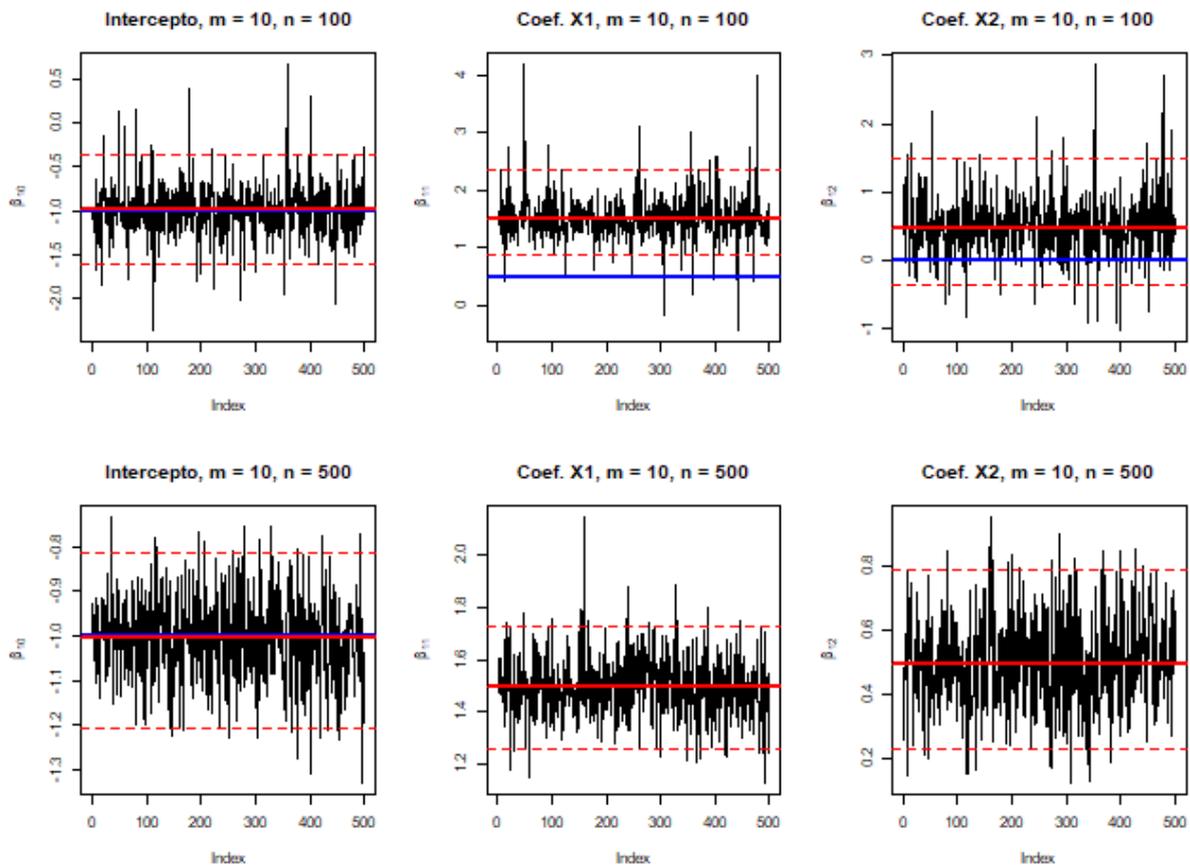


Figura 8 – Estimativas dos parâmetros de regressão do primeiro componente da mistura para 500 amostras com escala de 10 notas – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.

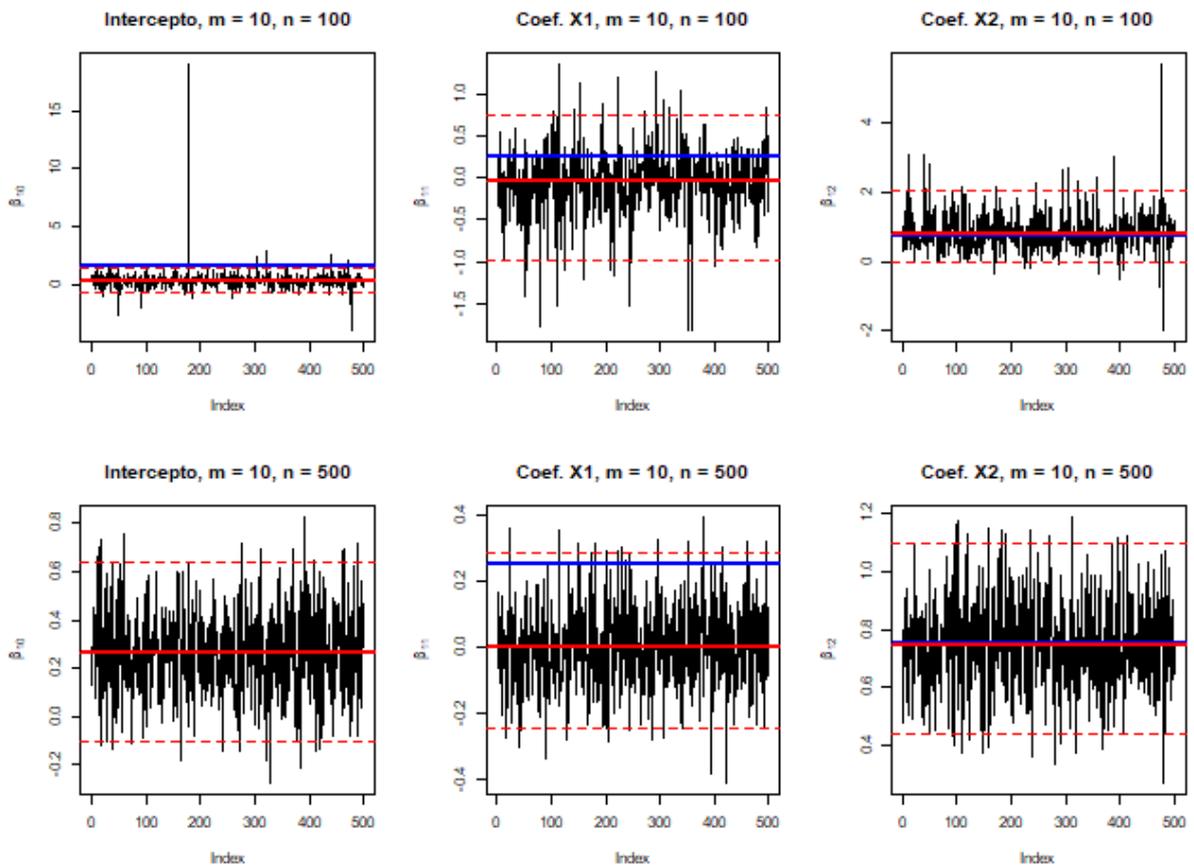


Figura 9 – Estimativas dos parâmetros de regressão do segundo componente da mistura para 500 amostras com escala de 10 notas – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.

A Figura 10 mostra o efeito do tamanho da escala de notas na estimativa dos parâmetros de mistura do modelo binomial redimensionado e deslocado. Como na simulação anterior, as estimativas estão centradas no “valor real” dos parâmetros e apresentam uma diminuição de amplitude conforme o tamanho da escala aumenta.

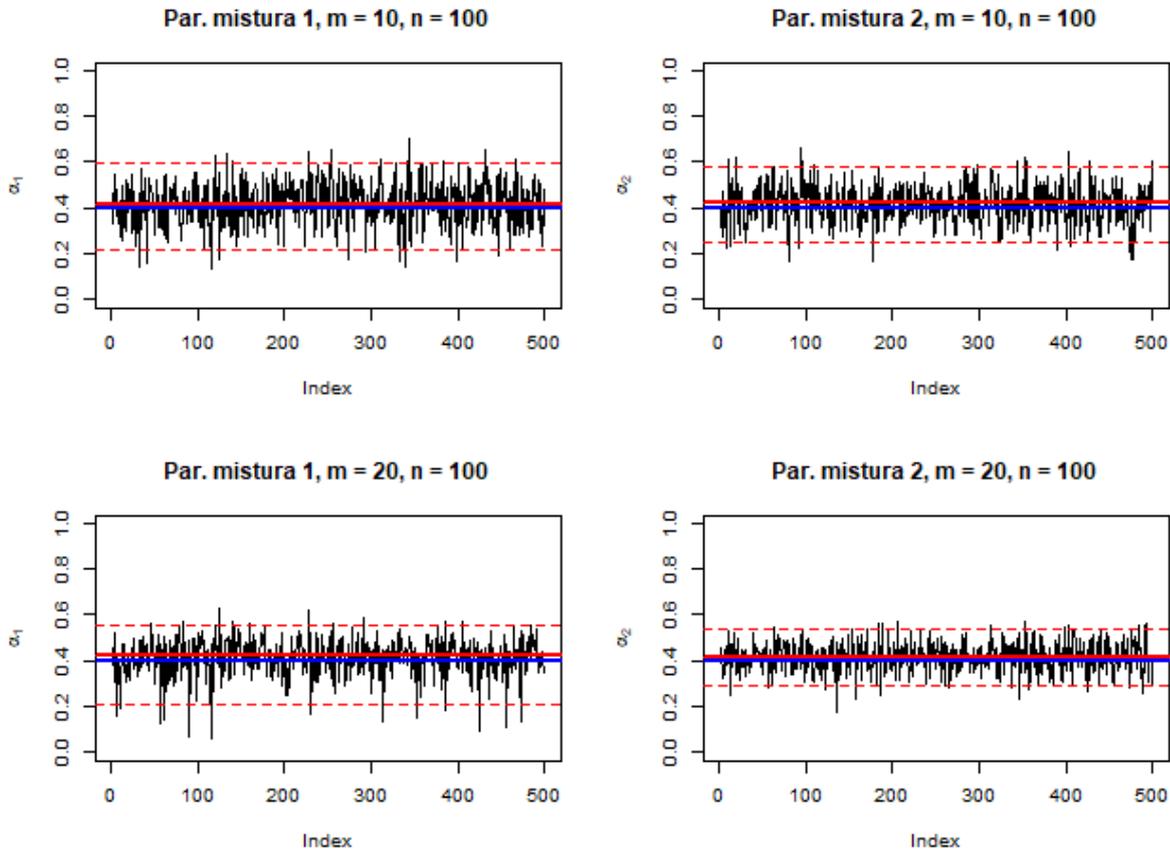


Figura 10 – Estimativas dos parâmetros de mistura para 500 amostras de tamanho 100 – efeito do tamanho da escala de notas para o modelo binomial redimensionado e deslocado.

As Figuras 11 e 12 apresentam o efeito do tamanho da escala de notas para a estimação dos parâmetros de regressão dos dois componentes da mistura. A estimativa do intercepto do modelo apresenta uma degradação no segundo componente quando a escala é de $m = 10$ notas e o tamanho amostral é $n = 100$. Observe que, neste caso, existem 10 observações amostrais para cada valor da escala. Novamente, como observado na simulação anterior, se tal relação é pequena os coeficientes de regressão podem sofrer tal degradação. Esta degradação também é observada no componente relativo à variável contínua (X_2).

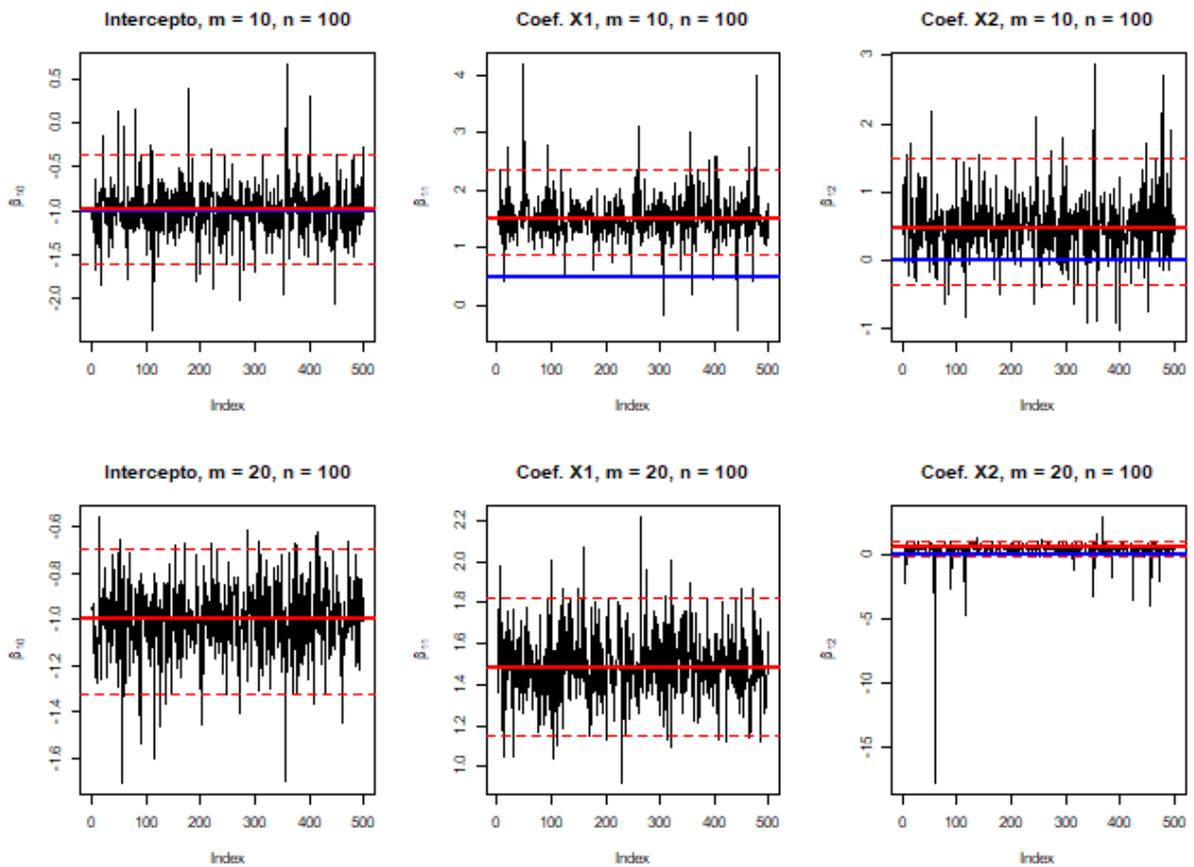


Figura 11 – Estimativas dos parâmetros de regressão do primeiro componente da mistura para 500 amostras de tamanho 100 – efeito do tamanho da escala de notas para o modelo binomial redimensionado e deslocado.

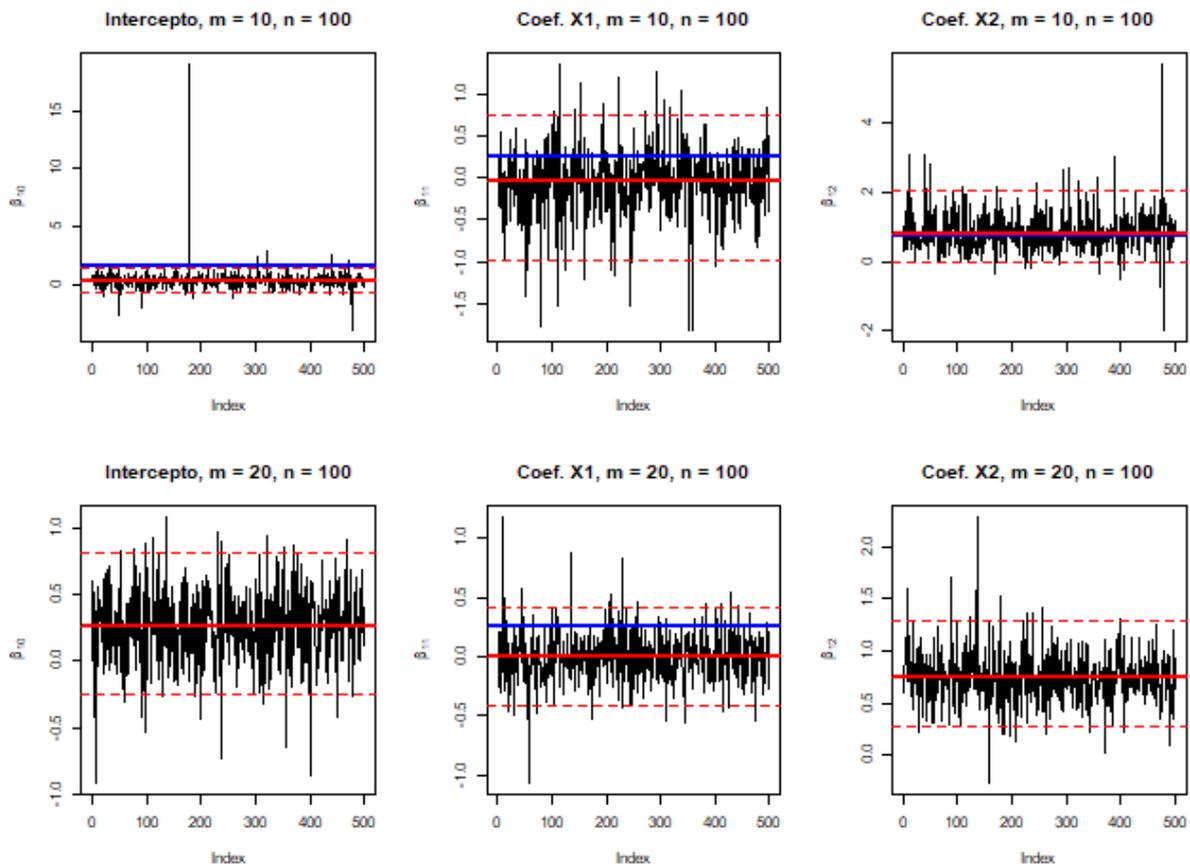


Figura 12 – Estimativas dos parâmetros de regressão do segundo componente da mistura para 500 amostras de tamanho 100 notas – efeito do tamanho escala de notas para o modelo binomial redimensionado e deslocado.

Modelo binomial redimensionado e deslocado com inflação

As Figuras 13, 14 e 15 apresentam as estimativas dos parâmetros de mistura referente ao componente uniforme discreto e degenerado (inflação) para o modelo binomial redimensionado e deslocado com inflação (seção 2.3). Para a simulação do modelo foram geradas 500 amostras com $n = 100$ e $n = 500$ do modelo com os componentes sem covariáveis com preditores lineares dados por $\eta_1 = -1.0$ e $\eta_2 = 1.0$. O objetivo é identificar a relação entre os coeficientes de mistura associado aos termos não binomiais.

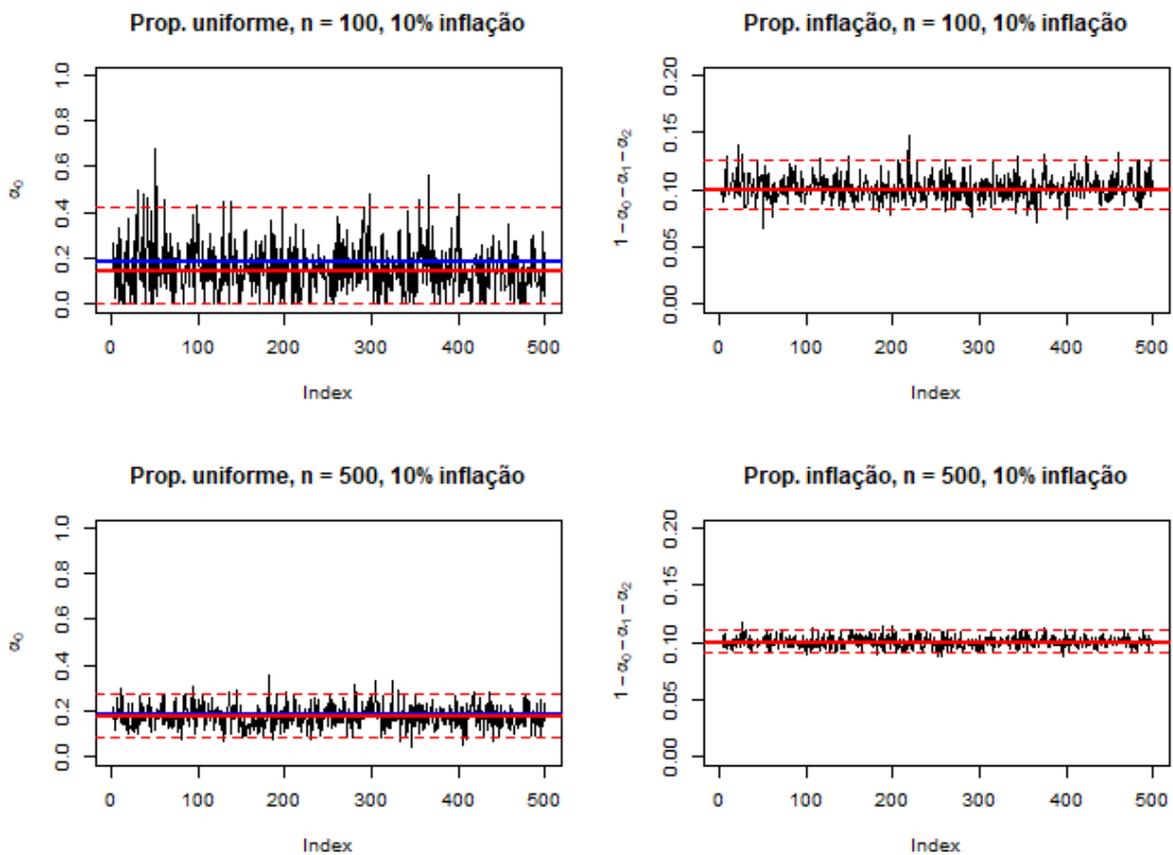


Figura 13 – Estimativas dos parâmetros de mistura, referente ao termo de avaliação ao acaso e inflacionada, para 500 amostras com 10% de inflação na maior nota da escala – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.

Observa-se que o coeficiente associado ao termo de avaliação ao acaso (componente uniforme discreto) tende a ser levemente subestimado, independente da proporção de inflação. A proporção de inflação têm as estimativas centradas no “valor real” do parâmetro para ambos os tamanhos amostrais. O aumento do tamanho amostral, como esperado, causa uma diminuição na amplitude das estimativas.

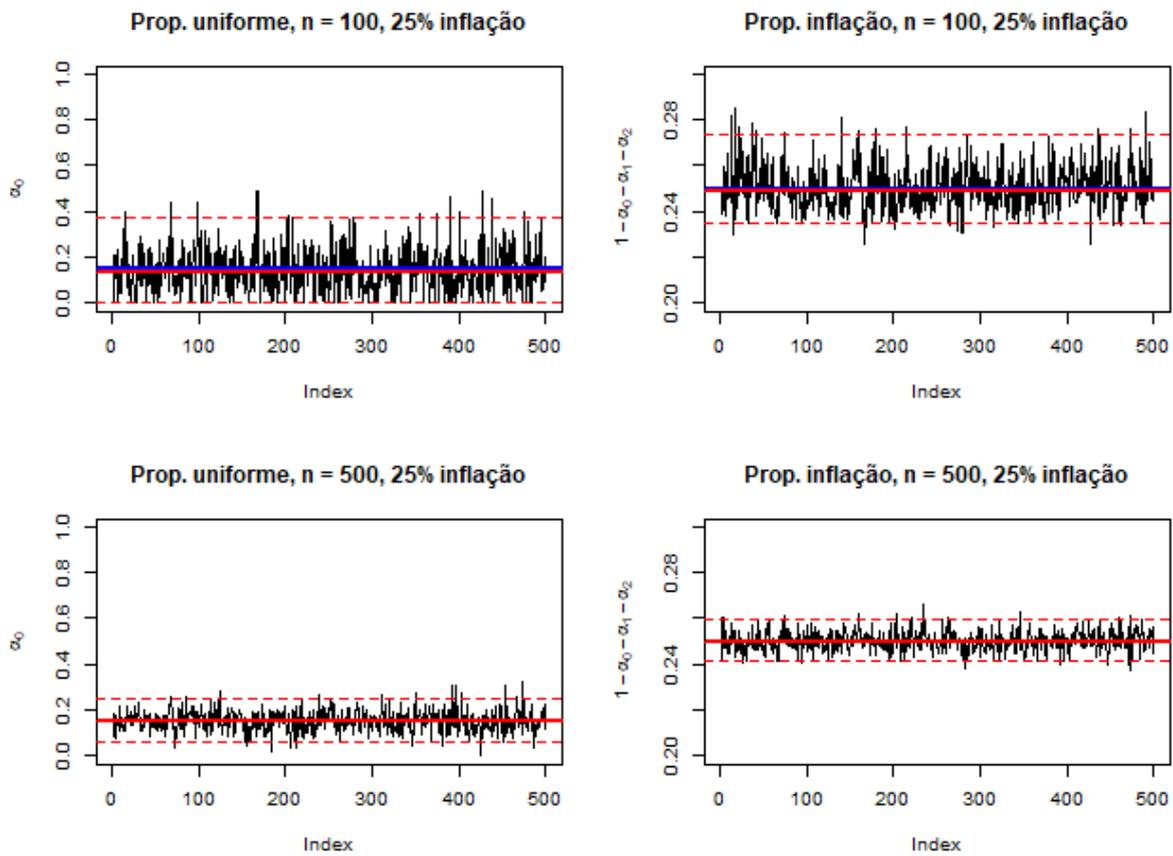


Figura 14 – Estimativas dos parâmetros de mistura, referente ao termo de avaliação ao acaso e inflacionada, para 500 amostras com 25% de inflação na maior nota da escala – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.

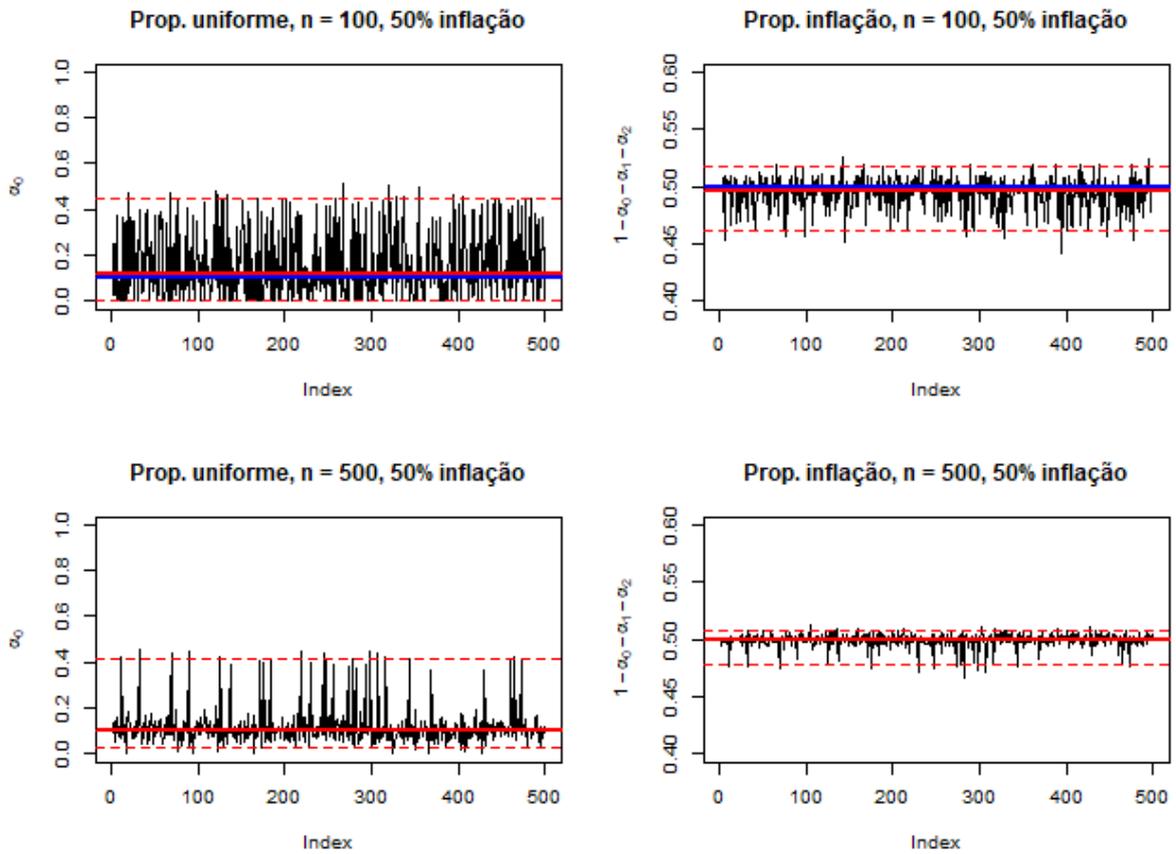


Figura 15 – Estimativas dos parâmetros de mistura, referente ao termo de avaliação ao acaso e inflacionada, para 500 amostras com 50% de inflação na maior nota da escala – efeito do tamanho amostral para o modelo binomial redimensionado e deslocado.

Modelo binomial duplo

Para avaliar o modelo binomial duplo foram simulados 2 cenários: no primeiro cenário o modelo é composto por um componente binomial usual, ou seja, binomial duplo com $\theta = 1$ com parâmetros $m = 25$ e $\pi = 0.25$ e um componente binomial duplo com parâmetros $m = 25$, $\pi = 0.75$ e $\theta = 0.50$; no segundo cenário o modelo é composto por um componente binomial usual com parâmetros $m = 25$ e $\pi = 0.25$ e um componente binomial duplo com parâmetros $m = 25$, $\pi = 0.75$ e $\theta = 0.75$. O parâmetro $\theta < 1.0$ faz com que as extremidades da distribuição tenha maior probabilidade em detrimento dos valores no centro.

Para cada cenário foram geradas 100 amostras de tamanho $n = 500$. Os parâmetros foram escolhidos de forma que a distribuição resultante seja bimodal.

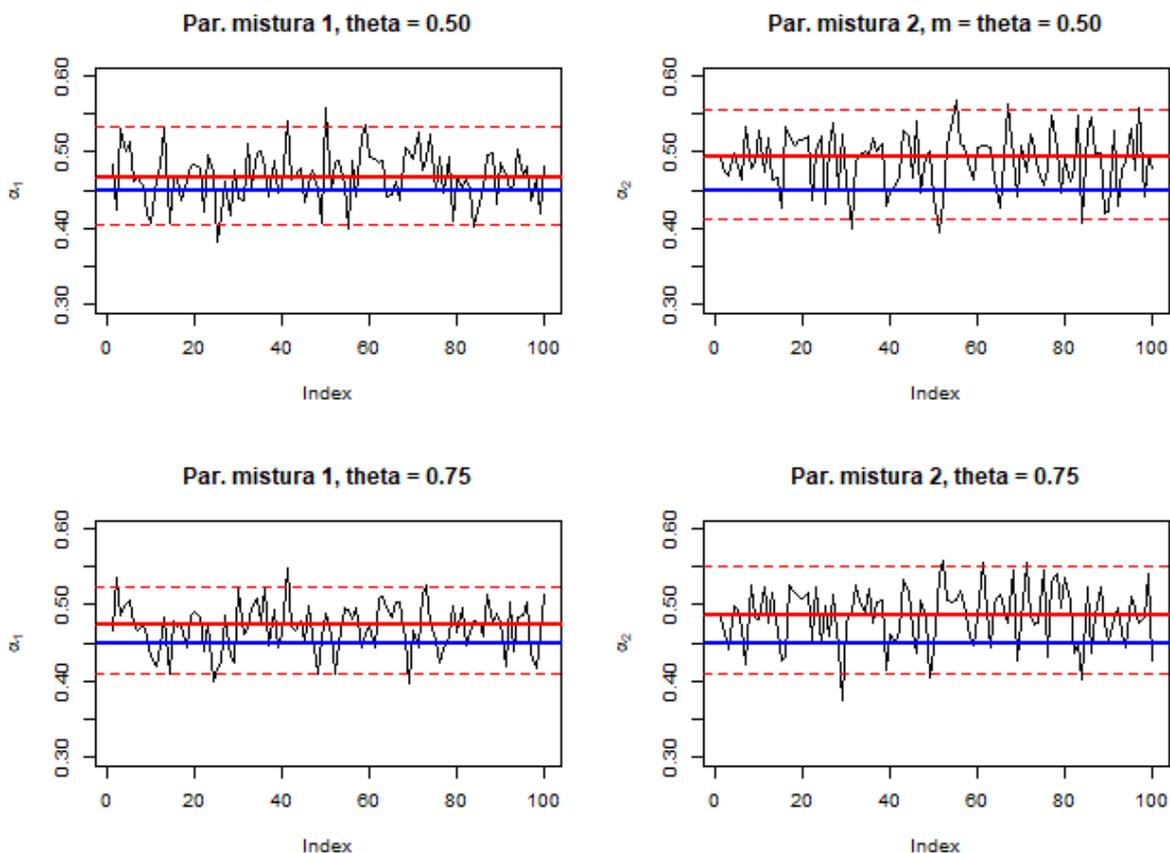


Figura 16 – Estimativas dos parâmetros de mistura, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.

A Figura 19 mostra as estimativas dos parâmetros de mistura para as amostras comparando os 2 cenários. Observa-se que a amplitude das estimativas se mantém constante com o aumento do parâmetro de variância, embora se observe um leve distanciamento na mediana das estimativas para o parâmetro do componente 1 em relação à respectiva mediana, e o efeito contrário no componente 2.

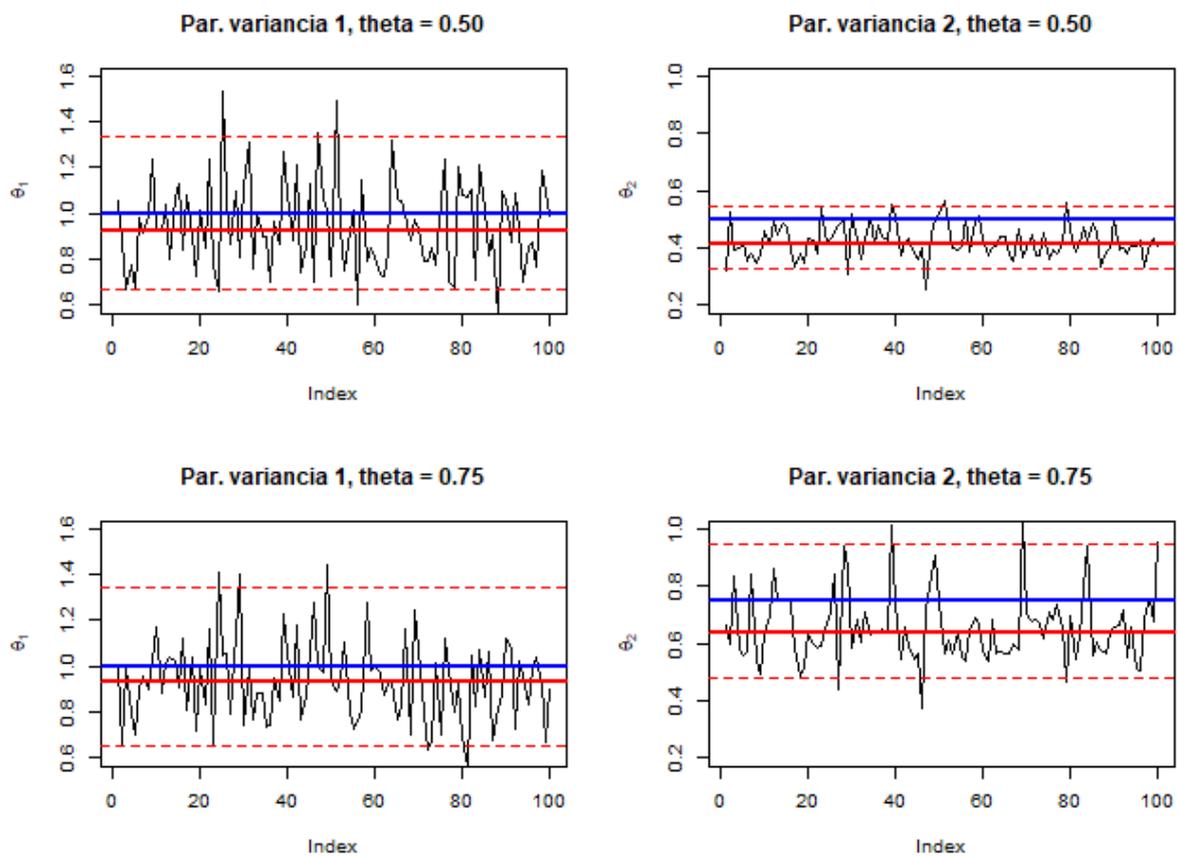


Figura 17 – Estimativas dos parâmetros de variância, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.

A Figura 20 mostra as estimativas dos parâmetros de variância para as amostras comparando os 2 cenários. A amplitude das estimativas do parâmetro para o componente 1 se mostra mais volátil no cenário 1, sofrendo uma redução em sua variabilidade com o aumento do coeficiente de variância do componente 2. As estimativas do parâmetro de variância do segundo componente é mais volátil no segundo componente quando o valor real é $\theta = 0.75$. Não se observa efeitos significativos quanto a mediana dos valores estimados. Em todos os cenários os valores estimados estão levemente abaixo daqueles usados para gerar as amostras.

Por fim, a Figura 21 mostra as estimativas dos interceptos dos componentes para os 2 cenários considerados. Observa-se um aumento de amplitude nas estimativas dos interceptos com relação ao parâmetro de variância dos modelos. Quando o segundo componente apresenta parâmetro $\theta = 0.50$ as estimativas dos interceptos do primeiro componente se mostra com maior variabilidade do que as estimativas correspondentes para $\theta = 0.75$. O efeito observado nas estimativas do intercepto do segundo componente sofre o efeito inverso, ou seja, à medida que o parâmetro de variância aumenta de $\theta = 0.50$ para $\theta = 0.75$ a variabilidade das estimativas aumenta. Não se observa variações significativas quando a mediana das estimativas.

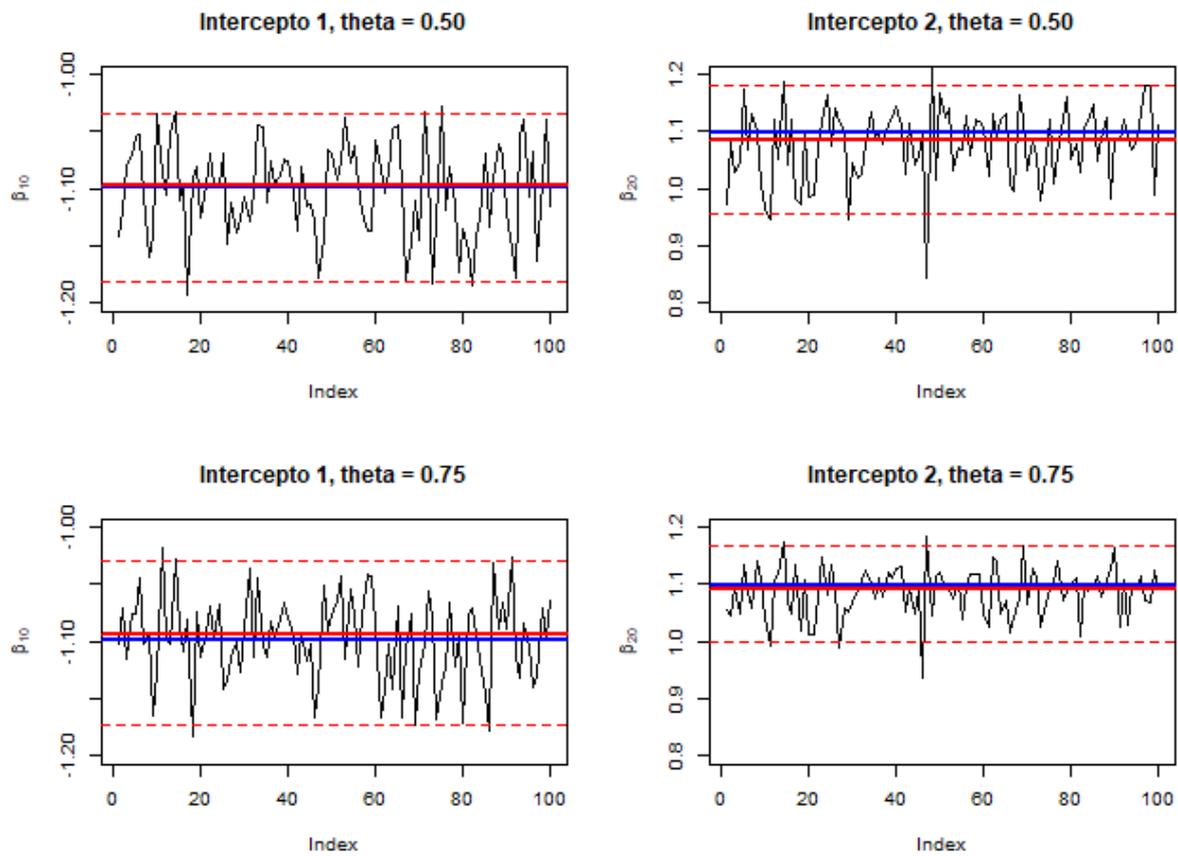


Figura 18 – Estimativas dos parâmetros de regressão, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.

Modelo binomial multiplicativo

Para avaliar o modelo binomial duplo foram simulados 2 cenários: no primeiro cenário o modelo é composto por um componente binomial usual, ou seja, binomial duplo com $\theta = 1$ com parâmetros $m = 25$ e $\pi = 0.25$ e um componente binomial duplo com parâmetros $m = 25$, $\pi = 0.75$ e $\theta = 0.50$; no segundo cenário o modelo é composto por um componente binomial usual com parâmetros $m = 25$ e $\pi = 0.25$ e um componente binomial duplo com parâmetros $m = 25$, $\pi = 0.75$ e $\theta = 0.75$. O parâmetro $\theta < 1.0$ faz com que as extremidades da distribuição tenha maior probabilidade em detrimento dos valores no centro.

Para cada cenário foram geradas 100 amostras de tamanho $n = 500$. Os parâmetros foram escolhidos de forma que a distribuição resultante seja bimodal.

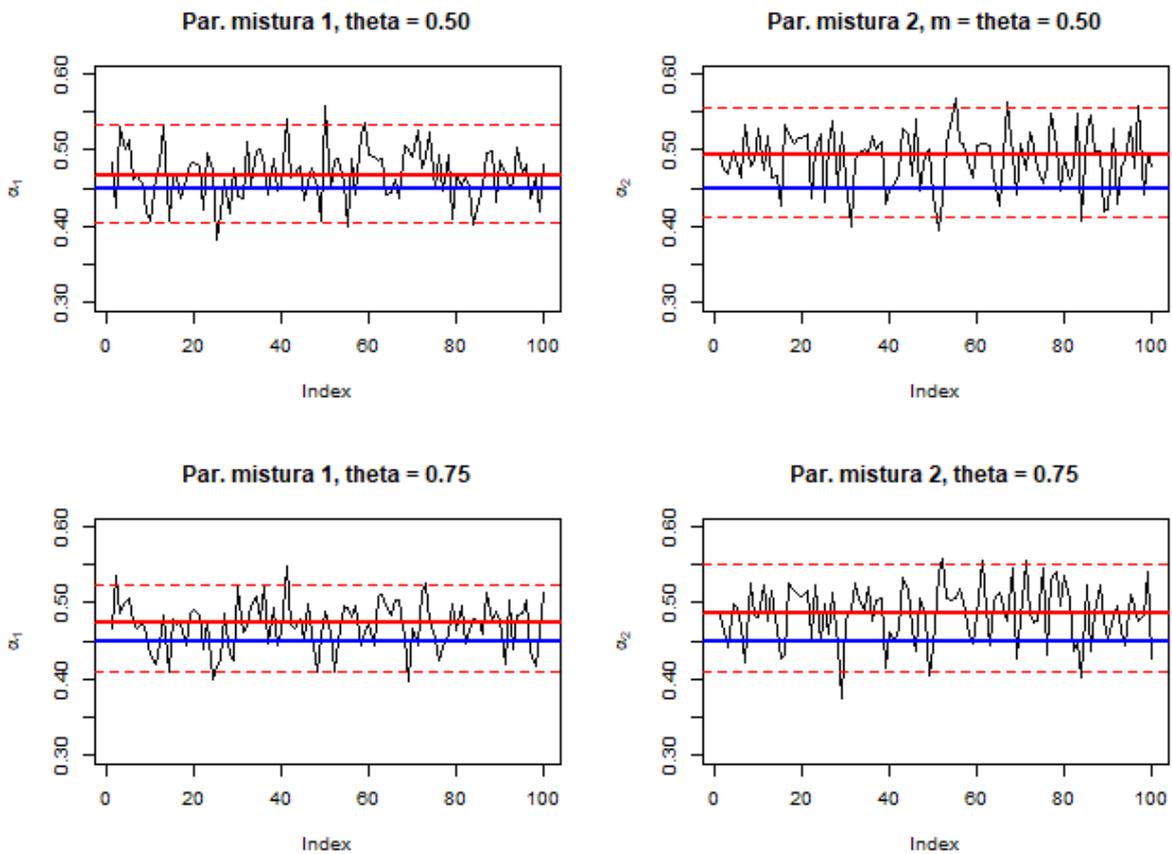


Figura 19 – Estimativas dos parâmetros de mistura, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.

A Figura 19 mostra as estimativas dos parâmetros de mistura para as amostras comparando os 2 cenários. Observa-se que a amplitude das estimativas se mantém constante com o aumento do parâmetro de variância, embora se observe um leve distanciamento na mediana das estimativas para o parâmetro do componente 1 em relação à respectiva mediana, e o efeito contrário no componente 2.

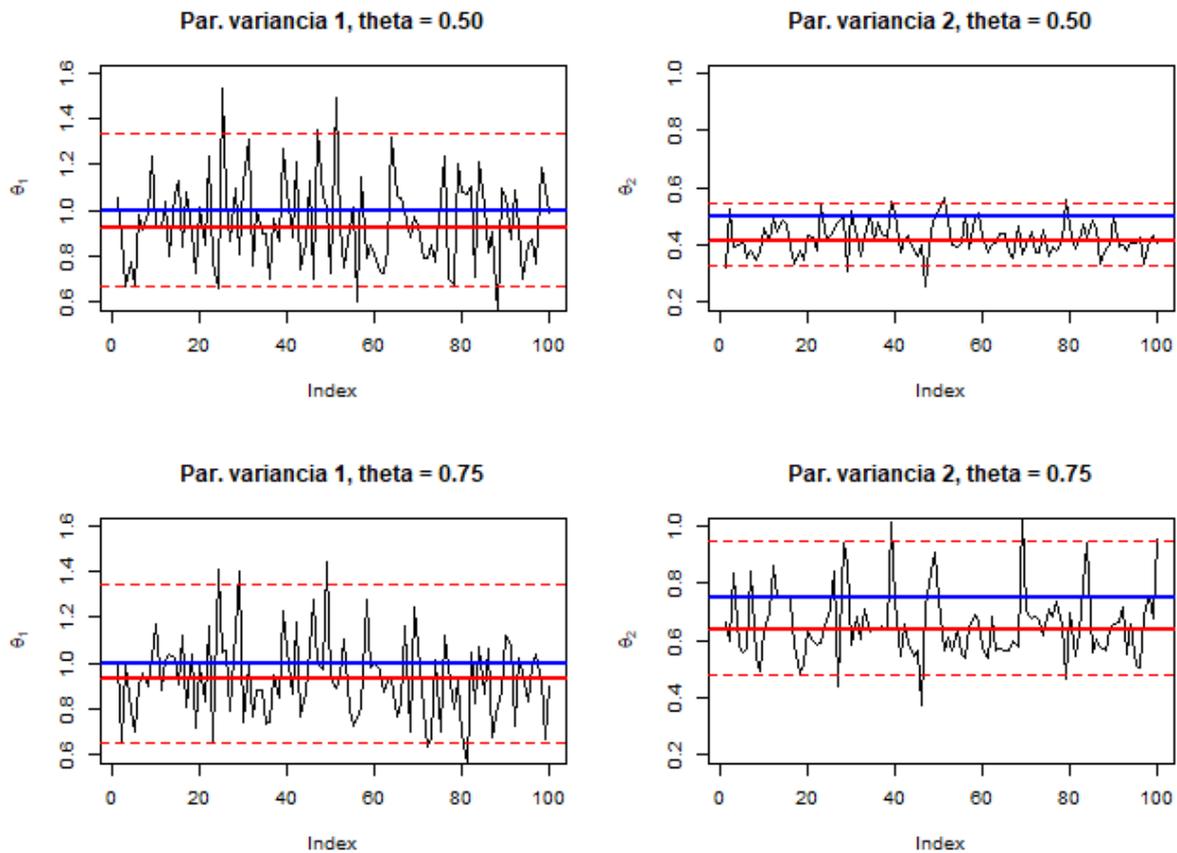


Figura 20 – Estimativas dos parâmetros de variância, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.

A Figura 20 mostra as estimativas dos parâmetros de variância para as amostras comparando os 2 cenários. A amplitude das estimativas do parâmetro para o componente 1 se mostra mais volátil no cenário 1, sofrendo uma redução em sua variabilidade com o aumento do coeficiente de variância do componente 2. As estimativas do parâmetro de variância do segundo componente é mais volátil no segundo componente quando o valor real é $\theta = 0.75$. Não se observa efeitos significativos quanto a mediana dos valores estimados. Em todos os cenários os valores estimados estão levemente abaixo daqueles usados para gerar as amostras.

Por fim, a Figura 21 mostra as estimativas dos interceptos dos componentes para os 2 cenários considerados. Observa-se um aumento de amplitude nas estimativas dos interceptos com relação ao parâmetro de variância dos modelos. Quando o segundo componente apresenta parâmetro $\theta = 0.50$ as estimativas dos interceptos do primeiro componente se mostra com maior variabilidade do que as estimativas correspondentes para $\theta = 0.75$. O efeito observado nas estimativas do intercepto do segundo componente sofre o efeito inverso, ou seja, à medida que o parâmetro de variância aumenta de $\theta = 0.50$ para $\theta = 0.75$ a variabilidade das estimativas aumenta. Não se observa variações significativas quando a mediana das estimativas.

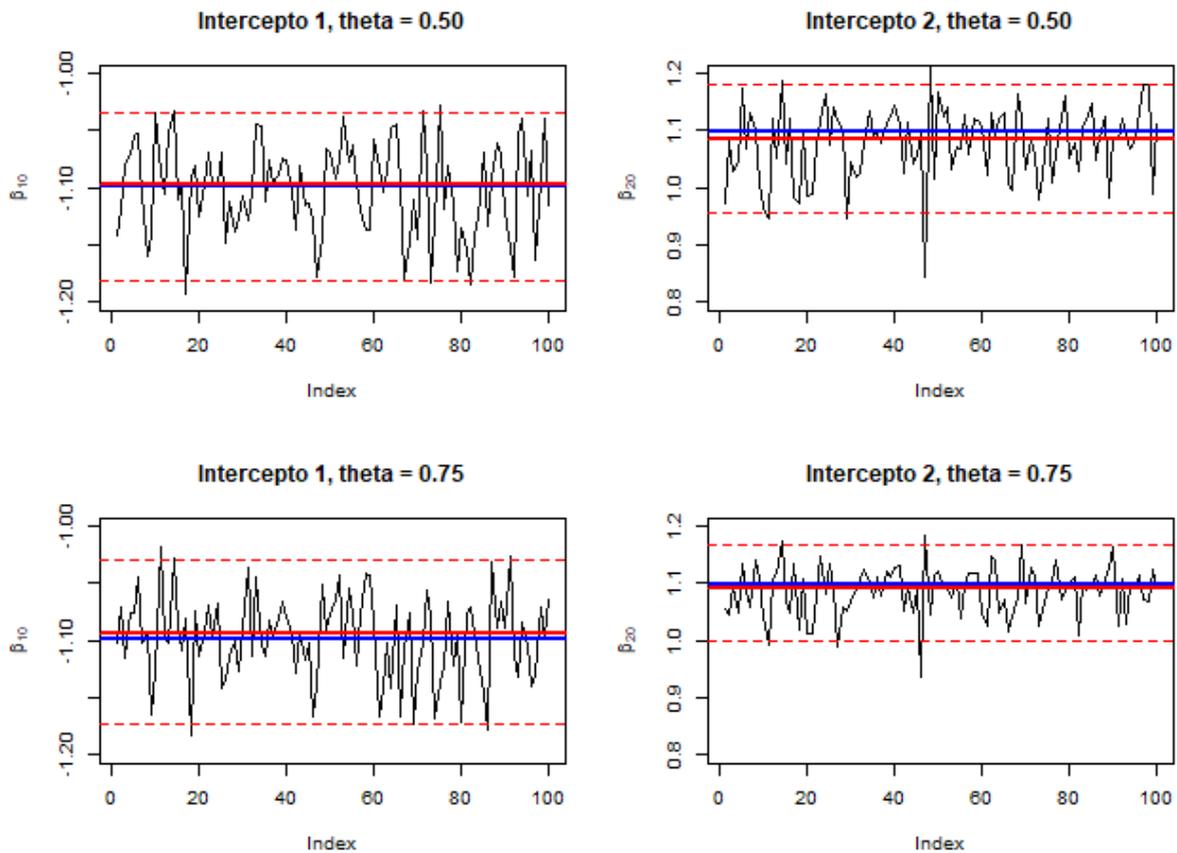


Figura 21 – Estimativas dos parâmetros de regressão, referente aos termos binomiais, para 100 amostras com parâmetro de variância no segundo componente igual a 0.50 – efeito do parâmetro de variância para o modelo binomial duplo.

5.2 Aplicação a dados reais

Para demonstrar o modelo aplicado a dados reais, serão ajustados os modelos apresentados neste trabalho para o conjunto de dados da *escala de ansiedade manifestada de Taylor* (Taylor, 1953). O conjunto considerado foi obtido do site *Open source psychometrics project* (disponível na Internet em http://openpsychometrics.org/_rawdata/). Foi considerado o conjunto de dados com última atualização no dia 22 de julho de 2012 com 5410 observações.

Para avaliação dos modelos, serão ajustados os modelos para explicar o escore de ansiedade manifestada de Taylor (Taylor, 1953) com relação a idade e ao sexo dos respondentes.

O conjunto original de dados tem 5410 observações dos quais, 2378 são respondentes do sexo masculino, 2993 são respondentes do sexo feminino, 30 responderam ter ‘outro’ sexo e 9 não responderam. Os elementos amostrais referentes ao sexo ‘outros’ e sem resposta para sexo foram eliminados do conjunto de dados, reduzindo o tamanho amostral para 5371 observações. O objetivo é que esta variável seja dicotômica, por isto a eliminação da categoria ‘outros’.

No conjunto de dados resultante haviam 2 respondentes que reportaram idades, respectivamente, iguais a 255 e 500 anos. Tais elementos foram eliminados da amostra, resultando em

um conjunto final com 5369 observações.

Idade	Sexo					
	Masculino			Feminino		
	mínimo	mediana	máximo	mínimo	mediana	máximo
(10,20]	4.0	34.0	49.0	5.0	38.0	50.0
(20,30]	2.0	33.0	49.0	1.0	36.0	49.0
(30,40]	0.0	32.0	49.0	3.0	35.0	49.0
(40,50]	4.0	31.0	47.0	3.0	33.0	50.0
(50,60]	2.0	30.0	47.0	5.0	31.0	47.0
(60,70]	2.0	25.0	48.0	1.0	25.0	46.0
(70,80]	1.0	20.5	46.0	7.0	30.5	45.0
(80,90]	–	–	–	–	–	–
(90,100]	–	–	–	20.0	20.0	20.0

Tabela 2 – Descrição da amostra de ansiedade manifestada de Taylor. Conjunto de dados atualizado até 22 de julho de 2012.

A Tabela 2 mostra a distribuição dos escores de ansiedade por sexo e idade. O escore é uma variável aleatória que assume valores inteiros, com intervalo 1 entre os valores, variando de 0 até 50. Observa-se que em ambos os sexos o nível de ansiedade apresenta medianas decrescentes com a idade. Não há qualquer evidência de que tal efeito ocorra com os limites mínimo e máximo dos valores. Praticamente para todas as faixas de idade os respondentes do sexo feminino apresentaram níveis maiores de ansiedade. Desta forma, é esperado que o nível de ansiedade decresça com a idade e que seja maior para as mulheres.

A Figura 22 apresenta os escores de ansiedade manifestada de Taylor. Observe que a distribuição dos escores é assimétrica e unimodal, com mediana em torno do escore 35. Para analisar estes dados serão ajustados modelos com um 1 e 2 componentes binomiais e um (único) componente uniforme discreto. No fim serão realizados testes de hipóteses para verificar se os modelos podem ser reduzidos.

A idade mínima observada na amostra foi 14 anos. Desta forma, a variável ‘idade do respondente’ foi subtraída deste valor e a variável ‘sexo do respondente’ foi codificada como uma dicotômica que vale 0 se sexo masculino e 1 se sexo feminino. Desta forma, o intercepto dos modelos se referem aos respondentes do sexo masculino com 14 anos de idade.

Ajuste dos modelos aos dados

Modelo binomial redimensionado e deslocado sem inflação

A Tabela 3 mostra o ajuste do modelo binomial redimensionado e deslocado sem inflação com 2 componentes aos dados de ansiedade manifestada de Taylor. No modelo ajustado, estima-se que a proporção de respondentes que atribuem notas de acordo com o componente 1 é igual a 29.85% (IC 95%: 27.98%; 31.72%), a proporção de respondentes que atribuem notas de acordo com o componente 2 é igual a 39.80% (IC 95%: 37.90%; 41.72%) e a proporção de respondentes

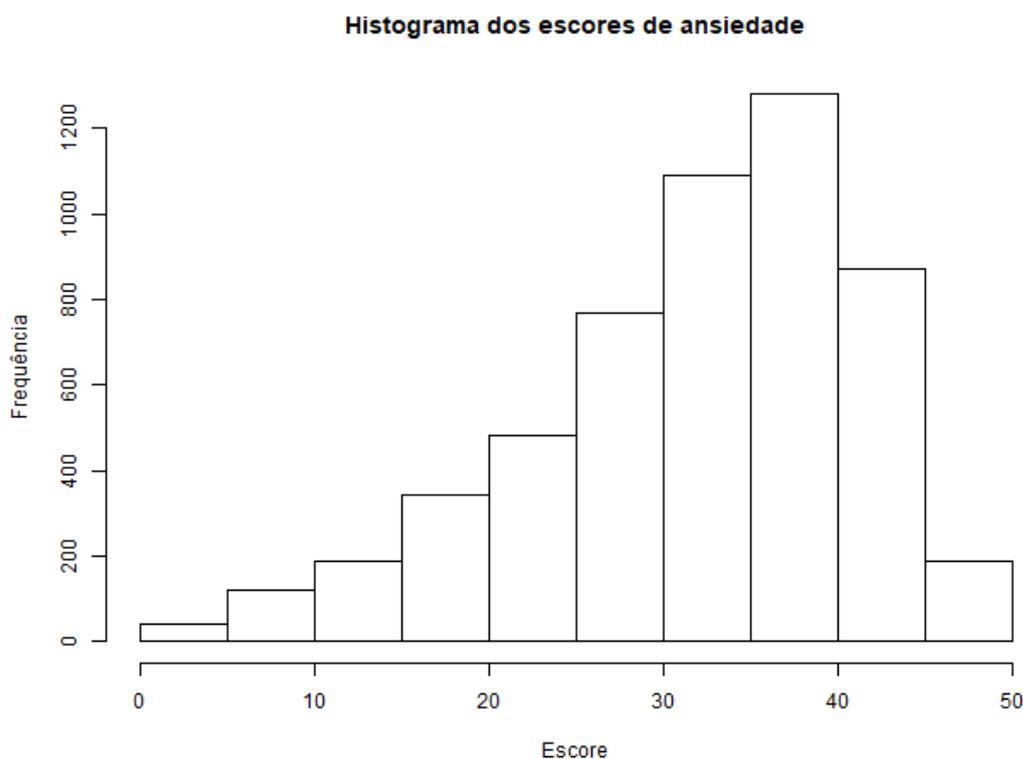


Figura 22 – Histograma dos escores de ansiedade manifestada de Taylor.

que atribuem notas ao acaso é calculado como 30.3%. O primeiro componente do modelo aponta um escore médio de 29.8346 (IC 95%: 29.3144; 30.3505) para os respondentes do sexo masculino com 14 anos de idade. A idade apresenta um efeito negativo no escore médio (exponencial), de tal forma que uma pessoa com 24 anos de idade do sexo masculino terá um escore médio de 27.9573 (IC 95%: 27.2142; 28.6953). O sexo feminino apresenta um aumento no escore médio, de modo que o escore médio para uma pessoa do sexo feminino com 14 anos de idade é 33.2672 (IC 95%: 32.3193; 34.1891). O segundo componente do modelo aponta um escore médio de 39.1506 (IC 95%: 38.8085; 39.4850) para os respondentes do sexo masculino com 14 anos de idade. O escore médio para uma pessoa com 24 anos de idade do sexo masculino é 38.1057 (IC 95%: 37.5996; 38.5973). Uma pessoa do sexo feminino com 14 anos de idade apresenta um escore médio de 41.2197 (IC 95%: 40.6436; 41.7674).

A Tabela 4 mostra o ajuste do modelo binomial redimensionado e deslocado sem inflação com 1 componente aos dados de ansiedade manifestada de Taylor. No modelo, a proporção de respondentes que atribuem notas com avaliações por qualidade é igual a 43.52% (IC 95%: 41.52%; 45.53%). O escore médio para respondentes do sexo masculino com 14 anos de idade é 36.5165 (IC 95%: 36.1900; 36.8381), o efeito da idade é tal que uma pessoa com 24 anos de idade do sexo masculino terá um escore médio de 35.4888 (IC 95%: 35.0039; 35.9644) e uma pessoa do sexo feminino com 14 anos de idade terá o escore médio igual a 38.6909 (IC 95%: 38.1163; 39.2456).

A log-verossimilhança para o modelo binomial redimensionado e deslocado sem inflação

Componente	Parâmetro	Estimativa	Err. Padrão
1	Mistura	0.2985	0.009350
	Intercepto	0.3917	0.021531
	Idade	-0.0154	0.000852
	Sexo (F)	0.2955	0.020472
2	Mistura	0.3981	0.009545
	Intercepto	1.2833	0.019908
	Idade	-0.0119	0.000761
	Sexo (F)	0.2631	0.018899

Tabela 3 – Ajuste do modelo binomial redimensionado e deslocado sem inflação com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.

Parâmetro	Estimativa	Err. Padrão
Mistura	0.4352	0.010025
Intercepto	0.9963	0.016454
Idade	-0.0102	0.000686
Sexo (F)	0.2337	0.015807

Tabela 4 – Ajuste do modelo binomial redimensionado e deslocado sem inflação com 1 componente binomial aos dados de ansiedade manifestada de Taylor.

com 2 componentes binomiais é igual a -19370.63 e para o modelo com 1 componente binomial é igual a -19947.89 .

Para testar a hipótese de homogeneidade entre os componentes binomiais do modelo é possível aplicar o teste de razão de verossimilhança. Para isso, calcula-se as razões das verossimilhanças. A estatística do teste é $\lambda = 1154.51$ que segue uma distribuição qui-quadrado com 4 graus de liberdade. O teste é tal que $p < 0.0001$ e, assim, rejeita-se a hipótese de homogeneidade entre os componentes no teste de razão de verossimilhanças.

Para proceder com o teste de Wald é necessário construir um vetor de parâmetros sob a hipótese nula de tal forma que o parâmetro de mistura de cada componente seja o valor ajustado no modelo com 1 componente binomial, dividido pelo número de componentes binomiais no modelo não restrito (ou seja, dividido por 2) e os parâmetros de regressão são iguais para ambos os componentes sob a hipótese nula. Procedendo com o cálculo, obtém-se uma estatística de teste igual a $Q = 3.9292$ com distribuição qui-quadrado com 4 graus de liberdade. A hipótese de homogeneidade entre os componentes não é rejeitada neste caso com $p = 0.4157$.

De forma similar, para proceder com o teste gradiente utiliza-se o mesmo vetor calculado anteriormente para o teste de Wald. A estatística do teste é $S = 0.1273$ com distribuição qui-quadrado com 4 graus de liberdade. A hipótese de homogeneidade entre os componentes não é rejeitada com $p = 0.9981$.

Modelo binomial redimensionado e deslocado com inflação

A Tabela 5 apresenta o ajuste aos dados de ansiedade para o modelo binomial redimensionado e deslocado com inflação com 2 componentes binomiais. O escore mais frequente na amostra foi o 39 e, desta forma, assumiu-se que tal valor estava inflacionado (não havia evidências). No modelo ajustado, estima-se que a proporção de respondentes que atribuem notas de acordo com o componente 1 é igual a 29.85% (IC 95%: 27.98%; 31.72%), a proporção de respondentes que atribuem notas de acordo com o componente 2 é igual a 39.81% (IC 95%: 37.73%; 41.89%) e a proporção de respondentes que atribuem notas ao acaso como 30.34% (IC 95%: 28.24%; 32.44%). Os efeitos observados são bastante parecidos com os do modelo sem inflação, havendo apenas muito pequenas nas estimativas na última casa decimal. Observe que a proporção de inflação estimada será $100.0\% - 29.85\% - 39.81\% - 30.34\% \approx 0.00\%$, indicando que o modelo ajustado coincide com aquele não inflacionado.

Componente	Parâmetro	Estimativa	Err. Padrão
1	Mistura	0.2985	0.009362
	Intercepto	0.3918	0.021554
	Idade	-0.0154	0.000853
	Sexo (F)	0.2955	0.020480
2	Mistura	0.3981	0.010410
	Intercepto	1.2833	0.019949
	Idade	-0.0119	0.000777
	Sexo (F)	0.2631	0.019176
Uniforme	Mistura	0.3034	0.010459

Tabela 5 – Ajuste do modelo binomial redimensionado e deslocado com inflação com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.

A log-verossimilhança para o modelo binomial redimensionado e deslocado sem inflação com 2 componentes binomiais é igual a -19370.63 e para o modelo similar com inflação é igual a -19370.64 .

Para testar a hipótese de que não há inflação nos dados, pode ser aplicado o teste de razão de verossimilhança. Para isso, calcula-se as razões das verossimilhanças. A estatística do teste é $\lambda = 0.0059$ que segue uma distribuição qui-quadrado com 1 graus de liberdade. O teste é tal que $p = 0.9388$ e, assim, não é possível rejeitar a hipótese de não inflação nos dados.

Para proceder com o teste de Wald é necessário construir um vetor de parâmetros sob a hipótese nula de tal forma que a soma dos parâmetros de mistura de no modelo completo seja igual a 1. Para isto, basta tomar $\tau_0 = 1 - \tau_1 - \tau_2$. Procedendo com o cálculo, obtém-se uma estatística de teste igual a $Q < 0.0001$ com distribuição qui-quadrado com 1 graus de liberdade. A hipótese de ausência de inflação não é rejeitada neste caso com $p = 0.9999$.

De forma similar, para proceder com o teste gradiente utiliza-se o mesmo vetor calculado anteriormente para o teste de Wald. A estatística do teste é $S = 0.0029$ com distribuição qui-

quadrado com 1 grau de liberdade. A hipótese de homogeneidade entre os componentes não é rejeitada com $p = 0.9571$.

Modelo binomial duplo

As Tabelas 6 e 7 apresentam, respectivamente, os ajustes do modelo binomial duplo aos dados de ansiedade.

No modelo com 2 componentes, estima-se que a proporção de respondentes que atribuem notas de acordo com o componente 1 é igual a 38.97% (IC 95%: 36.23%; 41.71%), a proporção de respondentes que atribuem notas de acordo com o componente 2 é igual a 61.02% (IC 95%: 58.76%; 63.28%) e a proporção de respondentes que atribuem notas ao acaso é calculado como 0.01%. O primeiro componente do modelo aponta um escore médio de 28.6475 (IC 95%: 28.4941; 28.8007) para os respondentes do sexo masculino com 14 anos de idade. A idade apresenta um efeito negativo no escore médio tal que uma pessoa com 24 anos de idade do sexo masculino terá escore igual a 25.4987 (IC 95%: 25.2751; 25.7221). O sexo feminino apresenta um aumento no escore médio, de modo que o escore para uma pessoa do sexo feminino com 14 anos de idade é 32.8843 (IC 95%: 32.6010; 33.1653). O segundo componente do modelo aponta um escore médio de 36.2839 (IC 95%: 36.1207; 36.4458) para os respondentes do sexo masculino com 14 anos de idade. O escore médio para uma pessoa com 24 anos de idade do sexo masculino é 35.6480 (IC 95%: 35.4179; 35.8758). Uma pessoa do sexo feminino com 14 anos de idade apresenta um escore médio de 38.2105 (IC 95%: 37.9186; 38.4975).

No modelo com 1 componente, o escore médio é 33.9633 (IC 95%: 33.8820; 34.0443) para os respondentes do sexo masculino com 14 anos de idade. Uma pessoa com 24 anos de idade do sexo masculino terá um escore médio de 32.3217 (IC 95%: 32.2048; 32.4382). O escore médio para uma pessoa do sexo feminino com 14 anos de idade é 36.6774 (IC 95%: 36.5331; 36.8207).

Componente	Parâmetro	Estimativa	Err. Padrão
1	Mistura	0.3897	0.013707
	Variância	0.1294	0.005394
	Intercepto	0.2939	0.006264
	Idade	-0.0254	0.000268
	Sexo (F)	0.3591	0.006269
2	Mistura	0.6102	0.011316
	Variância	0.2792	0.010108
	Intercepto	0.9728	0.008167
	Idade	-0.0063	0.000302
	Sexo (F)	0.2031	0.007897

Tabela 6 – Ajuste do modelo binomial duplo com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.

A log-verossimilhança para o modelo binomial duplo com 2 componentes binomiais é -19048.63 e para o modelo com 1 componente binomial é igual a -19261.44 .

Parâmetro	Estimativa	Err. Padrão
Mistura	0.9340	0.008455
Variância	0.1534	0.003714
Intercepto	0.7504	0.003725
Idade	-0.0147	0.000138
Sexo (F)	0.2623	0.003632

Tabela 7 – Ajuste do modelo binomial duplo com 1 componente binomial aos dados de ansiedade manifestada de Taylor.

O teste de razão de verossimilhanças para testar a homogeneidade entre os componentes binomiais é tal que a estatística do teste é $\lambda = 425.6306$ que segue uma distribuição qui-quadrado com 5 graus de liberdade. A hipótese de homogeneidade é rejeitada com $p < 0.0001$.

A estatística do teste de Wald para a hipótese de homogeneidade é igual a $Q = 0.2479$ com distribuição qui-quadrado com 4 graus de liberdade. A hipótese de homogeneidade entre os componentes não é rejeitada neste caso com $p = 0.0015$.

De forma similar, a estatística do teste gradiente é $S = 16.6983$ com distribuição qui-quadrado com 5 graus de liberdade. A hipótese de homogeneidade entre os componentes é rejeitada com $p = 0.0051$.

Modelo binomial multiplicativo

As Tabelas 8 apresenta os ajustes do modelo binomial multiplicativo aos dados de ansiedade. Observe que, ao contrário dos modelos anteriores, a diferença entre os componentes do modelo se refletiu em um intercepto negativo no primeiro componente. A idade do respondente que, em todos os demais modelos, apresentavam efeito negativo (ou seja, a ansiedade diminui com a idade), neste modelo apresenta efeito positivo. Observa-se também que o modelo apresenta erros padrão bem maiores que aqueles dos demais modelos.

No modelo com 2 componentes, estima-se que a proporção de respondentes que atribuem notas de acordo com o componente 1 é igual a 3.72% (IC 95%: 2.33%; 5.11%), a proporção de respondentes que atribuem notas de acordo com o componente 2 é igual a 19.51% (IC 95%: 17.77%; 21.25%) e a proporção de respondentes que atribuem notas ao acaso é calculado como 76.77%. Observe que o componente predominante no modelo foi o de avaliações ao acaso, mostrando que o modelo pode ter falhado em ajustar as avaliações não realizadas ao acaso. Isto explica, em parte, o comportamento distinto dos modelos (em relação aos anteriores) e os valores de erro padrão bastante elevados. O primeiro componente do modelo aponta um escore médio de 9.8472 (IC 95%: 3.9003; 20.7754) para os respondentes do sexo masculino com 14 anos de idade. A idade apresenta um efeito positivo no escore médio tal que uma pessoa com 24 anos de idade do sexo masculino terá escore igual a 18.2526 (IC 95%: 8.2567; 31.2816). O sexo feminino apresenta um aumento no escore médio, de modo que o escore para uma pessoa do sexo feminino com 14 anos de idade é 10.6252 (IC 95%: 4.2300; 22.0343). O segundo componente do modelo aponta um escore médio de 41.4680 (IC 95%: 39.0994; 43.4087) para os respondentes do sexo

masculino com 14 anos de idade. O escore médio para uma pessoa com 24 anos de idade do sexo masculino é 46.1581 (IC 95%: 44.2782; 47.4558). Uma pessoa do sexo feminino com 14 anos de idade apresenta um escore médio de 42.0795 (IC 95%: 39.0902; 44.3677).

O modelo com apenas 1 componente binomial apresentou um ajuste ruim, indicando 0,04% de observações geradas pelo componente binomial e intervalos de confiança para os escores médios cobrindo todos os valores da escala. Apenas para efeito de demonstração, o ajuste é apresentado na Tabela 9. Os testes de hipóteses serão omitidos.

Componente	Parâmetro	Estimativa	Err. Padrão
1	Mistura	0.0372	0.006968
	Log-variância	0.0945	0.049802
	Intercepto	-1.4055	0.532130
	Idade	0.0852	0.031809
	Sexo (F)	0.0956	0.195820
2	Mistura	0.1951	0.008679
	Log-variância	0.0811	0.010210
	Intercepto	1.5811	0.151906
	Idade	0.0905	0.006804
	Sexo (F)	0.0890	0.045040

Tabela 8 – Ajuste do modelo binomial multiplicativo com 2 componentes binomiais aos dados de ansiedade manifestada de Taylor.

Componente	Parâmetro	Estimativa	Err. Padrão
1	Mistura	0.0037	0.007144
	Log-variância	0.0816	0.501374
	Intercepto	-1.4184	5.947966
	Idade	0.0889	0.367928
	Sexo (F)	0.0869	2.007498

Tabela 9 – Ajuste do modelo binomial multiplicativo com 1 componente binomial aos dados de ansiedade manifestada de Taylor.

A log-verossimilhança para o modelo binomial multiplicativo com 2 componentes binomiais é -20698.25 e para o modelo com 1 componente binomial é igual a -21109.47 .

Avaliação da qualidade de ajuste

Para avaliar as medidas de qualidade de ajuste dos modelos, foram calculadas as medidas de alavanca, influência e influência local para os modelos binomial redimensionado e deslocado (Figuras 23 e 26), binomial duplo (Figuras 24 e 27) e binomial multiplicativo (Figuras 25 e 28).

Comparando as Figuras 23, 24 e 25 observa-se que, como esperado, a maioria dos pontos estão concentrados na base do gráfico e alguns poucos se afastam da nuvem. Observando as posições destes pontos que se afastam das núvens nos diferentes gráficos observa-se que estão na

mesma região. Isto indica, não foram devidamente identificados, que podem se tratar dos mesmos pontos que, independente do modelo escolhido, causam perturbação no ajuste dos modelos.

É interessante também observar que a medida de alavanca se mantpem em uma faixa entre 1.0 e 3.5 em todos os modelos avaliados.

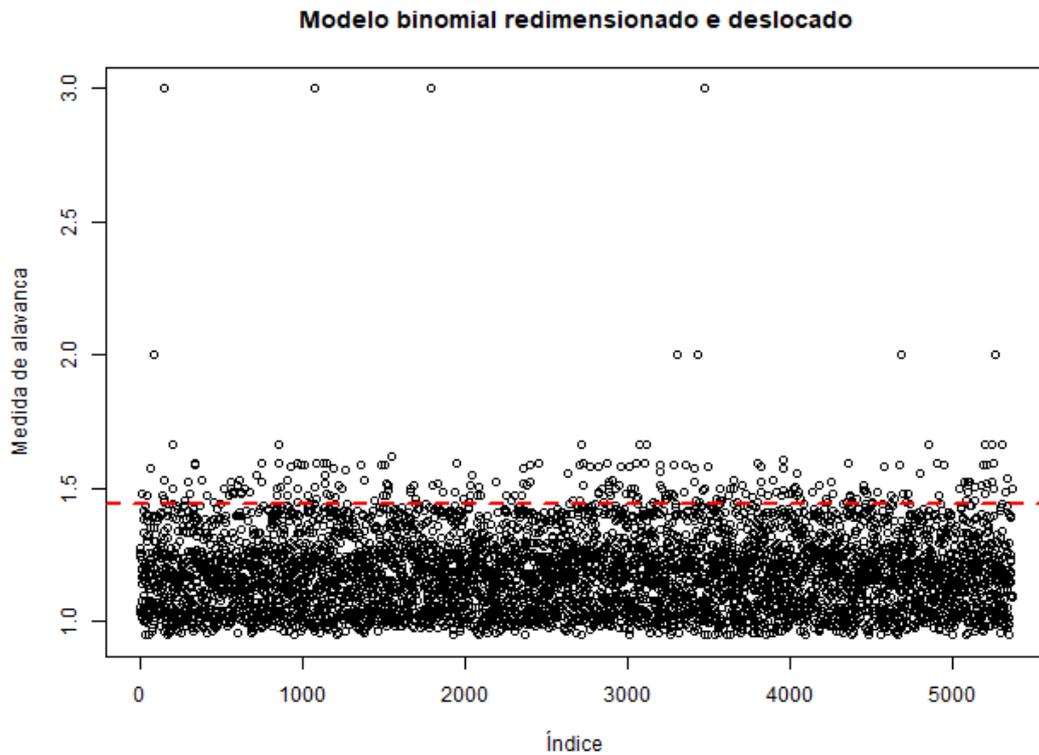


Figura 23 – Medidas de alavanca para os dados de ansiedade manifestada de Taylor – modelo binomial redimensionado e deslocado.

As Figuras 26, 27 e 28 apresentam as medidas de influência e influência local para cada um dos modelos. Comparando as medidas para cada componente dos modelos, observa-se que a ordem de grandeza das medidas permanecem aproximadamente iguais. Isto indica que os componentes ajustados apresentam aproximadamente a mesma medida de verossimilhança para cada elemento amostral.

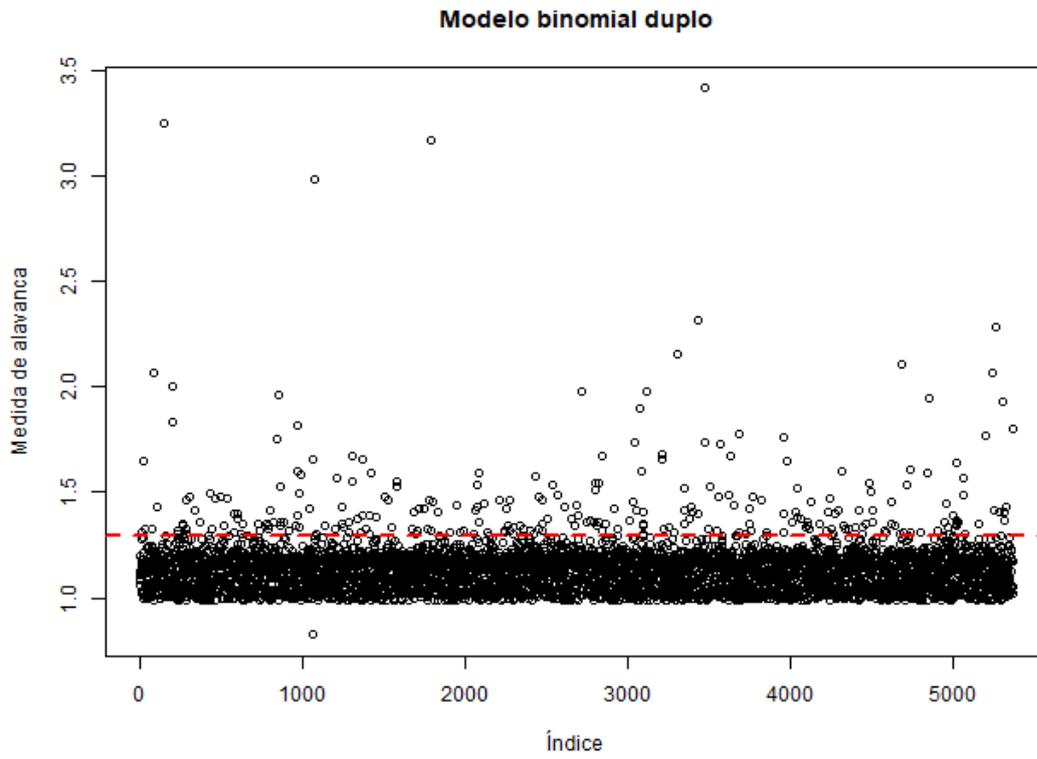


Figura 24 – Medidas de alavanca para os dados de ansiedade manifestada de Taylor – modelo binomial duplo.

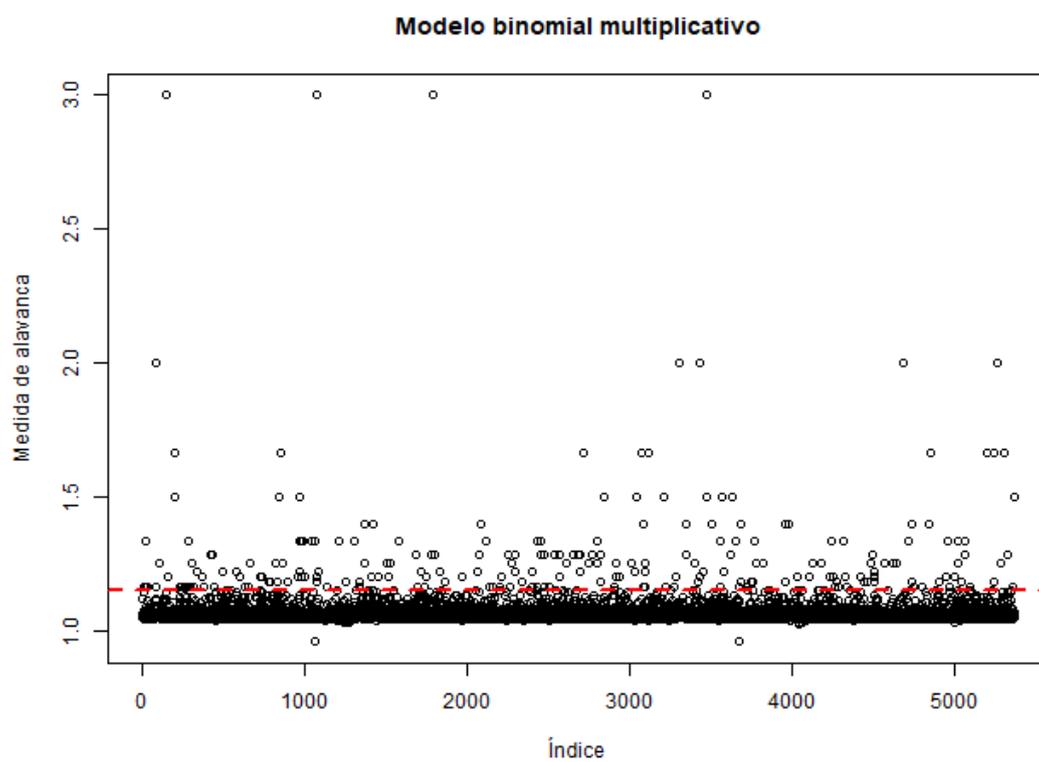


Figura 25 – Medidas de alavanca para os dados de ansiedade manifestada de Taylor – modelo binomial multiplicativo.

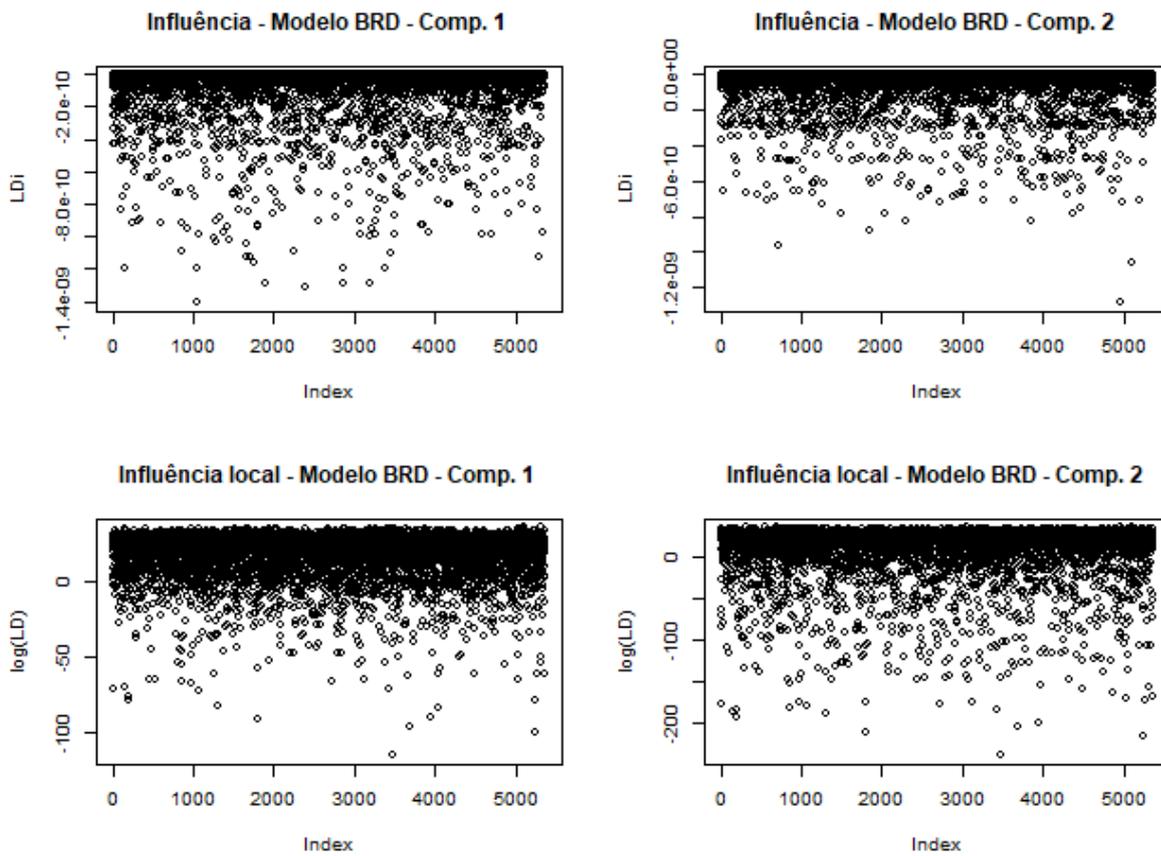


Figura 26 – Medidas de influência para os dados de ansiedade manifestada de Taylor – modelo binomial redimensionado e deslocado.

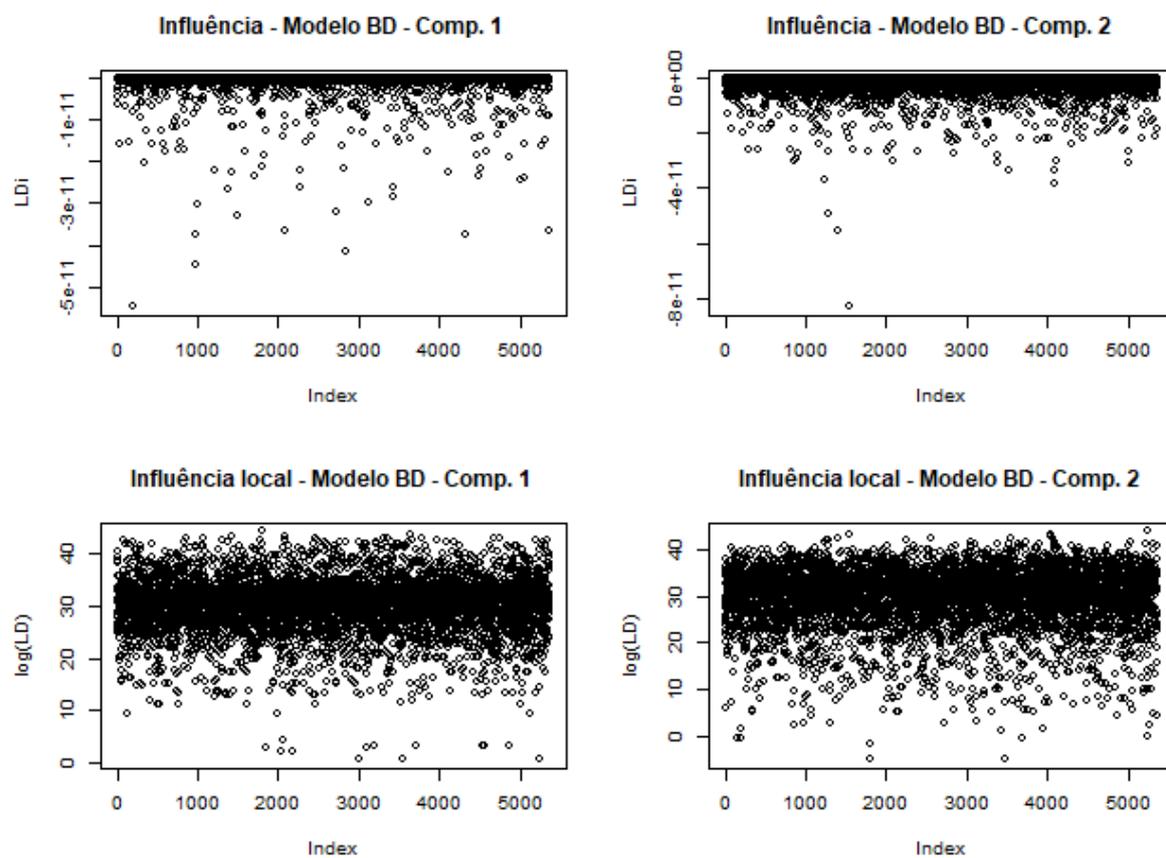


Figura 27 – Medidas de influência para os dados de ansiedade manifestada de Taylor – modelo binomial duplo.

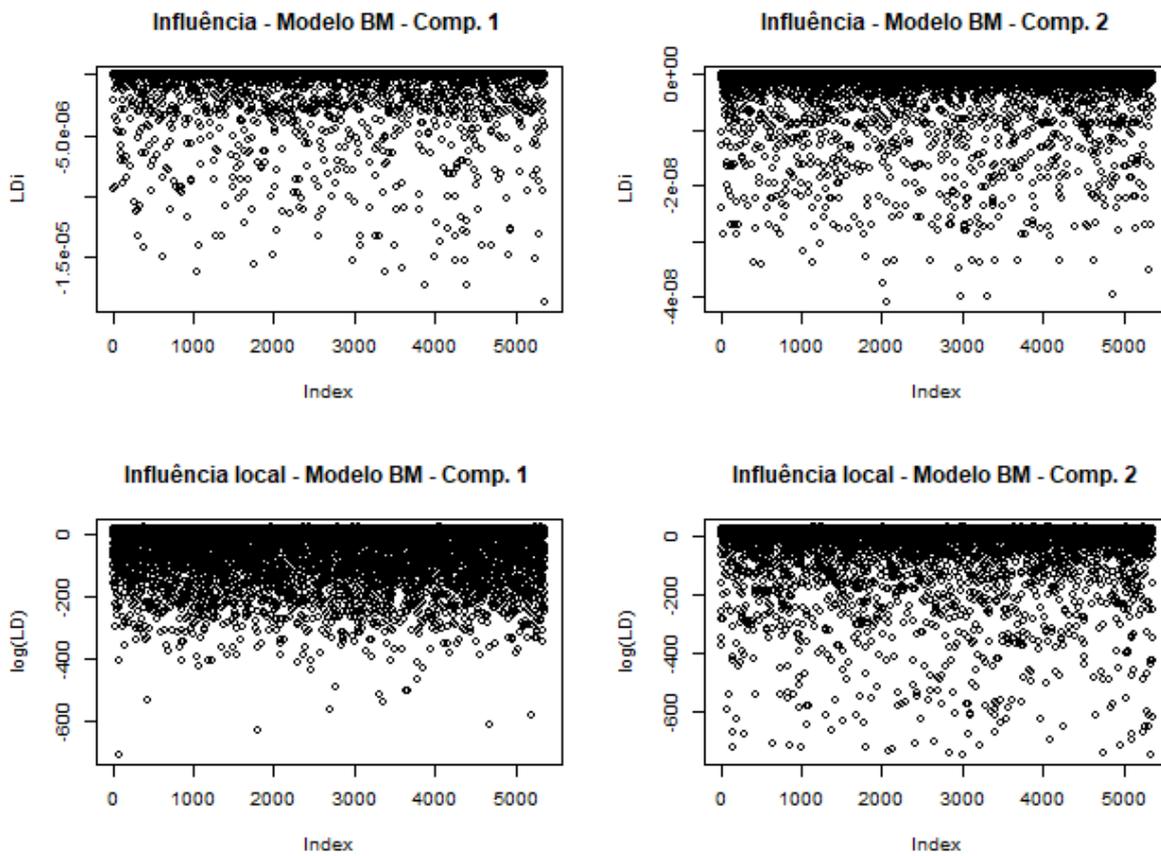


Figura 28 – Medidas de influência para os dados de ansiedade manifestada de Taylor – modelo binomial multiplicativo.

6 Conclusão

O sucesso de um produto ou de um serviço é resultado de uma série de fatores, dentre os quais a percepção da qualidade pelos clientes ou usuários é um dos mais importantes. Um produto ou serviço que não atenda às necessidades dos clientes em geral não permanece no mercado por muito tempo. Desta forma, a avaliação destes produtos e serviços é uma necessidade para se traçar estratégias de melhoria continuada, sobretudo quando se planeja um novo serviço, se desenvolve novos produtos ou, até mesmo, com o objetivo de direcionar campanhas de marketing. A avaliação da percepção da qualidade por parte dos clientes permite às empresas identificar pontos de melhoria.

No entanto, o processo de avaliação destes produtos (serviços) pelo público pode incorrer em diferentes estratégias de atribuição de notas (ou ordenação por preferência), inclusive com avaliação ao acaso, sem que as características sob avaliação sejam de fato averiguadas.

Este trabalho apresenta uma extensão para o modelo CUB (*Combined Uniform and shifted Binomial*) para a avaliação de produtos pela atribuição de notas discretas. Além disto, são apresentados 2 modelos para a situação em que a suposição de que a variância da resposta é determinada unicamente pela média não seja adequada.

Os modelos citados, bem como as medidas de qualidade de ajuste e métodos para o teste de hipóteses são apresentados. Os modelos, também, são avaliados por meio de um estudo de simulação onde algumas características observadas são comentadas.

Diversos outros modelos podem ser construídos sob as mesmas premissas – dentre eles é de interesse o desenvolvimento de modelos para escalas de notas não limitadas (que ocorrem em cenários de avaliação seriadas, com a atribuição de 1 ponto por avaliação, com um número ilimitado de possíveis avaliações). Modelos multivariados também podem ser definidos para os cenários nos quais o avaliador quantifica diversas características dos produtos, resultando em um vetor de notas.

Referências

- Altham, P. M. E. Two generalizations of the binomial distribution. *JRSS C*, v. 27, p. 162 – 167, 1978. Citado 2 vezes nas páginas 35 e 39.
- Bailey, W. N.; Terry, M. E. The Rank Analysis of Incomplete Block Designs. I. The Method of Paired Comparisons. *Biometrika*, v. 39, 1952. Citado na página 18.
- Beaver, R. J. Weighted least squares analysis of several univariate Bradley-Terry models. *J. Amer. Statist. Assoc.*, v. 72, p. 629 – 634, 1977. Citado na página 19.
- Bockenholt, U. Application to Thurstonian models to ranking data. In: A., F. M.; Verducci, J. S. (Ed.). *Probability Models and Statistical Analyses of Ranking Data*. New York: Springer, 1993. p. 157 – 172. Citado na página 19.
- Bradley, R. A. The Rank Analysis of Incomplete Block Designs. III. Some Large-Samples Results on Estimation and Power for a Method of Paired Comparisons. *Biometrika*, v. 42, p. 450 – 470, 1955. Citado na página 19.
- Bradley, R. A.; El-Helbawy, A. T. Treatment Contrasts in Paired Comparisons: Basic Procedures with Application to Factorials. *Biometrika*, v. 63, p. 255 – 262, 1976. Citado na página 19.
- Cohen, J. B.; Goldberd, M. E. The Dissonance Model in Post-Decision Product Evaluation. *Journal of Marketing Research*, v. 7, n. 3, p. 315 – 321, 1970. Citado na página 17.
- Cook, R. D. Assessment of local influence (with discussion). *JRSS B*, v. 48, p. 133 – 169, 1986. Citado na página 50.
- Davidson, R. R.; Beaver, R. J. On Extending the Bradley-Terry Model to Incorporate Within-Pair Order Effects. *Biometrics*, v. 33, n. 4, p. 693 – 702, 1977. Citado na página 19.
- D'Elia, A. A shifted binomial model for ranks. In: Nunez-Antón, V.; Ferreira, E. (Ed.). *Proceedings of the 15th International Workshop on Statistical Modelling*. Bilbao: Servicio Editorial de la Universidad del País Vasco, 2000. p. 412 – 416. Citado 2 vezes nas páginas 20 e 23.
- D'Elia, A.; Piccolo, D. A mixture model for preferences data analysis. *Computational Statistics and Data Analysis*, v. 49, p. 917 – 934, 2005. Citado 3 vezes nas páginas 20, 23 e 26.
- Dempster, A. P.; Laird, N. M.; Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *JRSS B*, v. 39, n. 1, p. 1 – 38, 1977. Citado 2 vezes nas páginas 30 e 33.
- Dodds, W. B.; Monroe, K. B.; Grewal, D. Effects of Price, Brand, and Store Information on Buyer's Product Evaluation. *Evaluation and Program Planning*, v. 2, p. 197 – 207, 1979. Citado na página 17.
- Efron, B. Double exponential families and their use in Generalized Linear Regression. *Journal of the American Statistical Association*, v. 395, p. 709 – 721, 1986. Citado 3 vezes nas páginas 35, 36 e 38.
- Fligner, M. A.; Verducci, J. S. *Probability Models and Statistical Analyses of Ranking Data*. New York: Springer, 1993. Citado na página 18.
- Ford Jr., L. R. Solution of a Ranking Problem from Binary Comparisons. *Amer. Math. Monthly*, v. 64, p. 28 – 33, 1957. Citado na página 19.

Gottard, A.; Iannario, M.; Piccolo, D. Varying uncertainty in CUB models. *Advances in Data Analysis and Classification*, v. 1 – 20, 2016. Citado 3 vezes nas páginas 20, 23 e 27.

Grilli, L. et al. Latent class CUB models. Disponível na internet em www.researchgate.net/publication/257705512. 2014. Citado na página 20.

Grizzle, J. E.; Starmer, F.; Koch, G. G. Analysis of Categorical Data by Linear Models. *Biometrics*, v. 25, n. 3, p. 489 – 504, 1969. Citado na página 19.

Homburg, C.; Hoyer, W. D.; Koschate, N. Consumers' Reaction to Price Increases: Do Customers Satisfaction and Perceived Motive Fairness Matter? *Journal of Academy of Marketing Science*, v. 53, n. 1, p. 36 – 49, 2005. Citado na página 17.

Hui, M. K.; Tse, D. K. What to Tell Consumers in Waits of Different Lengths: An Integrative Model of Service Evaluation. *Journal of Marketing*, v. 60, n. 2, p. 81 – 90, 1996. Citado na página 18.

Iannario, M. On the identifiability of a mixture model for ordinal data. *Metron*, LXVIII, n. 1, p. 87 – 94, 2010. Citado na página 20.

Iannario, M. Hierarchical CUB models for ordinal variables. *Communications in Statistics, Theory and Methods*, v. 41, p. 3110 – 3125, 2012. Citado na página 20.

Iannario, M. Modelling shelter choices in a class of mixture models for ordinal responses. *Statistical Methods and Applications*, v. 21, p. 1 – 22, 2012. Citado 3 vezes nas páginas 20, 23 e 30.

Iannario, M. Modeling uncertainty and overdispersion in ordinal data. *Communications in Statistics, Theory and Methods*, v. 43, n. 4, 2014. Citado na página 20.

Iannario, M.; Piccolo, D. A theorem on CUB models for rank data. *Statistics and Probability Letters*, v. 91, p. 27 – 31, 2014. Citado na página 21.

Larsen, D. L. et al. Assessment of Client/Patient Satisfaction: Development of a General Scale. *Evaluation and Program Planning*, v. 2, p. 197 – 207, 1979. Citado na página 17.

Lehmann, E. L.; Romano, J. P. *Testing Statistical Hypotheses*. New York: Springer, 2005. Citado na página 57.

Manisera, M.; Zuccolotto, P. Modeling rating data with nonlinear CUB models. *Computational Statistics and Data Analysis*, v. 78, p. 100 – 118, 2014. Citado na página 20.

Marden, J. I. *Analyzing and Modelling Rank Data*. London: Chapman & Hall, 1995. Citado na página 18.

Murphy, T. B.; Martin, D. Mixtures of distance-based models for ranking data. *Comput. Statist. Data Anal.*, v. 41, p. 645 – 655, 2003. Citado na página 19.

Nelder, J. A.; Wedderburn, R. W. M. Generalized linear models. *JRSS, A* – 135, p. 370 – 384, 1972. Citado na página 24.

Piccolo, D. Observed information matrix for MUB models. *Quaderni di Statistica*, v. 8, p. 33 – 78, 2006. Citado 2 vezes nas páginas 20 e 23.

Pregibon, D. Logistic regression diagnostics. *Annals of Statistics*, v. 9, p. 705 – 724, 1981. Citado na página 48.

- Rao, C. R. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, v. 44, n. 1, p. 50 – 57, 1948. Citado 2 vezes nas páginas 51 e 60.
- Springall, A. Response Surface Fitting using a Generalization of the Bradley-Terry Paired Comparison Model. *Appl. Statist.*, v. 22, p. 1973, 1973. Citado na página 19.
- Stern, H. Probability models on rankings and the electoral process. In: Fligner, M. A.; Verducci, J. S. (Ed.). *Probability Models and Statistical Analyses of Ranking Data*. New York: Springer, 1993. p. 173 – 195. Citado na página 19.
- Sweeney, J.; Soutar, G. N. Consumer Perceived Value: The Development of a Multiple Scale Item. *Journal of Retailing*, v. 77, p. 203 – 220, 2001. Citado na página 17.
- Taylor, J. A personality scale for manifest anxiety. *The Journal of Abnormal and Social Psychology*, v. 48, n. 2, p. 285 – 290, 1953. Citado na página 93.
- Taylor, S. A.; Baker, T. L. An Assessment of the Relationship Between Service Quality and Consumer Satisfaction in the Formation of Consumers' Purchase Intentions. *Journal of Retailing*, v. 70, n. 2, p. 163 – 178, 1994. Citado na página 18.
- Terrell, G. R. The gradient statistic. *Computing Science and Statistics*, v. 34, p. 206 – 215, 2002. Citado 2 vezes nas páginas 51 e 64.
- Wald, A. Test of statistical hypothesis concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, v. 54, n. 3, p. 426 – 482, 1943. Citado 2 vezes nas páginas 51 e 57.
- Wilks, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, v. 9, p. 60 – 62, 1938. Citado 2 vezes nas páginas 51 e 52.
- Zeithaml, V. A. Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, v. 52, n. 3, p. 2 – 22, 1988. Citado na página 17.

Anexos

ANEXO A – Códigos para estimação dos modelos

Modelo binomial redimensionado e deslocado sem inflação

```

# -----#
# Biblioteca usada para inverter a matriz de      #
# informacao de Fisher.                          #
# -----#
library(MASS)

# -----#
# Modelo binomial redimensionado e deslocado sem inflacao. #
#
# formula – formula do modelo ( $Y \sim X1 + X2$ )      #
# data – conjunto de dados                          #
# t – numero de componentes                          #
# m – tamanho da escala                             #
# d – deslocamento da escala                        #
# s – intervalo de valores na escala                #
# tolerance – tolerancia na estimacao (padrao 1e-6) #
# maxiter – maximo de iteracoes (padrao 500)       #
# -----#
binom.model = function(formula , data , t = 3, m = 5, d = 1, s = 2,
  tolerance = 1e-6, maxiter = 500, hessian = FALSE) {

  # Checagem da formula
  if (class(formula) != "formula")
    stop("'formula' must be a valid formula object.")
  if(attr(terms(formula), "response") != 1)
    stop("Model does not have a response.")

  # Monta o data frame de acordo com a formula
  df = model.frame(formula , data = data)

  # Separa respostas e covariaveis
  z = df[,1]
  X = as.matrix(df[, -1])

```

```

# Adiciona intercepto se houver
if(attr(terms(formula), "intercept") == 1) X = cbind(1, X)

# Valores iniciais dos parametros
alpha = rep(1/(t + 1), t)
beta = matrix(0, nrow = t, ncol = ncol(X))
beta[,1] = seq(-2.5, 2.5, length.out = t)

# Variaveis para controle de convergencia
count = 0
loglik = -Inf
converged = TRUE

# Probabilidade da avaliacao ao acaso
p0 = rep(1/(m - d + 1), nrow(X))

# Probabilidades dos termos binomiais (matriz)
pr = matrix(0, nrow = nrow(X), ncol = t)

# Estimacao
repeat {
  # Calcula os componentes binomiais
  for (k in 1:t) {
    eta = X %*% beta[k,]
    pr[,k] = dbinom(x = z/s - d, size = m - d,
                    prob = exp(eta)/(1 + exp(eta)))
  }

  # Calcula as probabilidades
  pp = (1 - sum(alpha)) * p0
  for (k in 1:t) pp = pp + alpha[k] * pr[,k]

  # Atualiza os pesos
  W = matrix(0, nrow = nrow(X), ncol = length(alpha))
  for (k in 1:t) W[,k] = (alpha[k] * pr[,k])/pp

  # Atualiza os parametros de mistura
  alpha.old = alpha
  alpha = colMeans(W)
  alpha.dt = max(abs(alpha - alpha.old))
}

```

```

# Atualiza os parametros de regressao
beta.old = beta
for (k in 1:t) {
  eta = X %*% beta[k,]

  U = colSums(W[,k] * X * as.vector((((z/s - d) -
    (m - d) * exp(eta)/(1 + exp(eta))))))
  K = (t(sqrt(W[,k]) * X * as.vector((m - d) *
    exp(eta)/((exp(eta) + 1)^2))) %*%
    (sqrt(W[,k]) * X))
  beta[k,] = beta[k,] + as.vector(ginv(K) %*% U)
}
beta.dt = max(abs(beta - beta.old))

# Calcula verossimilhanca
loglik.old = loglik
loglik = sum(log(pp))
loglik.dt = abs(loglik - loglik.old)

# Contagem de iteracoes
count = count + 1

# Verifica falta de convergencia
converged = (count <= maxiter)
if (!converged) break

# Regras de parada
if (max(alpha.dt, beta.dt) < tolerance) break
if (loglik.dt < tolerance) break
}

# Calcula os escores
pp = (1 - sum(alpha)) * p0
for (k in 1:t) pp = pp + alpha[k] * pr[,k]

U.A = matrix(0, nrow = nrow(X), ncol = t)
for (r in 1:t) U.A[,r] = (pr[,r] - p0)/pp

U.B = array(0, dim = c(t, nrow(X), ncol(X)))

for (r in 1:t) {
  eta = X %*% beta[r,]

```

```

temp.1 = (X * ((alpha[r] * pr[,r])/pp))
temp.2 = (z/s - d - ((m - d) *
           exp(eta))/(1 + exp(eta)))

U.B[r, ,] = temp.1 * as.vector(temp.2)
}

# Calcula a matriz hessiana
H = diag((ncol(X) + 1) * t)
if (hessian) {
  H[1:t, 1:t] = t(U.A) %*% U.A
  for (r in 1:t) {
    pos.r = t + (r - 1)*ncol(X) + 1
    H[1:t, pos.r:(pos.r + ncol(X) - 1)] =
      t(U.A) %*% U.B[r, ,]
    H[pos.r:(pos.r + ncol(X) - 1), 1:t] =
      t(H[1:t, pos.r:(pos.r + ncol(X) - 1)])

    for (s in 1:t) {
      pos.s = t + (s - 1)*ncol(X) + 1
      H[pos.r:(pos.r + ncol(X) - 1),
        pos.s:(pos.s + ncol(X) - 1)] =
        t(U.B[r, ,]) %*% U.B[s, ,]
    }
  }
}

# Exporta os resultados
result = list(alpha = alpha, beta = beta, loglik = loglik,
              converged = converged, U.A = U.A, U.B = U.B)

if (hessian) {
  result$H = H
  result$positive.definite = all(eigen(H)$values > 0)
}
result
}

```

Modelo binomial redimensionado e deslocado com inflação

```

# -----#
# Biblioteca usada para inverter a matriz de      #
# informacao de Fisher.                          #
# -----#
library(MASS)

# -----#
# Modelo binomial redimensionado e deslocado com inflacao. #
#
# formula – formula do modelo ( $Y \sim X1 + X2$ )      #
# data – conjunto de dados                          #
# t – numero de componentes                          #
# m – tamanho da escala                             #
# d – deslocamento da escala                        #
# s – intervalo de valores na escala                #
# C – valor inflacionado (padrao  $m * s$ )              #
# tolerance – tolerancia na estimacao (padrao  $1e-6$ )   #
# maxiter – maximo de iteracoes (padrao 500)       #
# -----#
inf.binom.model = function(formula, data, t = 3, m = 5, d = 1,
  s = 2, C = m * s, tolerance = 1e-6, maxiter = 500,
  hessian = FALSE) {

  # Checagem da formula
  if (class(formula) != "formula")
    stop("'formula' must be a valid formula object.")
  if(attr(terms(formula), "response") != 1)
    stop("Model does not have a response.")

  # Monta o data frame de acordo com a formula
  df = model.frame(formula, data = data)

  # Separa respostas e covariaveis
  z = df[,1]
  X = as.matrix(df[,-1])

  # Adiciona intercepto se houver
  if(attr(terms(formula), "intercept") == 1) X = cbind(1, X)

  # Valores iniciais dos parametros

```

```

tau = rep(1/(t + 2), t + 1)
beta = matrix(0, nrow = t, ncol = ncol(X))
beta[,1] = seq(-2.5, 2.5, length.out = t)

# Variaveis para controle de convergencia
count = 0
loglik = -Inf
converged = TRUE

# Probabilidade de avaliacao ao acaso
p0 = rep(1/(m - d + 1), nrow(X))

# Probabilidades dos termos binomiais (matriz)
pr = matrix(0, nrow = nrow(X), ncol = t)

# Estimacao
repeat {
  # Calcula os componentes binomiais
  for (k in 1:t) {
    eta = X %*% beta[k,]
    pr[,k] = dbinom(x = z/s - d, size = m - d,
                    prob = exp(eta)/(1 + exp(eta)))
  }

  # Calcula as probabilidades
  pp = (1 - sum(tau)) * (z == C) + tau[t + 1] * p0
  for (k in 1:t) pp = pp + tau[k] * pr[,k]

  # Atualiza os pesos
  W = matrix(0, nrow = nrow(X), ncol = length(tau))
  for (k in 1:t) W[,k] = (tau[k] * pr[,k])/pp
  W[,t+1] = (tau[t + 1] * p0)/pp

  # Atualiza os parametros de mistura
  tau.old = tau
  tau = colMeans(W)
  tau.dt = max(abs(tau - tau.old))

  # Atualiza os parametros de regressao
  beta.old = beta
  for (k in 1:t) {
    eta = X %*% beta[k,]

```

```

    U = colSums(W[,k] * X * as.vector((((z/s - d) -
      (m - d) * exp(eta)/(1 + exp(eta))))))
    K = (t(sqrt(W[,k]) * X * as.vector((m - d) *
      exp(eta)/((exp(eta) + 1)^2))) %*%
      (sqrt(W[,k]) * X))

    beta[k,] = beta[k,] + as.vector(ginv(K) %*% U)
  }
  beta.dt = max(abs(beta - beta.old))

# Calcula a verossimilhanca
  loglik.old = loglik
  loglik = sum(log(pp))
  loglik.dt = abs(loglik - loglik.old)

# Contagem de iteracoes
  count = count + 1

# Verifica falta de convergencia
  converged = (count <= maxiter)
  if (!converged) break

# Regra de parada
  if (max(tau.dt, beta.dt) < tolerance) break
  if (loglik.dt < tolerance) break
}

# Calcula os escores
pp = (1 - sum(tau)) * (z == C) + tau[t + 1] * p0
for (k in 1:t) pp = pp + tau[k] * pr[,k]

U.T = matrix(0, nrow = nrow(X), ncol = t + 1)
for (r in 1:t) U.T[,r] = (pr[,r] - (z == C))/pp
U.T[, t + 1] = (p0 - (z == C))/pp

U.B = array(0, dim = c(t, nrow(X), ncol(X)))

for (r in 1:t) {
  eta      = X %*% beta[r,]
  temp.1 = (X * ((tau[r] * pr[,r])/pp))
  temp.2 = (z/s - d - ((m - d) * exp(eta))/(1 + exp(eta)))

```

```

    U.B[r , ,] = temp.1 * as.vector(temp.2)
  }

# Calcula a matriz hessiana
H = diag(t * (ncol(X) + 1) + 1)
if (hessian) {
  H[1:(t+1),1:(t+1)] = t(U.T) %*% U.T
  for (r in 1:t) {
    pos.r = (t + 1) + (r - 1)*ncol(X) + 1
    H[1:(t+1),pos.r:(pos.r + ncol(X) - 1)] =
      t(U.T) %*% U.B[r , ,]
    H[pos.r:(pos.r + ncol(X) - 1),1:(t+1)] =
      t(H[1:(t+1),pos.r:(pos.r + ncol(X) - 1)])

    for (s in 1:t) {
      pos.s = (t + 1) + (s - 1)*ncol(X) + 1
      H[pos.r:(pos.r + ncol(X) - 1),
        pos.s:(pos.s + ncol(X) - 1)] =
        t(U.B[r , ,]) %*% U.B[s , ,]
    }
  }
}

# Exporta os resultados
result = list(tau =tau, beta = beta, loglik = loglik,
              converged = converged, U.T = U.T, U.B = U.B)

if (hessian) {
  result$H = H
  result$positive.definite = all(eigen(H)$values > 0)
}
result
}

```

Modelo binomial duplo

```

# -----#
# Biblioteca usada para inverter a matriz de      #
# informacao de Fisher.                          #
# -----#
library(MASS)

# -----#
# Biblioteca usada para calcular a probabilidade no modelo #
# binomial duplo.                                  #
# -----#
library(rmutil)

# -----#
# Modelo binomial duplo.                          #
#
# formula - formula do modelo ( $Y \sim X1 + X2$ )    #
# data - conjunto de dados                        #
# t - numero de componentes                       #
# m - tamanho da escala                          #
# tolerance - tolerancia na estimacao (padrao 1e-4) #
# maxiter - maximo de iteracoes (padrao 500)     #
# min.phi - menor valor do parametro de variancia #
# -----#
dbinom.model = function(formula, data, t = 3, m = 5,
  tol.param = 1e-6, tol.like = 1e-4,
  maxiter = 100, hessian = FALSE) {

  # Checagem da formula
  if (class(formula) != "formula")
    stop("'formula' must be a valid formula object.")
  if(attr(terms(formula), "response") != 1)
    stop("Model does not have a response.")

  # Monta o data frame de acordo com a formula
  df = model.frame(formula, data = data)

  # Separa respostas e covariaveis
  Z = df[,1]
  X = as.matrix(df[,-1])

```

```

# Adiciona intercepto se houver
if(attr(terms(formula), "intercept") == 1) X = cbind(1, X)

# Valores iniciais dos parametros
alpha = rep(1/(t + 1), t)
theta = rep(1, t)
beta = matrix(0, nrow = t, ncol = ncol(X))
beta[,1] = seq(-1.5, 1.5, length.out = t)

# Variaveis para controle de convergencia
count = 0
loglik = -Inf
converged = TRUE

# Probabilidade de avaliacao ao acaso
p0 = rep(1/(m + 1), nrow(X))

# Probabilidade dos termos binomiais (matriz)
pr = matrix(0, nrow = nrow(X), ncol = t)

# Estimacao
repeat {
  loglik.old = loglik

  # Loop interno
  repeat {
    # Calcula os componentes binomiais
    for (r in 1:t) {
      eta = X %*% beta[r,]
      pr[,r] = ddoublebinom(y = Z, size = m,
        m = exp(eta)/(1 + exp(eta)),
        s = theta[r])
    }

    # Calcula as probabilidades
    pp = (1 - sum(alpha)) * p0
    for (r in 1:t) pp = pp + alpha[r] * pr[,r]

    # Atualiza os pesos
    W = matrix(0, nrow = nrow(X), ncol = length(alpha))
    for (r in 1:t) W[,r] = (alpha[r] * pr[,r])/pp
  }
}

```

```

# Atualiza a verossimilhanca (passo anterior)
loglik = sum(log(pp))

# Atualiza os parametros de regressao
beta.old = beta
for (r in 1:t) {
  eta = X %*% beta[r,]

  U = colSums(W[,r] * X *
    as.vector((Z - m * exp(eta)/(1 + exp(eta))))))
  K = (t(sqrt(W[,r]) * X *
    as.vector(m * exp(eta)/((exp(eta) + 1)^2))) %*%
    (sqrt(W[,r]) * X))
  beta[r,] = beta[r,] + as.vector(ginv(K) %*% U)
}
beta.dt = max(abs(beta - beta.old))

# Atualiza os parametros de mistura
alpha.old = alpha
alpha = colMeans(W)
alpha.dt = max(abs(alpha - alpha.old))

# Regra de parada interna
if(max(alpha.dt, beta.dt) < tol.param) break
}

# Calcula o coeficiente de variancia
theta.old = theta
for (k in 1:t) {
  eta = X %*% beta[k,]
  theta[k] = (sum(W[,k])/2) /
    (sum(W[,k] * (log(((Z/m)^Z) * ((1 - Z/m)^(m - Z))) -
      Z * eta + m * log(1 + exp(eta))))))
}

# Calculo para controle de convergencia
theta.dt = max(abs(theta - theta.old))
loglik.dt = abs(loglik - loglik.old)

# Contagem de iteracoes
count = count + 1

```

```

# Verifica falta de convergencia
converged = (count <= maxiter)
if (!converged) break

# Regra de parada
if (loglik.dt < tol.like) break
}

# Calcula os escores
pp = (1 - sum(alpha)) * p0
for (k in 1:t) pp = pp + alpha[k] * pr[,k]

U.A = matrix(0, nrow = nrow(X), ncol = t)
for (r in 1:t) U.A[,r] = (pr[,r] - p0)/pp

U.T = matrix(0, nrow = nrow(X), ncol = t)
for (r in 1:t) {
  eta = X %*% beta[r,]

  U.T[,r] = ((alpha[r] * pr[,r])/pp) *
    (1/(2 * theta[r]) - (log(((Z/m)^Z) *
      ((1 - Z/m)^(m - Z)))) -
      Z * eta + m * log(1 + exp(eta)))
}

U.B = array(0, dim = c(t, nrow(X), ncol(X)))

for (r in 1:t) {
  eta = X %*% beta[r,]
  temp.1 = (X * ((alpha[r] * pr[,r])/pp))
  temp.2 = (Z - m * exp(eta)/(1 + exp(eta)))

  U.B[r,,] = temp.1 * as.vector(temp.2)
}

# Calcula a matriz hessiana
H = diag((ncol(X) + 2) * t)
if (hessian) {
  H[1:t,1:t] = t(U.A) %*% U.A
  H[(t+1):(2*t),(t+1):(2*t)] = t(U.T) %*% U.T

  for (r in 1:t) {

```

```
        pos.r = 2*t + (r - 1)*ncol(X) + 1
        for (s in 1:t) {
            pos.s = 2*t + (s - 1)*ncol(X) + 1
            H[pos.r:(pos.r + ncol(X) - 1),
              pos.s:(pos.s + ncol(X) - 1)] =
              t(U.B[r, ,]) %*% U.B[s, ,]
        }
    }
}

# Expora os resultados
result = list(alpha = alpha, theta = theta, beta = beta,
              loglik = loglik, converged = converged,
              U.A = U.A, U.B = U.B, U.T = U.T)

if (hessian) {
    result$H = H
    result$positive.definite = all(eigen(H)$values > 0)
}
result
}
```

Modelo binomial multiplicativo

```

# -----#
# Biblioteca usada para inverter a matriz de          #
# informacao de Fisher.                               #
# -----#
library(MASS)

# -----#
# Biblioteca usada para calcular a probabilidade no modelo #
# binomial multiplicativo.                            #
# -----#
library(rmutil)

# -----#
# Modelo binomial multiplicativo.                     #
#                                                     #
# formula - formula do modelo ( $Y \sim X1 + X2$ )      #
# data - conjunto de dados                            #
# t - numero de componentes                           #
# m - tamanho da escala                              #
# tolerance - tolerancia na estimacao (padrao  $1e-4$ )  #
# maxiter - maximo de iteracoes (padrao 500)         #
# min.phi - menor valor do parametro de variancia    #
# -----#
mbinom.model = function(formula, data, t = 3, m = 5,
  tol.param = 1e-6, tol.like = 1e-4, maxiter = 100,
  hessian = FALSE) {

  # Checagem da formula
  if (class(formula) != "formula")
    stop("'formula' must be a valid formula object.")
  if(attr(terms(formula), "response") != 1)
    stop("Model does not have a response.")

  # Monta o data frame de acordo com a formula
  df = model.frame(formula, data = data)

  # Separa resposta e covariaveis
  Z = df[,1]
  X = as.matrix(df[,-1])

```

```

# Variaveis auxiliares
zmz = (0:m) * (m - 0:m)
zmz2 = zmz^2
z = (0:m)
z2 = z^2

# Adiciona intercepto se houver
if(attr(terms(formula), "intercept") == 1) X = cbind(1, X)

# Valores iniciais dos parametros
alpha = rep(1/(t + 1), t)
phi = rep(0, t)
beta = matrix(runif(t * ncol(X), min = -.01, max = .01),
              nrow = t, ncol = ncol(X))
beta[,1] = seq(-1.5, 1.5, length.out = t)

# Variaveis para controle de convergencia
count = 0
loglik = -Inf
converged = TRUE

# Probabilidade de avaliacao ao acaso
p0 = rep(1/(m + 1), nrow(X))

# Probabilidade dos termos binomiais (matriz)
pr = matrix(0, nrow = nrow(X), ncol = t)

# Estimacao
repeat {
  # Calcula os componentes binomiais
  for (r in 1:t) {
    eta = X %*% beta[r,]
    dpp = apply(X = eta, MARGIN = 1, FUN = function(x)
               exp(log(choose(m, z)) + z * x - m * log(1 + exp(x)) + zmz)
    )
    dpp = apply(X = t(dpp), MARGIN = 1, FUN = sum)

    pr[,r] = exp(log(choose(m, Z)) + Z * eta -
                 m * log(1 + exp(eta)) + Z * (m - Z) * phi[r])/dpp
  }
}

# Calcula as probabilidades
pp = (1 - sum(alpha)) * p0

```

```

for (r in 1:t) pp = pp + alpha[r] * pr[,r]

# Atualiza os pesos
W = matrix(0, nrow = nrow(X), ncol = length(alpha))
for (r in 1:t) W[,r] = (alpha[r] * pr[,r])/pp

# Calcula a verossimilhanca
loglik.old = loglik
loglik = sum(log(pp))
loglik.dt = abs(loglik - loglik.old)

# Atualiza conjuntamente os parametros de regressao
# e variancia
beta.old = beta
phi.old = phi
for (r in 1:t) {
  eta = X %*% beta[r,]

  dpp = apply(X = eta, MARGIN = 1, FUN = function(x)
    exp(log(choose(m, z)) + z * x -
      m * log(1 + exp(x)) + zmz * phi[r]))

  dpp = t(dpp)/apply(X = t(dpp), MARGIN = 1, FUN = sum)

  EZ = apply(dpp, MARGIN = 1,
    function(prob) sum(z * prob))
  EZ2 = apply(dpp, MARGIN = 1,
    function(prob) sum(z2 * prob))
  EZmZ = apply(dpp, MARGIN = 1,
    function(prob) sum(zmz * prob))
  EZmZ2 = apply(dpp, MARGIN = 1,
    function(prob) sum(zmz2 * prob))

  VZ = (EZ2 - EZ^2)
  VZmZ = (EZmZ2 - EZmZ^2)

  U.phi = sum(W[,r] * as.vector(Z * (m - Z) - EZmZ))
  K.phi = sum(W[,r] * as.vector(VZmZ))

  U.beta = colSums(W[,r] * X * as.vector(Z - EZ))
  K.beta = (t(sqrt(W[,r]) * X * as.vector(VZ)) %*%
    (sqrt(W[,r]) * X))

```

```

        beta[r,] = beta[r,] + as.vector(ginv(K.phi) %% U.phi)
        phi[r] = phi[r] + U.phi/K.phi
    }
    beta.dt = max(abs(beta - beta.old))
    phi.dt = max(abs(phi - phi.old))

# Atualiza os parametros de mistura
alpha.old = alpha
alpha = colMeans(W)
alpha.dt = max(abs(alpha - alpha.old))

# Contagem de iteracoes
count = count + 1

# Verifica falta de convergencia
converged = (count <= maxiter)
if (!converged) break

# Regras de parada
if(max(alpha.dt, beta.dt, phi.dt) < tol.param) break
if (loglik.dt < tol.like) break
}

# Calcula os escores
pp = (1 - sum(alpha)) * p0
for (k in 1:t) pp = pp + alpha[k] * pr[,k]

U.A = matrix(0, nrow = nrow(X), ncol = t)
for (r in 1:t) U.A[,r] = (pr[,r] - p0)/pp

U.B = array(0, dim = c(t, nrow(X), ncol(X)))
U.T = matrix(0, nrow = nrow(X), ncol = t)
for (r in 1:t) {
    eta = X %% beta[r,]
    dpp = apply(X = eta, MARGIN = 1, FUN = function(x)
        exp(log(choose(m, z)) + z * x - m * log(1 + exp(x)) +
            zmz * phi[r]))
    dpp = t(dpp)/apply(X = t(dpp), MARGIN = 1, FUN = sum)
    EZ = apply(dpp, MARGIN = 1,

```

```

        function(prob) sum((0:m) * prob))
EZmZ = apply(dpp, MARGIN = 1,
            function(prob) sum(((0:m) * (m - 0:m)) * prob))

U.T[,r] = ((alpha[r] * pr[,r])/pp) * (Z * (m - Z) - EZmZ)

temp.1 = (X * ((alpha[r] * pr[,r])/pp))
temp.2 = Z - EZ

U.B[r,,] = temp.1 * as.vector(temp.2)
}

# Calcula a matriz hessiana
H = diag((ncol(X) + 2) * t)
if (hessian) {
  H[1:t,1:t] = t(U.A) %% U.A
  H[1:t,(t+1):(2*t)] = t(U.A) %% U.T
  H[(t+1):(2*t),1:t] = t(H[1:t,(t+1):(2*t)])
  H[(t+1):(2*t),(t+1):(2*t)] = t(U.T) %% U.T

  for (r in 1:t) {
    pos.r = 2*t + (r - 1)*ncol(X) + 1

    H[1:t,pos.r:(pos.r + ncol(X) - 1)] =
      t(U.A) %% U.B[r,,]
    H[pos.r:(pos.r + ncol(X) - 1),1:t] =
      t(H[1:t,pos.r:(pos.r + ncol(X) - 1)])

    H[(t+1):(2*t),pos.r:(pos.r + ncol(X) - 1)] =
      t(U.T) %% U.B[r,,]
    H[pos.r:(pos.r + ncol(X) - 1),(t+1):(2*t)] =
      t(H[(t+1):(2*t),pos.r:(pos.r + ncol(X) - 1)])

    for (s in 1:t) {
      pos.s = 2*t + (s - 1)*ncol(X) + 1
      H[pos.r:(pos.r + ncol(X) - 1),
        pos.s:(pos.s + ncol(X) - 1)] =
        t(U.B[r,,]) %% U.B[s,,]
    }
  }
}
}

```

```
result = list(alpha = alpha, phi = phi, beta = beta, loglik = loglik,
              converged = converged, U.A = U.A, U.B = U.B, U.T = U.T)

if (hessian) {
  result$H = H
  result$positive.definite = all(eigen(H)$values > 0)
}
result
}
```