
ESTIMAÇÃO DO POSTO DA MATRIZ DOS PARÂMETROS DO MODELO DE REGRESSÃO DIRICHLET

TATIANE FERREIRA DO NASCIMENTO MELO DA SILVA

Orientador: Prof. Dr. Klaus Leite Pinto Vasconcellos

Área de concentração: Estatística Matemática

Dissertação submetida como requerimento parcial para obtenção do grau de Mestre em Estatística pela Universidade Federal de Pernambuco

Recife/PE
novembro, 2004

Universidade Federal de Pernambuco
Mestrado em Estatística

26 de novembro de 2004

(data)

Nós recomendamos que a dissertação de mestrado de autoria de

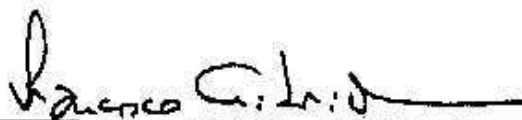
Tatiane Ferreira do Nascimento Melo

Intitulada

Estimação do posto da matriz dos parâmetros do modelo de regressão

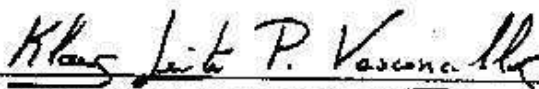
Dirichlet

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



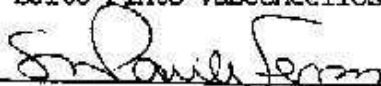
Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

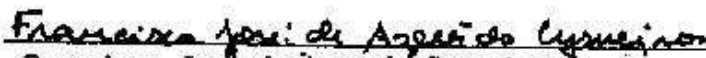


Klaus Leite Pinto Vasconcellos

orientador



Silvia Lopes de Paula Ferrari (USP)



Francisco José de Azevedo Cysneiros



Prof. Francisco C. L. Neto
Coordenador de Mestrado
em Estatística da UFPE

Este documento será anexado à versão final da dissertação.

Dedico este trabalho aos meus pais, Pedro e Rosângela, aos meus irmãos, Cristiano e Patricia, e ao meu esposo, Alessandro.

Agradecimentos

A DEUS, que é minha base, meu refúgio, minha certeza, Aquele em que tudo posso, e tudo confio. Pois sem Ele nada seria possível e tudo seria vazio.

Ao meu orientador, Klaus Leite Pinto Vasconcellos, pela atenção, pela paciência, pela competência e por ter acreditado na minha capacidade.

Aos meus pais e aos meus irmãos pelo apoio, pelo amor, pela força, por nunca deixarem eu desistir e pelas palavras de conforto.

Ao meu esposo pela compreensão, pelo amor incondicional, pelo carinho e por abdicar do nosso tempo juntos para que mais essa etapa em minha vida fosse concluída.

Ao professor Francisco Cribari-Neto, pela competência, pela seriedade, pelo profissionalismo, pelo incentivo e pela confiança.

A toda família do Alessandro que sempre esteve do nosso lado quando decidimos vir para Recife.

À tia Santina pelo carinho, pela força e por participar de todas as fases de minha vida.

À Fátima de Sousa Maior pela amizade, pelo apoio, pela força e confiança.

Às minhas grandes amigas: Ana Rosa, Flávia, Karla, Márcia, Sigreice e Tatiana pela força, pelo apoio e pela confiança. Ao meu amigo Cleyton pela força.

Aos meus amigos de mestrado de que nunca vou esquecer. Às pessoas com que pude compartilhar esses dois anos, pessoas que sempre me ajudaram a conhecer a Estatística mais de perto e que me receberam com muito carinho nesta nova área de estudo, que para mim era até então desconhecida.

Ao André pela amizade, pelos momentos de descontrações, pela alegria e pelas sugestões.

À Andréa pela amizade fortalecida no segundo ano de mestrado, por deixar eu conhecer a pessoa meiga, simples e doce que é.

Ao Braga Junior pela amizade, pela paciência, pela alegria e por sempre me ajudar com suas sugestões.

À Gecy pela amizade, pela alegria em que vê a vida, por sempre me ajudar neste mestrado, pelo apoio, pelo entusiasmo, otimismo e coragem.

À Renata pela amizade, pelo alto astral, por sempre me ouvir nos momentos tristes e alegres, pelo carinho, pela alegria e empolgação nos estudos.

À minha amiga Sandra Maria, pela serenidade, centralidade, pelo exemplo de vida, pela paciência e por sempre me ajudar nas dúvidas estatísticas.

À Sandra Rêgo pela amizade, paciência, confiança, perseverança, por me permitir fazer parte de sua vida e me ajudar nos estudos.

Ao meu colega Cherubino (em memória) pelas idéias compartilhadas.

Aos meus colegas de mestrado, Artur, Camilo, Carlos Gadelha, Carlos Tomé, Daniela, Denis, Francisco, Hernando, Katya, Lenaldo, Milena, Polyane, Themis e Tiago.

À Keila Mara pelo incentivo para estudar na Universidade Federal de Pernambuco.

À Tatiane pela paciência, amizade e atenção com que me recebeu em Recife.

Ao Gilson, Patrícia Leal, Moisés e Silvia pela amizade e respeito.

Aos professores Jesus Carlos da Mota e Maurílio Márcio Melo da Universidade Federal de Goiás pelo incentivo a estudar Estatística.

Às professoras Maria Cristina Falcão Raposo, Viviana Giampaoli e Cláudia Regina pelos momentos de descontração e aprendizado proporcionados.

Aos professores do Programa de Mestrado em Estatística da UFPE.

À Valéria Bittencourt, pela competência com que sempre cuidou dos assuntos relacionados a pós-graduação, pela compreensão e por sempre me ouvir nos momentos tristes e alegres.

A todos que direta ou indiretamente colaboraram para que mais esta etapa de minha vida fosse concluída.

Aos participantes da banca examinadora pelas sugestões.

À CAPES, pelo apoio financeiro.

Resumo

O modelo de regressão Dirichlet é útil, por exemplo, na modelagem de taxas e proporções, onde a soma das componentes de cada vetor de observações é igual a um. Os coeficientes deste modelo de regressão constituem uma matriz. Se esta matriz não tem posto completo, ou seja, se alguns de seus elementos podem ser escritos como combinações lineares de outros, então a quantidade de parâmetros do modelo a serem estimados é menor. Nosso objetivo é estimar o posto desta matriz de parâmetros, utilizando uma estatística de teste proposta por Ratsimalahelo (2003), através do procedimento de teste seqüencial e dos critérios de informação BIC (*Bayesian Information Criteria*) e HQIC (*Hannan Quinn Information Criteria*). Em seguida, avaliamos o desempenho dos estimadores do posto da matriz de coeficientes, baseados nestes procedimentos.

Neste trabalho consideramos dois modelos de regressão Dirichlet. Através dos resultados de simulação de Monte Carlo, observamos que quando utilizamos o procedimento de teste seqüencial, para estimar o posto da matriz de coeficientes dos modelos de regressão Dirichlet, o desempenho dos estimadores, em geral, é melhor em termos de viés e erro quadrático médio do que quando utilizamos os critérios de informação BIC e HQIC.

Abstract

The Dirichlet regression model is useful in the modelling of rates and proportions, where the sum of the components of each vector of observations is one. The coefficients of this regression model constitute a matrix. If this matrix does not have full rank, in other words, if some of its rows or columns can be written as linear combinations of others, then the quantity of estimated parameters of the model is reduced. Our goal is to first estimate the rank of this matrix of parameters, using a test statistic proposed by Ratsimalahelo (2003), via the sequential testing strategy, the Bayesian information criterion (BIC), and Hannan Quinn information criterion (HQIC). We then evaluate the performance of the estimator of the rank of the coefficient matrix based on these procedures.

In this work we consider two Dirichlet regression models. From the Monte Carlo simulation results we observe that, when we use the sequential test procedure to estimate the rank of the coefficient matrix of the Dirichlet regression models, the performance of these estimators is generally better in terms of bias and mean square error than those of the information criteria BIC and HQIC.

Índice

1.	Introdução	1
1.1.	Introdução	1
1.1.1.	Organização da dissertação	2
1.1.2.	Suporte computacional	2
1.2.	Estimadores de máxima verossimilhança	3
2.	O modelo de regressão Dirichlet	4
2.1.	Modelo beta	4
2.1.1.	Definição e estimação	4
2.2.	Modelo Dirichlet	11
2.2.1.	Definição e estimação	11
2.2.2.	Modelo 1	18
2.2.3.	Modelo 2	25
3.	Estimador do posto	27
3.1.	Considerações iniciais	27
3.1.1.	Teste de hipótese	28
3.1.2.	Decomposição em valores singulares e a estatística de teste	29
3.2.	Procedimento de teste seqüencial	31
3.3.	Critérios de informação BIC e HQIC	32
4.	Avaliação numérica	33
4.1.	Resultados numéricos para o Modelo 1	34

4.2.	Resultados numéricos para o Modelo 2	54
5.	Aplicação	80
5.1.	Dados do pólen fossilizado em flores	80
5.2.	Dados da composição de leucócitos no sangue	81
6.	Conclusões	85
▷	Apêndice 1	86
▷	Apêndice 2	114
▷	Apêndice 3	122
▷	Apêndice 4	124
▷	Apêndice 5	124
A5.1.	Programas principais	124
A5.2.	Biblioteca de funções	135
A5.3.	Programas de ligação	150
▷	Referências bibliográficas	151

Capítulo 1

Introdução

1.1 Introdução

Neste trabalho estudamos um estimador para o posto da matriz dos coeficientes de uma regressão em que as observações seguem uma distribuição Dirichlet, que é uma extensão multivariada da distribuição beta. Esta última distribuição é útil na modelagem de dados estatísticos que assumem valores no intervalo $(0, 1)$, enquanto que a primeira se presta à modelagem probabilística de um vetor de variáveis compreendidas no intervalo $(0, 1)$ cuja soma é igual a um. Com o objetivo de definir uma estrutura de regressão para dados com distribuição Dirichlet, consideramos duas parametrizações, que denotaremos por “modelo 1” e “modelo 2”. Na primeira parametrização, a média da variável resposta é modelada juntamente com um parâmetro de precisão, ou seja, a j -ésima componente, μ_j , da resposta média é modelada através de uma estrutura de regressão juntamente com um parâmetro de precisão ϕ . Na segunda parametrização a variável resposta se distribui como Dirichlet com parâmetros $\alpha_1, \alpha_2, \dots, \alpha_p$, onde estes parâmetros obedecem estrutura de regressão.

Os coeficientes na estrutura de regressão multivariada do modelo de regressão Dirichlet constituem uma matriz, que denotaremos por B . Quando essa matriz não tem posto completo, algumas de suas colunas (ou linhas) são linearmente dependentes, ou seja, alguns de seus parâmetros podem ser escritos como combinações lineares de outros. Pensando nisso, nosso objetivo principal é estimar o posto da matriz dos coeficientes nos dois modelos de regressão Dirichlet considerados. Vários testes foram propostos para estimação do posto de matrizes desconhecidas. Em particular, Ratsimalahelo (2003) desenvolveu um teste usando a teoria da perturbação de matrizes. O teste consiste em determinar quantos valores singulares de uma matriz estimada são significantemente diferentes de zero, já que o posto de uma matriz é igual ao número de valores singulares diferentes de zero. Outros pesquisadores, como Stewart (1984), Gill e Lewbel (1992), Robin e Smith (2000) também desenvolveram testes para posto de matrizes, mas o teste desenvolvido por Ratsimalahelo (2003) tem uma grande vantagem sobre os outros testes, a facilidade de seu cálculo.

Na estimação do posto da matriz B usamos a estatística de teste proposta por Ratsimalahelo (2003); associados a esta estatística de teste, usamos três procedimentos para estimação do posto, a saber: o teste seqüencial e os critérios de informação BIC (*Bayesian Information Criteria*) e HQIC (*Hannan Quinn Information Criteria*). Usando estes procedimentos, Ratsimalahelo (2003) estabelece um estimador fortemente consistente para o posto de B , a partir de um estimador fortemente consistente de B . Para os resultados numéricos, utilizamos o método de simulação Monte Carlo. Com tais resultados avaliamos o desempenho numérico do teste seqüencial e dos critérios de informação BIC e HQIC para as duas parametrizações consideradas quando a matriz B tem ou não posto completo. Além disso, aplicamos o segundo modelo de regressão Dirichlet considerado a dois conjuntos de dados

reais e estimamos o posto da matriz de parâmetros desse modelo usando a estatística de teste de Ratsimalahelo (2003) juntamente com os três procedimentos citados acima.

1.1.1 Organização da dissertação

A estrutura do restante desta dissertação é composta da seguinte maneira: no capítulo 2 apresentamos as distribuições beta e Dirichlet, algumas propriedades amostrais de ambas e consideramos ainda dois modelos de regressão Dirichlet, mostrando como obter as estimativas dos parâmetros desconhecidos. No capítulo 3, apresentamos a estatística de teste proposta por Ratsimalahelo (2003) juntamente com o procedimento de teste seqüencial e os critérios de informação BIC e HQIC para estimação do posto da matriz de coeficientes da regressão Dirichlet. No quarto capítulo, com base em simulações de Monte Carlo para os modelos de regressão Dirichlet, investigamos o comportamento do procedimento de teste seqüencial e dos critérios de informação BIC e HQIC na estimação do posto. No capítulo 5, aplicamos o modelo de regressão Dirichlet a dois conjuntos de dados reais e estimamos o posto da matriz de coeficientes do modelo. As conclusões desta dissertação são apresentadas no capítulo 6.

1.1.2 Suporte computacional

As ferramentas computacionais utilizadas para o desenvolvimento deste trabalho foram a linguagem de programação `Ox` e o programa `R`.

Para as avaliações numéricas, utilizamos a linguagem de programação matricial `Ox`, criada por Jurgen Doornik em 1994, que permite a implementação de técnicas estatísticas com facilidade e atende a requisitos como precisão e eficiência, o que tem contribuído para sua ampla utilização no campo da computação numérica. A versão utilizada aqui foi a 3.30 para sistemas operacionais Linux e está disponível gratuitamente para uso acadêmico no endereço <http://www.nuff.ox.ac.uk/Users/Doornik>. Detalhes sobre esta linguagem de programação podem ser encontrados em Doornik (2001) e Cribari-Neto e Zarkos (2003). Os gráficos apresentados neste trabalho foram feitos utilizando o programa `R`, que é baseado em uma linguagem de alto nível desenvolvida para análise, manipulação e apresentação gráfica de dados. Ver Cribari-Neto e Zarkos (1999) para detalhes.

1.2 Estimadores de máxima verossimilhança

Suponha que x é o valor observado de um vetor aleatório $X = (X_1, \dots, X_N)^\top$ cuja distribuição é caracterizada por uma função de probabilidade ou densidade com forma analítica $f(x; \theta)$ conhecida, mas dependente de um vetor $\theta = (\theta_1, \dots, \theta_p)^\top$ de parâmetros desconhecidos. Seja $\Theta \subset \mathbb{R}^p$ o espaço paramétrico, que representa o conjunto de valores possíveis para o vetor θ .

A função de verossimilhança $L(\theta; x)$ é definida como sendo igual à função do modelo, embora seja interpretada diferentemente, como função de θ . Assim $L(\theta; x) = f(x; \theta)$. Usualmente, trabalha-se com a log-verossimilhança $\ell(\theta; x) = \log L(\theta; x)$. A função de verossimilhança informa a ordem natural de preferência entre diversas possibilidades para θ . Entre os possíveis candidatos para estimar o parâmetro verdadeiro, θ_0 , a partir dos mesmos dados x , o vetor de parâmetros mais plausível é aquele de maior verossimilhança. Neste sentido, o método de máxima verossimilhança objetiva escolher o valor do vetor θ de parâmetros que fornece a chance maior de ocorrer novamente os mesmos dados que ocorreram. Assim, para estimar o vetor verdadeiro θ_0 de parâmetros, escolhe-se aquele vetor de parâmetros que maximiza a função de verossimilhança no espaço paramétrico Θ . O estimador de máxima verossimilhança (EMV) de θ é portanto o vetor que maximiza $L(\theta; x)$ em Θ . Como a função logaritmo é monótona crescente, maximizar $L(\theta; x)$ e $\ell(\theta; x)$ em Θ são processos equivalentes. Então o EMV $\hat{\theta}$ é definido como o vetor

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} L(\theta; x) = \operatorname{argmax}_{\theta \in \Theta} \ell(\theta; x).$$

Se Θ é um conjunto discreto, calcula-se $\ell(\theta; x)$ para diversos θ 's e escolhe-se $\hat{\theta}$ como aquele valor de θ correspondente ao máximo $\ell(\theta; x)$. Quando $\ell(\theta; x)$ é contínua e diferenciável em Θ , o EMV $\hat{\theta}$ pode ser obtido resolvendo-se o sistema de equações simultâneas

$$\frac{\partial \ell(\theta; x)}{\partial \theta_j} = 0, \quad \text{com } j = 1, \dots, p$$

desde que θ não se encontre na fronteira do espaço paramétrico. A primeira derivada da função log-verossimilhança $U(\theta) = \frac{\partial \ell(\theta; x)}{\partial \theta}$ é chamada função escore. As equações de máxima verossimilhança são usualmente não-lineares e nestes casos as soluções de $U(\hat{\theta}) = 0$ devem ser obtidas por técnicas iterativas, como as apresentadas no apêndice 3.

Os estimadores de máxima verossimilhança possuem algumas propriedades bem conhecidas, ver Cordeiro (1992). Uma propriedade importante é que, sob condições bastante gerais, os EMV's são assintoticamente normais, ou seja,

$$\sqrt{N}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, I(\theta_0)^{-1}),$$

onde θ_0 é o valor verdadeiro do parâmetro, o símbolo $\xrightarrow{d}$ significa convergência em distribuição e $I(\theta_0)$ é a informação de Fisher,

$$I(\theta) \equiv E_\theta \left(\frac{\partial \ell(\theta; x)}{\partial \theta} \frac{\partial \ell(\theta; x)}{\partial \theta^\top} \right)$$

Esta propriedade será utilizada em nosso trabalho.

Capítulo 2

O modelo de regressão Dirichlet

2.1 Modelo beta

Na prática, o uso de modelos de regressão é bastante comum para analisar dados que possivelmente têm algum tipo de relação com outras variáveis. O modelo mais utilizado é o de regressão normal linear, mas este não é apropriado para situações onde a variável resposta está restrita ao intervalo $(0, 1)$. Uma solução é usar um modelo de regressão onde a resposta é modelada por uma distribuição beta, usando, por exemplo, uma parametrização definida pela média e por um parâmetro de dispersão, ver Ferrari & Cribari-Neto (2004). A distribuição beta tem sido bastante utilizada para se definir modelos probabilísticos. Bury (1999) lista um conjunto de aplicações da distribuição beta em engenharia. Janardan & Padmanabhan (1986) modelam variáveis hidrológicas usando a distribuição beta. A seguir, consideramos algumas propriedades da distribuição beta, como sua função de densidade e seus primeiros momentos.

2.1.1. Definição e estimação

Seja Y uma variável aleatória. Dizemos que Y tem distribuição beta com parâmetros $\alpha_1, \alpha_2 > 0$, denotado por $Y \sim \mathcal{B}(\alpha_1, \alpha_2)$, se sua função de densidade é dada por

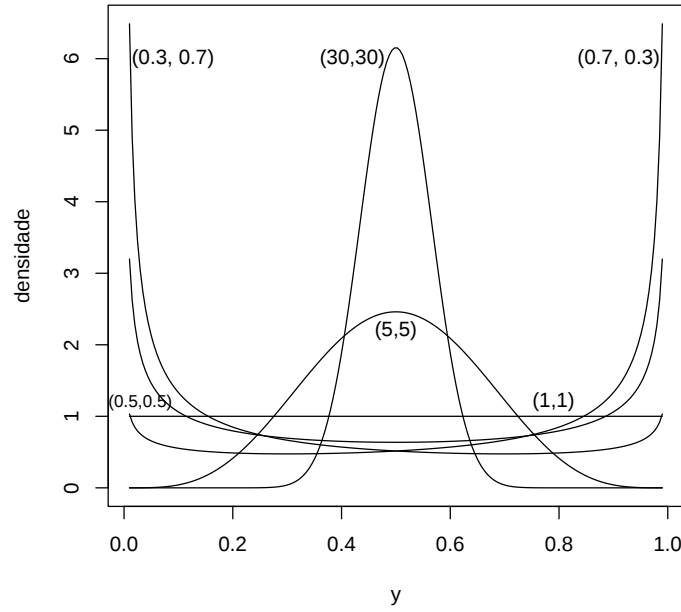
$$f(y; \alpha_1, \alpha_2) = \begin{cases} \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y^{\alpha_1 - 1} (1 - y)^{\alpha_2 - 1} & \text{se } 0 < y < 1, \\ 0 & \text{caso contrário,} \end{cases} \quad (2.1.1.1)$$

onde $\Gamma(\cdot)$ é a função gama.

Os parâmetros α_1 e α_2 são parâmetros de ajuste, pois através de α_1 e α_2 podem ser obtidas diferentes distribuições em $(0, 1)$. Quando $\alpha_1 = \alpha_2 = 0.5$, a distribuição beta se reduz à chamada “distribuição arco seno”, a qual é útil no estudo de passeios aleatórios. Distribuições beta para as quais $\alpha_1 + \alpha_2 = 1$ com $\alpha_1 \neq 0.5$ são conhecidas como distribuições “arco seno generalizadas”. Se $\alpha_1 = \alpha_2$ obtemos densidades simétricas em torno de 0.5 e quando $\alpha_1 = \alpha_2 = 1$ temos a distribuição uniforme. Logo, a distribuição beta é uma família de distribuições, como pode ser visto na Figura 2.1.1.1.

Para estudar o comportamento da distribuição beta, temos que conhecer algumas propriedades da distribuição como, por exemplo, a média e a variância. O n -ésimo momento em torno de zero da distribuição $\mathcal{B}(\alpha_1, \alpha_2)$ é

Figura 2.1.1.1: Gráfico da densidade beta para diferentes valores de α_1 e α_2 .



$$\begin{aligned}
 E(Y^n) &= \int_0^1 y^n f(y; \alpha_1, \alpha_2) dy \\
 &= \int_0^1 y^n \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y^{\alpha_1-1} (1-y)^{\alpha_2-1} dy \\
 &= \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \left[\int_0^1 y^{n+\alpha_1-1} (1-y)^{\alpha_2-1} dy \right].
 \end{aligned}$$

A integral entre colchetes acima é facilmente obtida usando-se a função beta definida por

$$B(s, r) = \frac{1}{a^{s+r-1}} \int_0^a z^{s-1} (a-z)^{r-1} dz, \quad \forall a \in (0, 1)$$

ou ainda,

$$\int_0^a z^{s-1} (a-z)^{r-1} dz = B(s, r) a^{s+r-1}. \quad (2.1.1.2)$$

Fazendo $a = 1$, $s = n + \alpha_1$, $r = \alpha_2$ e $z = y$ em (2.1.1.2) temos que

$$\int_0^1 y^{n+\alpha_1-1} (1-y)^{\alpha_2-1} dy = B(n + \alpha_1, \alpha_2).$$

Tendo em vista a propriedade da função Beta,

$$B(s, r) = \frac{\Gamma(s)\Gamma(r)}{\Gamma(s+r)}, \quad (2.1.1.3)$$

temos

$$\int_0^1 y^{n+\alpha_1-1} (1-y)^{\alpha_2-1} dy = \frac{\Gamma(n+\alpha_1)\Gamma(\alpha_2)}{\Gamma(n+\alpha_1+\alpha_2)},$$

portanto

$$\begin{aligned} E(Y^n) &= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{\Gamma(n+\alpha_1)\Gamma(\alpha_2)}{\Gamma(n+\alpha_1+\alpha_2)} \\ &= \frac{\Gamma(\alpha_1+\alpha_2)\Gamma(n+\alpha_1)}{\Gamma(\alpha_1)\Gamma(n+\alpha_1+\alpha_2)} \end{aligned}$$

A expressão acima pode ser simplificada aplicando-se uma propriedade muito conhecida da função gamma, a de que $\Gamma(z+1) = z\Gamma(z)$ temos

$$\begin{aligned} \Gamma(z+n) &= \Gamma[(z+(n-1))+1] \\ &= (z+(n-1))\Gamma[z+(n-1)] \\ &= (z+(n-1))\Gamma[(z+(n-2))+1] \\ &= (z+(n-1))(z+(n-2))\Gamma[z+(n-2)] \\ &\quad \dots \\ &= (z+(n-1))(z+(n-2))\dots(z+1)\Gamma(z+1) \\ &= (z+(n-1))(z+(n-2))\dots(z+1)z\Gamma(z), \end{aligned} \quad (2.1.1.4)$$

Então usando (2.1.1.4) temos

$$E(Y^n) = \frac{\alpha_1(\alpha_1+1)(\alpha_1+2)\dots(\alpha_1+n-1)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)(\alpha_1+\alpha_2+2)\dots(\alpha_1+\alpha_2+n-1)}.$$

Logo,

$$E(Y) = \frac{\alpha_1}{\alpha_1+\alpha_2},$$

$$E(Y^2) = \frac{\alpha_1(\alpha_1+1)}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)}.$$

Então , a variância é

$$\begin{aligned} \text{Var}(Y) &= E(Y^2) - [E(Y)]^2 \\ &= \frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}. \end{aligned}$$

Podemos observar na Figura 2.1.1.1 que, quando α_1 e α_2 aumentam, a variabilidade da distribuição diminui, o que é confirmado pela expressão da variância.

Para obter uma estrutura de regressão que permita modelar a média de uma variável resposta sujeita a um parâmetro de precisão, usamos uma parametrização diferente da mostrada na densidade (2.1.1.1). Seja $\mu = \alpha_1/(\alpha_1 + \alpha_2)$ e $\phi = \alpha_1 + \alpha_2$, ou seja, $\alpha_1 = \mu\phi$ e $\alpha_2 = (1 - \mu)\phi$. A função de densidade de Y pode ser reescrita, na nova parametrização, como

$$f(y; \mu, \phi) = \begin{cases} \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1} & \text{se } 0 < y < 1, \\ 0 & \text{caso contrário,} \end{cases} \quad (2.1.1.5)$$

onde $0 < \mu < 1$ e $\phi > 0$. O valor esperado e a variância de Y são expressos, na nova parametrização, como:

$$E(Y) = \mu,$$

$$\text{Var}(Y) = \frac{V(\mu)}{1 + \phi}, \quad (2.1.1.6)$$

onde $V(\mu) = \mu(1 - \mu)$ denota a função de variância, e onde ϕ pode ser interpretado como um parâmetro de precisão, no sentido de que, para μ fixo, se o valor de ϕ aumenta, a variância da variável resposta diminui. A Figura 2.1.1.2 apresenta diferentes densidades beta sob a nova parametrização, correspondentes a alguns valores dos parâmetros (μ, ϕ) . Observamos que quando $\mu = 0.5$ as densidades se apresentam de forma simétrica; quando $\mu \neq 0.5$ as formas são não simétricas.

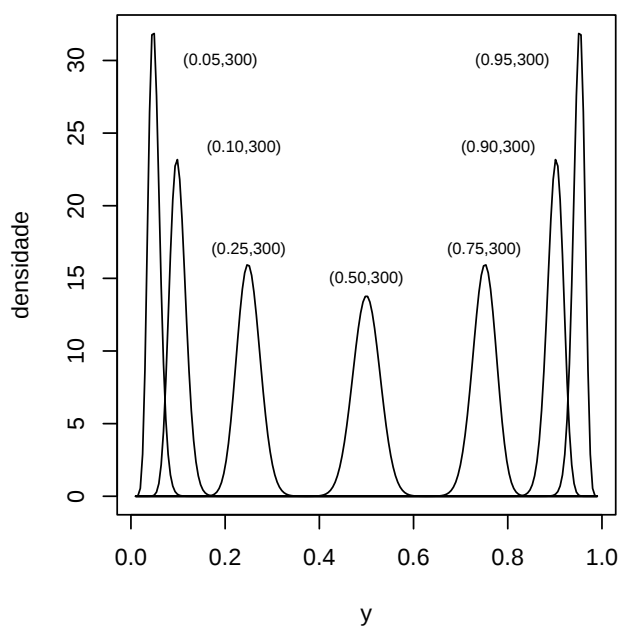
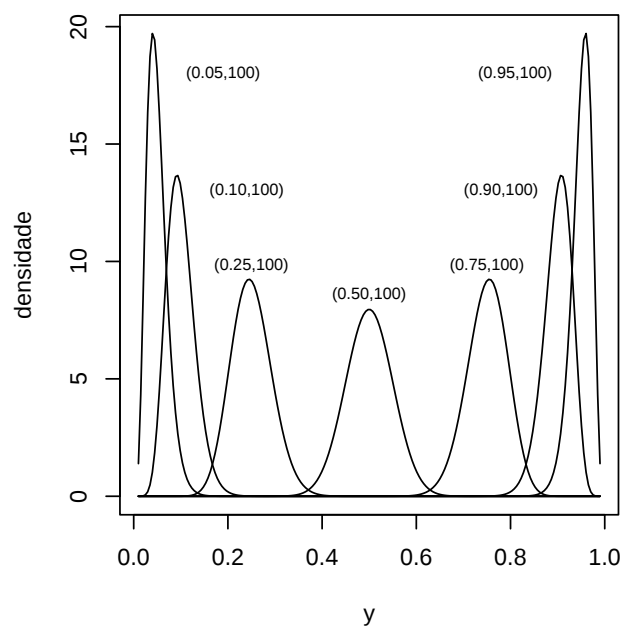
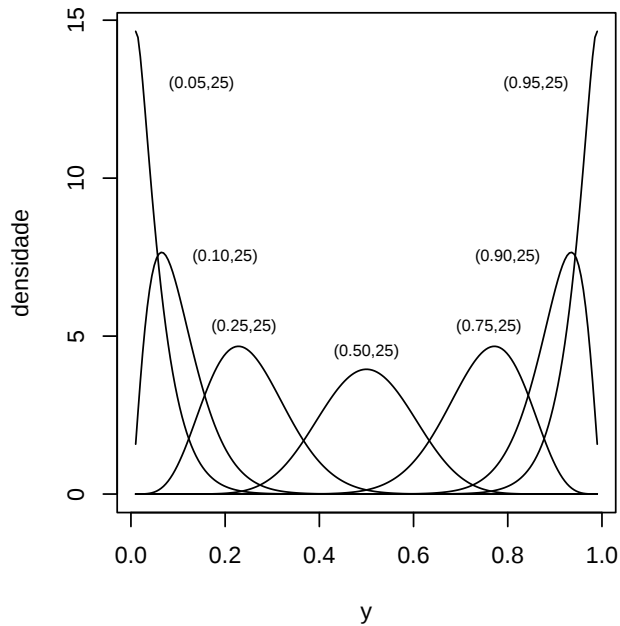
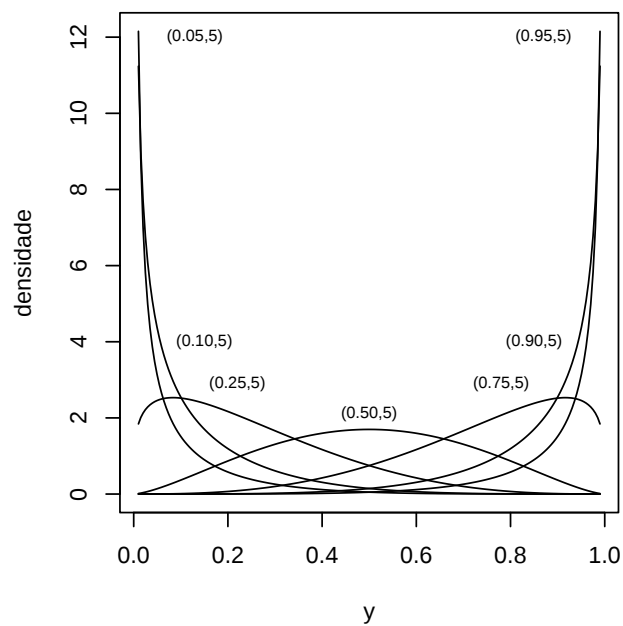
Quando a variável resposta assume valores no intervalo arbitrário (a, b) , com a e b conhecidos, pode-se modelar a variável $(Y - a)/(b - a)$, a qual é uma padronização que permite trabalhar com a função de densidade (2.1.1.5), no intervalo $(0, 1)$. O modelo proposto por Ferrari & Cribari-Neto (2004) é definido estabelecendo uma relação entre variáveis aleatórias com distribuição beta e algumas variáveis explicativas; nele uma relação é estabelecida entre μ e estas variáveis. Todo o desenvolvimento apresentado no restante desta seção está descrito em Ferrari & Cribari-Neto (2004).

Sejam $Y_1, \dots, Y_N$ variáveis aleatórias independentes, onde cada Y_i , $i = 1, \dots, N$, tem distribuição beta (2.1.1.5), isto é, $Y_i \sim \mathcal{B}(\mu_i, \phi)$. O modelo é obtido supondo que a média de Y_i pode ser escrita através de

$$h(\mu_i) = \beta_1 x_{i1} + \dots + \beta_k x_{ik} = \eta_i,$$

onde $h : (0, 1) \rightarrow \mathbb{R}$ é uma função de ligação estritamente monótona e duplamente diferenciável. Os parâmetros da regressão $\beta_1, \dots, \beta_k$ são desconhecidos e $x_{i1}, \dots, x_{ik}$ são observações em k covariáveis ($k < N$), que assumem valores fixos e conhecidos. Existem diversas relações funcionais para $h(\cdot)$, mas uma função bastante útil é a função logit definida por $h(\mu_i) = \log[\mu_i/(1 - \mu_i)]$. Com esta função, pode-se reescrever μ_i como

Figura 2.1.1.2: Gráficos das densidades beta para diferentes valores de (μ, ϕ) .



$$\mu_i = \frac{e^{-x_i^\top \beta}}{1 + e^{-x_i^\top \beta}}, \quad (2.1.1.7)$$

onde $x_i^\top = (x_{i1}, \dots, x_{ik})$ e $\beta = (\beta_1, \dots, \beta_k)$, $i = 1, \dots, N$.

Considerando a função de densidade dada em (2.1.1.5), a função de log-verossimilhança baseada em uma amostra de N observações para o modelo proposto é dada por

$$\ell(\beta, \phi) = \sum_{i=1}^N \ell_i(\mu_i, \phi), \quad (2.1.1.8)$$

onde

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

μ_i , definida através de (2.1.1.7) como $\mu_i = h^{-1}(\eta_i)$, é uma função de β , e ϕ é o parâmetro de precisão.

A função de escore é obtida pela diferenciação da função de log-verossimilhança (2.1.1.8) com relação aos parâmetros desconhecidos. Segue que para $u = 1, \dots, k$,

$$\frac{\partial \ell(\beta, \phi)}{\partial \beta_u} = \sum_{i=1}^N \frac{\partial \ell_i(\mu_i, \phi)}{\partial \mu_i} \frac{d\mu_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta_u} \quad (2.1.1.9),$$

mas

$$\begin{aligned} \frac{d\mu_i}{d\eta_i} &= \frac{1}{h'(\mu_i)}, \\ \frac{\partial \eta_i}{\partial \beta_u} &= x_{iu} \end{aligned} \quad (2.1.1.10)$$

e

$$\begin{aligned} \frac{\partial \ell_i(\mu_i, \phi)}{\partial \mu_i} &= \phi \left[\log \frac{y_i}{1 - y_i} - \{\psi(\mu_i \phi) - \psi((1 - \mu_i)\phi)\} \right] \\ &= \phi [y_i^* - \mu_i^*], \end{aligned} \quad (2.1.1.11)$$

onde $\psi(\cdot)$ é a função digamma, ou seja, $\psi(z) = d \log \Gamma(z) / dz$ para $z > 0$, $y_i^* = \log\{y_i/(1-y_i)\}$ e $\mu_i^* = \{\psi(\mu_i \phi) - \psi((1 - \mu_i)\phi)\}$. Substituindo (2.1.1.10) e (2.1.1.11) em (2.1.1.9) temos

$$\frac{\partial \ell(\beta, \phi)}{\partial \beta_u} = U_\beta(\beta, \phi) = \phi \sum_{i=1}^N (y_i^* - \mu_i^*) \frac{1}{h'(\mu_i)} x_{iu}. \quad (2.1.1.12)$$

Temos ainda que

$$\frac{\partial \ell(\beta, \phi)}{\partial \phi} = U_\phi(\beta, \phi) = \sum_{i=1}^N \{ \mu_i [y_i^* - \mu_i^*] + \log(1 - y_i) + \psi(\phi) - \psi((1 - \mu_i)\phi) \}. \quad (2.1.1.13)$$

Pode-se reescrever as equações (2.1.1.12) e (2.1.1.13) na forma matricial como

$$U_\beta(\beta, \phi) = \phi X^\top T(y^* - \mu^*) \quad \text{e} \quad U_\phi(\beta, \phi) = (\{(y^* - \mu^*)^\top \text{diag}(\mu_i)\} - a^\top) \mathbf{1},$$

onde X é a matriz formada pelas covariáveis do modelo, com dimensão $N \times k$ e i -ésimo vetor linha $x_i^\top$, $y^* = (y_1^*, \dots, y_N^*)$, $\mu^* = (\mu_1^*, \dots, \mu_N^*)$, $T = \text{diag}\left(\frac{d\mu_i}{d\eta_i}\right)$, e de forma análoga, $a^\top = (\log(1-y_1) + \psi(\phi) - \psi((1-\mu_1)\phi), \dots, \log(1-y_N) + \psi(\phi) - \psi((1-\mu_N)\phi))$ e $\mathbf{1} = (1, \dots, 1)^\top$ é o vetor de uns de dimensão $N \times 1$.

A matriz de informação de Fisher é dada por

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

onde $K_{\beta\beta} = \phi X^\top W X$, $K_{\beta\phi} = K_{\phi\beta}^\top = X^\top T c$, $K_{\phi\phi} = \text{tr}(D)$, sendo W , c e D descritos a seguir. $W = \text{diag}(w_1, \dots, w_N)$, com

$$w_i = \phi \psi'(\mu_i \phi) + \psi'((1 - \mu_i)\phi) \frac{1}{[h'(\mu_i)]^2},$$

$c = (c_1, \dots, c_N)$, com $c_i = \phi[\psi'(\mu_i \phi)\mu_i - \psi'((1 - \mu_i)\phi)(1 - \mu_i)]$, onde $\psi'(\cdot)$ é a função trigamma e $D = \text{diag}(d_1, \dots, d_N)$, $d_i = \psi'(\mu_i \phi)\mu_i^2 - \psi'((1 - \mu_i)\phi)(1 - \mu_i)^2 - \psi'(\phi)$.

Diante das condições de regularidade para estimação por máxima verossimilhança discutidas em Cordeiro (1999), se o tamanho da amostra é grande, tem-se

$$\begin{pmatrix} \hat{\beta} \\ \hat{\phi} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \beta \\ \phi \end{pmatrix}, K^{-1} \right),$$

aproximadamente, onde $\hat{\beta}$ e $\hat{\phi}$ são os estimadores de máxima verossimilhança de β e ϕ , respectivamente. Estes estimadores são obtidos pela solução do sistema

$$\begin{aligned} U_\beta(\beta, \phi) &= 0, \\ U_\phi(\beta, \phi) &= 0, \end{aligned}$$

e não têm forma fechada. Conseqüentemente eles devem ser obtidos pela maximização da função de log-verossimilhança, usando um algoritmo de otimização não-linear, como o algoritmo de Newton (Newton-Raphson, Escore de Fisher, etc.) ou algum algoritmo quasi-Newton (BFGS, entre outros). O algoritmo de otimização requer um valor inicial para ser usado no esquema iterativo. Para isso, Ferrari & Cribari-Neto (2004) sugerem como valor inicial para β a estimativa dos mínimos quadrados ordinários, $(X^\top X)^{-1} X^\top z$ com $z = (h(y_1), \dots, h(y_N))$, onde $h(y) = \log[y/(1 - y)]$. Sugerem também o valor inicial para ϕ dado pela expressão

$$\frac{1}{N} \sum_{i=1}^N \frac{\tilde{\mu}_i(1 - \tilde{\mu}_i)}{\tilde{\sigma}_i^2} - 1,$$

onde $\tilde{\mu}_i$ é obtido aplicando $h^{-1}(\cdot)$ ao i -ésimo valor ajustado da regressão linear de z em X , ou seja, $\tilde{\mu}_i = h^{-1}(x_i^\top (X^\top X)^{-1} X^\top z)$, e $\tilde{\sigma}_i^2 = \tilde{e}^\top \tilde{e} / [(N - k) \{g'(\tilde{\mu}_i)\}^2]$, sendo $\tilde{e} = z - (X^\top X)^{-1} X^\top z$ o vetor de resíduos de mínimos quadrados da regressão linear na variável resposta transformada.

2.2. Modelo Dirichlet

Agora consideraremos uma extensão para o modelo de regressão descrito anteriormente. Esta é baseada em uma extensão multivariada da distribuição beta também conhecida como distribuição Dirichlet.

A distribuição beta pode ser empregada no estudo de variáveis que possam ser colocadas no intervalo $(0, 1)$. A distribuição Dirichlet, por sua vez, pode ser usada para modelar a distribuição conjunta de variáveis que estão compreendidas no intervalo $(0, 1)$ e cuja soma é igual a um. Esta distribuição é muito útil em estudos de proporções, divisão aleatória de um intervalo (Mauldon, 1951), além de ser usada como distribuição a priori na análise bayesiana (Good, 1965). A seguir, consideramos algumas propriedades da distribuição Dirichlet, como sua função de densidade e seus primeiros momentos.

2.2.1. Definição e estimação

Consideremos um modelo de regressão com observações multivariadas, em que, para todo $i = 1, \dots, N$, a observação $(Y_{i1}, \dots, Y_{ip})$, com $Y_{i1} + \dots + Y_{ip} = 1$, tem distribuição Dirichlet, com parâmetros $\alpha_{i1} > 0, \alpha_{i2} > 0, \dots, \alpha_{ip} > 0$. O vetor aleatório $(Y_1, \dots, Y_p)$, com $Y_1 + \dots + Y_p = 1$, tem distribuição Dirichlet com parâmetros $\alpha_1 > 0, \alpha_2 > 0, \dots, \alpha_p > 0$, quando $(Y_1, \dots, Y_{p-1})$ admite a função de densidade

$$f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) = \begin{cases} \frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} \prod_{j=1}^{p-1} y_j^{\alpha_j-1} (1 - \sum_{j=1}^{p-1} y_j)^{\alpha_p-1} & y \in R, \\ 0 & \text{c.c.} \end{cases} \quad (2.2.1.1)$$

onde $\Gamma(\cdot)$ é a função gamma e a região R é dada por:

$$R = \left\{ (y_1, \dots, y_{p-1}) \in \mathbb{R}^{p-1}; y_j > 0 \text{ e } \sum_{j=1}^{p-1} y_j < 1 \right\}.$$

Mostraremos a seguir que $f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p)$ é efetivamente uma densidade. Inicialmente notamos que $f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) \geq 0$, já que $\Gamma(\cdot) > 0$, $(1 - \sum_{j=1}^{p-1} y_j) > 0$, pois $\sum_{j=1}^{p-1} y_j < 1$, e, como $y_j > 0$, temos então que $\prod_{j=1}^{p-1} y_j^{\alpha_j-1} > 0$, $j = 1, \dots, p-1$. A seguir mostraremos que a densidade (2.2.1.1) integra um. Por exemplo, se $p = 3$, temos que

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2; \alpha_1, \alpha_2, \alpha_3) dy_1 dy_2 &= \frac{\Gamma(\sum_{j=1}^3 \alpha_j)}{\prod_{j=1}^3 \Gamma(\alpha_j)} \int_0^1 \int_0^{1-y_2} \prod_{j=1}^2 y_j^{\alpha_j-1} \\ &\quad (1 - \sum_{j=1}^2 y_j)^{\alpha_3-1} dy_1 dy_2 \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \int_0^1 \int_0^{1-y_2} y_1^{\alpha_1-1} y_2^{\alpha_2-1} (1 - y_1 - y_2)^{\alpha_3-1} dy_1 dy_2 \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \int_0^1 \left[\int_0^{1-y_2} y_1^{\alpha_1-1} (1 - y_1 - y_2)^{\alpha_3-1} dy_1 \right] y_2^{\alpha_2-1} dy_2. \end{aligned} \quad (2.2.1.2)$$

A integral entre colchetes que aparece em (2.2.1.2) pode ser obtida usando (2.1.1.2). Se fizermos $a = 1 - y_2$, $s = \alpha_1$ e $r = \alpha_3$, então, teremos

$$\int_0^{1-y_2} y_1^{\alpha_1-1} (1 - y_1 - y_2)^{\alpha_3-1} dy_1 = B(\alpha_1, \alpha_3) (1 - y_2)^{(\alpha_1+\alpha_3-1)},$$

logo (2.2.1.2) pode ser reescrita como

$$\frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} B(\alpha_1, \alpha_3) \int_0^1 y_2^{\alpha_2-1} (1 - y_2)^{(\alpha_1+\alpha_3-1)} dy_2.$$

Usando novamente (2.1.1.2) fazendo $a = 1$, $s = \alpha_2$ e $r = \alpha_1 + \alpha_3$ temos

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2; \alpha_1, \alpha_2, \alpha_3) dy_1 dy_2 = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} B(\alpha_1, \alpha_3) B(\alpha_2, \alpha_1 + \alpha_3);$$

por (2.1.1.3) temos

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y_1, y_2; \alpha_1, \alpha_2, \alpha_3) dy_1 dy_2 &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \frac{\Gamma(\alpha_1)\Gamma(\alpha_3)}{\Gamma(\alpha_1 + \alpha_3)} \frac{\Gamma(\alpha_2)\Gamma(\alpha_1 + \alpha_3)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)} \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)} = 1. \end{aligned}$$

Generalizando o desenvolvimento acima, temos que

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) dy_1 dy_2 \dots dy_{p-1}$$

será igual a

$$\frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} \left[\int_0^1 \int_0^{1-y_{p-1}} \cdots \int_0^{1-y_2-\dots-y_{p-1}} \prod_{j=1}^{p-1} y_j^{\alpha_j-1} (1 - \sum_{j=1}^{p-1} y_j)^{\alpha_p-1} dy_1 \dots dy_{p-1} \right].$$

Logo, usando (2.1.1.2) $p-1$ vezes temos que a integral anterior entre os colchetes é igual a

$$B(\alpha_1, \alpha_p) B(\alpha_2, \alpha_1 + \alpha_p) \dots B(\alpha_{p-1}, \alpha_1 + \alpha_2 + \dots + \alpha_{p-2} + \alpha_p),$$

e, por (2.1.1.3), é equivalente a

$$\frac{\Gamma(\alpha_1)\Gamma(\alpha_p)}{\Gamma(\alpha_1 + \alpha_p)} \frac{\Gamma(\alpha_2)\Gamma(\alpha_1 + \alpha_p)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_p)} \cdots \frac{\Gamma(\alpha_{p-1})\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_{p-2} + \alpha_p)}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_{p-2} + \alpha_{p-1} + \alpha_p)},$$

que depois de algumas simplificações se torna

$$\frac{\Gamma(\alpha_1)\Gamma(\alpha_2) \dots \Gamma(\alpha_p)}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p)}.$$

Logo

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) dy_1 dy_2 \dots dy_{p-1} = \frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} \frac{\prod_{j=1}^p \Gamma(\alpha_j)}{\Gamma(\sum_{j=1}^p \alpha_j)}.$$

Portanto a função de densidade da distribuição Dirichlet integra um, ou seja,

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) dy_1 dy_2 \dots dy_{p-1} = 1$$

Algumas propriedades da distribuição Dirichlet como, por exemplo, a média, a variância e a covariância serão apresentadas a seguir. Inicialmente encontramos o valor esperado de $Y_1^{q_1} Y_2^{q_2} \dots Y_{p-1}^{q_{p-1}}$, que é dado por

$$\begin{aligned} E(Y_1^{q_1} Y_2^{q_2} \dots Y_{p-1}^{q_{p-1}}) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{p-1} y_j^{q_j} f(y_1, \dots, y_{p-1}; \alpha_1, \dots, \alpha_p) dy_1 \dots dy_{p-1} \\ &= \int_0^1 \int_0^{1-y_{p-1}} \cdots \int_0^{1-y_2-\dots-y_{p-1}} \prod_{j=1}^{p-1} y_j^{q_j} \frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} \prod_{j=1}^{p-1} y_j^{\alpha_j-1} (1 - \sum_{j=1}^{p-1} y_j)^{\alpha_p-1} dy_1 \dots dy_{p-1} \\ &= \frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} \int_0^1 \int_0^{1-y_{p-1}} \cdots \int_0^{1-y_2-\dots-y_{p-1}} \left\{ \prod_{j=1}^{p-1} y_j^{q_j+\alpha_j-1} (1 - \sum_{j=1}^{p-1} y_j)^{\alpha_p-1} dy_1 \dots dy_{p-1} \right\}. \end{aligned}$$

Usando a integral dada em (2.1.1.2) $p-1$ vezes temos

$$\begin{aligned} E(Y_1^{q_1} Y_2^{q_2} \dots Y_{p-1}^{q_{p-1}}) &= \frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} B(\alpha_1 + q_1, \alpha_p) B(\alpha_2 + q_2, \alpha_1 + \alpha_p + q_1) \dots \\ &\quad B(\alpha_{p-1} + q_{p-1}, \alpha_1 + \alpha_2 + \dots + \alpha_{p-2} + \alpha_p + q_1 + \dots + q_{p-2}) \end{aligned}$$

$$\begin{aligned}
&= \frac{\Gamma(\sum_{j=1}^p \alpha_j)}{\prod_{j=1}^p \Gamma(\alpha_j)} \frac{\Gamma(\alpha_1 + q_1)\Gamma(\alpha_p)}{\Gamma(\alpha_1 + \alpha_p + q_1)} \frac{\Gamma(\alpha_2 + q_2)\Gamma(\alpha_1 + \alpha_p + q_1)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_p + q_1 + q_2)} \dots \\
&\quad \frac{\Gamma(\alpha_{p-1} + q_{p-1})\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_{p-2} + \alpha_p + q_1 + \dots + q_{p-2})}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_{p-2} + \alpha_{p-1} + \alpha_p + q_1 + \dots + q_{p-2} + q_{p-1})} \\
&= \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p)\Gamma(\alpha_1 + q_1)\Gamma(\alpha_2 + q_2) \dots \Gamma(\alpha_{p-1} + q_{p-1})}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p + q_1 + \dots + q_{p-1})\Gamma(\alpha_1)\Gamma(\alpha_2) \dots \Gamma(\alpha_{p-1})}. \tag{2.2.1.4}
\end{aligned}$$

Se $q_1 = n$ e $q_2 = q_3 = \dots = q_{p-1} = 0$ temos

$$\begin{aligned}
E(Y_1^n) &= \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p)\Gamma(\alpha_1 + n)\Gamma(\alpha_2) \dots \Gamma(\alpha_{p-1})}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p + n)\Gamma(\alpha_1)\Gamma(\alpha_2) \dots \Gamma(\alpha_{p-1})} \\
&= \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p)\Gamma(\alpha_1 + n)}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p + n)\Gamma(\alpha_1)},
\end{aligned}$$

e, usando a relação (2.1.1.4), temos

$$\begin{aligned}
E(Y_1^n) &= \frac{\Gamma(\sum_{j=1}^p \alpha_j)(\alpha_1 + (n-1))(\alpha_1 + (n-2)) \dots (\alpha_1 + 1)\alpha_1 \Gamma(\alpha_1)(\Gamma(\alpha_1))^{-1}}{(\sum_{j=1}^p \alpha_j + (n-1))(\sum_{j=1}^p \alpha_j + (n-2)) \dots (\sum_{j=1}^p \alpha_j + 1)(\sum_{j=1}^p \alpha_j)\Gamma(\sum_{j=1}^p \alpha_j)} \\
&= \frac{(\alpha_1 + (n-1))(\alpha_1 + (n-2)) \dots (\alpha_1 + 1)\alpha_1}{(\sum_{j=1}^p \alpha_j + (n-1))(\sum_{j=1}^p \alpha_j + (n-2)) \dots (\sum_{j=1}^p \alpha_j + 1)(\sum_{j=1}^p \alpha_j)}.
\end{aligned}$$

Logo, o n -ésimo momento em torno de zero da distribuição Dirichlet, com $j = 1, 2, \dots, p-1$ será

$$E(Y_j^n) = \frac{(\alpha_j + (n-1))(\alpha_j + (n-2)) \dots (\alpha_j + 1)\alpha_j}{(\sum_{t=1}^p \alpha_t + (n-1))(\sum_{t=1}^p \alpha_t + (n-2)) \dots (\sum_{t=1}^p \alpha_t + 1)(\sum_{t=1}^p \alpha_t)}.$$

Então, o valor esperado de Y_j , com $j = 1, 2, \dots, p-1$, obtido quando $n = 1$ na expressão acima, é

$$E(Y_j) = \frac{\alpha_j}{\sum_{t=1}^p \alpha_t},$$

e, se $n = 2$,

$$E(Y_j^2) = \frac{\alpha_j(\alpha_j + 1)}{(\sum_{t=1}^p \alpha_t)(\sum_{t=1}^p \alpha_t + 1)},$$

logo, a variância de Y_j , com $j = 1, 2, \dots, p-1$ é

$$\begin{aligned}\text{Var}(Y_j) &= E(Y_j^2) - [E(Y_j)]^2 \\ &= \frac{\alpha_j(\alpha_j + 1)}{(\sum_{t=1}^p \alpha_t)(\sum_{t=1}^p \alpha_t + 1)} - \frac{\alpha_j^2}{(\sum_{t=1}^p \alpha_t)^2} \\ &= \frac{\alpha_j(\sum_{t \neq j}^p \alpha_t)}{(\sum_{t=1}^p \alpha_t)^2(\sum_{t=1}^p \alpha_t + 1)}.\end{aligned}$$

A Figura 2.2.1.2 apresenta diferentes densidades Dirichlet, quando $p = 3$, correspondentes a alguns valores dos parâmetros $(\alpha_1, \alpha_2, \alpha_3)$. Observamos que quando α_1 , α_2 e α_3 aumentam, a variabilidade da distribuição diminui, o que é confirmado pela expressão da variância.

Para o cálculo da covariância basta fazer $q_1 = q_2 = 1$ e $q_3 = \dots = q_{p-1} = 0$ em (2.2.1.4):

$$\begin{aligned}E(Y_1 Y_2) &= \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p) \Gamma(\alpha_1 + 1) \Gamma(\alpha_2 + 1) \Gamma(\alpha_3) \dots \Gamma(\alpha_{p-1})}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p + 1 + 1) \Gamma(\alpha_1) \Gamma(\alpha_2) \Gamma(\alpha_3) \dots \Gamma(\alpha_{p-1})} \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p) \Gamma(\alpha_1 + 1) \Gamma(\alpha_2 + 1)}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p + 1 + 1) \Gamma(\alpha_1) \Gamma(\alpha_2)} \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p) \alpha_1 \Gamma(\alpha_1) \alpha_2 \Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p + 1 + 1) \Gamma(\alpha_1) \Gamma(\alpha_2)} \\ &= \frac{\Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_p) \alpha_1 \alpha_2}{\Gamma[(\alpha_1 + \alpha_2 + \dots + \alpha_p + 1) + 1]} \\ &= \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2 + \dots + \alpha_p + 1)(\alpha_1 + \alpha_2 + \dots + \alpha_p)},\end{aligned}$$

então, para $j \neq j^*$, com $j, j^* = 1, \dots, p-1$

$$E(Y_j Y_{j^*}) = \frac{\alpha_j \alpha_{j^*}}{(\sum_{t=1}^p \alpha_t)(\sum_{t=1}^p \alpha_t + 1)},$$

portanto,

$$\begin{aligned}
\text{Cov}(Y_j, Y_{j*}) &= E(Y_j Y_{j*}) - E(Y_j)E(Y_{j*}) \\
&= \frac{\alpha_j \alpha_{j*}}{(\sum_{t=1}^p \alpha_t)(\sum_{t=1}^p \alpha_t + 1)} - \frac{\alpha_j \alpha_{j*}}{(\sum_{t=1}^p \alpha_t)^2} \\
&= -\frac{\alpha_j \alpha_{j*}}{(\sum_{t=1}^p \alpha_t)^2 (\sum_{t=1}^p \alpha_t + 1)}.
\end{aligned}$$

A log-densidade de (2.2.1.1) é da forma

$$\log \Gamma\left(\sum_{j=1}^p \alpha_j\right) - \sum_{j=1}^p \log \Gamma(\alpha_j) + \sum_{j=1}^p (\alpha_j - 1) \log y_j,$$

assim, as derivadas de primeira ordem da log-densidade da distribuição Dirichlet são

$$\frac{\partial \log f(y_1, \dots, y_{p-1}; \alpha_j)}{\partial \alpha_j} = \psi\left(\sum_{t=1}^p \alpha_t\right) - \psi(\alpha_j) + \log y_j,$$

com $j = 1, \dots, p$ e $\psi(\cdot)$ denotando a função digamma. Como o valor esperado da função escore é nulo, ou seja,

$$E\left\{\frac{\partial \log f}{\partial \alpha_j}\right\} = 0,$$

temos que

$$\begin{aligned}
E\left\{\frac{\partial \log f}{\partial \alpha_j}\right\} &= E\left\{\psi\left(\sum_{t=1}^p \alpha_t\right) - \psi(\alpha_j) + \log y_j\right\} \\
&= \psi\left(\sum_{t=1}^p \alpha_t\right) - \psi(\alpha_j) + E[\log y_j] = 0,
\end{aligned}$$

logo

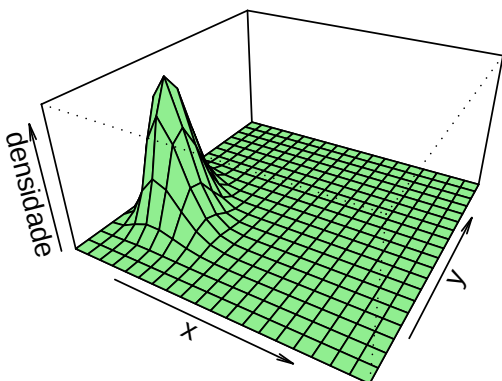
$$E[\log y_j] = \psi(\alpha_j) - \psi\left(\sum_{t=1}^p \alpha_t\right). \quad (2.2.1.5)$$

Considerando a função de densidade dada em (2.2.1.1), a função de log-verossimilhança baseada em uma amostra de N observações é dada por

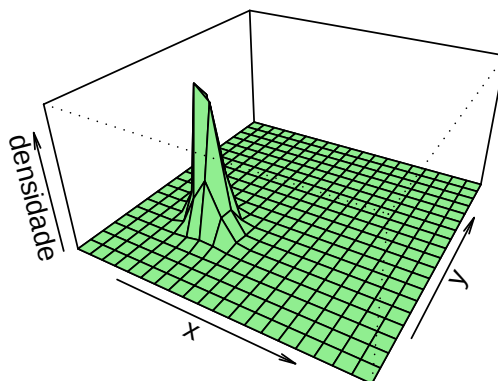
$$\ell(\alpha_1, \dots, \alpha_p) = \sum_{i=1}^N \ell_i(\alpha_{i1}, \dots, \alpha_{ip}), \quad (2.2.1.6)$$

Figura 2.2.1.2: Gráficos das densidades Dirichlet, quando $p = 3$, para diferentes valores de $(\alpha_1, \alpha_2, \alpha_3)$.

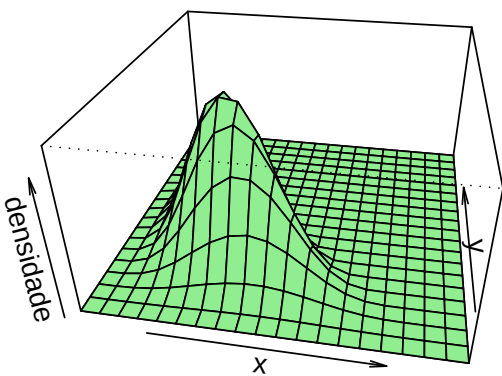
(5,10,15)



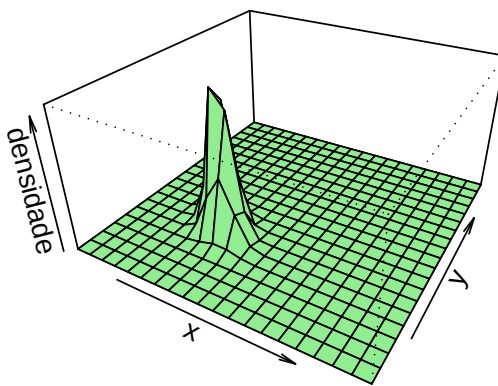
(30,35,40)



(5,5,5)



(30,30,30)



onde

$$\begin{aligned}
\ell_i(\alpha_{i1}, \dots, \alpha_{ip}) &= \log \left\{ \frac{\Gamma(\sum_{j=1}^p \alpha_{ij})}{\prod_{j=1}^p \Gamma(\alpha_{ij})} \prod_{j=1}^{p-1} y_{ij}^{\alpha_{ij}-1} \left[1 - \sum_{j=1}^{p-1} y_{ij} \right]^{\alpha_{ip}-1} \right\} \\
&= \log \left\{ \frac{\Gamma(\sum_{j=1}^p \alpha_{ij})}{\prod_{j=1}^p \Gamma(\alpha_{ij})} \prod_{j=1}^p y_{ij}^{\alpha_{ij}-1} \right\} \\
&= \log \Gamma \left(\sum_{j=1}^p \alpha_{ij} \right) - \sum_{j=1}^p \log \Gamma(\alpha_{ij}) + \sum_{j=1}^p (\alpha_{ij} - 1) \log y_{ij}.
\end{aligned}$$

Portanto,

$$\ell(\alpha_1, \dots, \alpha_p) = \sum_{i=1}^N \left\{ \log \Gamma \left(\sum_{j=1}^p \alpha_{ij} \right) - \sum_{j=1}^p \log \Gamma(\alpha_{ij}) + \sum_{j=1}^p (\alpha_{ij} - 1) \log y_{ij} \right\}. \quad (2.2.1.7)$$

Nas duas seções seguintes são considerados dois modelos de regressão para modelar variáveis aleatórias que seguem uma distribuição Dirichlet, a qual possui densidade dada por (2.2.1.1), indexada pelos parâmetros $\alpha_1, \dots, \alpha_p$.

2.2.2. Modelo 1

Para obter uma estrutura de regressão que permita modelar a média de uma variável resposta juntamente com um parâmetro de precisão, usamos uma parametrização diferente daquela mostrada na densidade (2.2.1.1). Seja $\phi = \alpha_1 + \alpha_2 + \dots + \alpha_p$ e $\mu_j = \alpha_j / \phi$, com $j = 1, 2, \dots, p-1$ e $\mu_p = 1 - \sum_{j=1}^{p-1} \mu_j$. A função de densidade de $(Y_1, \dots, Y_{p-1})$, dada em (2.2.1.1), pode ser reescrita, na nova parametrização, como

$$f(y_1, \dots, y_{p-1}; \mu_1, \dots, \mu_{p-1}; \phi) = \begin{cases} \frac{\Gamma(\phi)}{\prod_{j=1}^p \Gamma(\phi \mu_j)} \prod_{j=1}^{p-1} y_j^{\phi \mu_j - 1} (1 - \sum_{j=1}^{p-1} y_j)^{\phi \mu_p - 1} & y \in R, \\ 0 & \text{c.c.} \end{cases} \quad (2.2.2.1)$$

com $j = 1, 2, \dots, p-1$. O valor esperado e a variância de Y_j são dados por

$$E(Y_j) = \frac{\alpha_j}{\phi} = \frac{\phi \mu_j}{\phi} = \mu_j,$$

$$\text{Var}(Y_j) = \frac{\alpha_j(\phi - \alpha_j)}{\phi^2(\phi + 1)} = \frac{\phi \mu_j(\phi - \phi \mu_j)}{\phi^2(\phi + 1)} = \frac{\mu_j(1 - \mu_j)}{\phi + 1} = \frac{V(\mu_j)}{\phi + 1}, \quad (2.2.2.2)$$

a covariância entre Y_j e Y_{j*} para $j \neq j^*$ é dada por

$$\begin{aligned}\text{Cov}(Y_j, Y_{j*}) &= -\frac{\alpha_j \alpha_{j*}}{(\sum_{t=1}^p \alpha_t)^2 (\sum_{t=1}^p \alpha_t + 1)} \\ &= \frac{\phi \mu_j \phi \mu_{j*}}{\phi^2 (\phi + 1)} \\ &= \frac{\mu_j \mu_{j*}}{\phi + 1},\end{aligned}$$

onde μ_j é a média da variável resposta, $V(\mu_j) = \mu_j(1 - \mu_j)$ denota a função variância, com j e $j^* \in \{1, 2, \dots, p-1\}$, e ϕ pode ser interpretado como um parâmetro de precisão, no sentido de que, para μ_j fixo, se o valor de ϕ aumenta, a variância da variável resposta diminui. O modelo considerado aqui é obtido supondo que a média de cada Y_{ij} , com $i = 1, \dots, N$, pode ser escrita como

$$\begin{aligned}\mu_{ij} &= \begin{cases} h^{-1}(\eta_{ij}) & \text{se } j = 1 \\ h^{-1}(\eta_{ij})[1 - \sum_{t=1}^{j-1} \mu_{it}] & \text{se } j = 2, 3, \dots, p-1 \end{cases} \\ &= \begin{cases} g(\eta_{ij}) & \text{se } j = 1 \\ g(\eta_{ij})[1 - \sum_{t=1}^{j-1} \mu_{it}] & \text{se } j = 2, 3, \dots, p-1 \end{cases}\end{aligned}\tag{2.2.2.3}$$

onde $h^{-1}(\cdot) = g(\cdot)$ é uma função de ligação que assume valores reais em $(0, 1)$, estritamente monótona e duas vezes diferenciável e

$$\begin{aligned}\eta_{ij} &= x_{i1}\beta_{1j} + x_{i2}\beta_{2j} + x_{i3}\beta_{3j} + \dots + x_{ik}\beta_{kj} \\ &= x_i^\top \underline{\beta}_j,\end{aligned}$$

sendo $x_i^\top = (x_{i1}, \dots, x_{ik})$, com $i = 1, \dots, N$, e $\underline{\beta}_j = (\beta_{1j}, \beta_{2j}, \dots, \beta_{kj})^\top$. Os parâmetros da regressão, $\underline{\beta}_j$, com $j = 1, \dots, p-1$, são desconhecidos e $x_i^\top$, com $i = 1, \dots, N$, são observações em k covariáveis ($k < N$), que assumem valores fixos e conhecidos. A vantagem deste modelo é que a média da variável resposta é modelada diretamente, como se usa na prática.

A relação funcional usada para $h(\cdot)$ pode ser, por exemplo, a função logit, ou seja,

$$h^{-1}(\eta_{ij}) = g(\eta_{ij}) = \frac{e^{\eta_{ij}}}{1 + e^{\eta_{ij}}}, \quad i \in \{1, \dots, N\}, \quad j \in \{1, \dots, p-1\}.$$

Considerando a função de densidade dada em (2.2.2.1) e $B = [\underline{\beta}_1 \quad \underline{\beta}_2 \quad \dots \quad \underline{\beta}_{p-1}]$ a matriz dos parâmetros β 's, a função de log-verossimilhança baseada em uma amostra de N

observações é dada por

$$\ell(B, \phi) = \sum_{i=1}^N \ell_i(B, \phi)$$

onde

$$\begin{aligned} \ell_i(B, \phi) &= \log \left\{ \frac{\Gamma(\phi)}{\prod_{j=1}^p \Gamma(\phi \mu_{ij})} \prod_{j=1}^{p-1} y_{ij}^{\phi \mu_{ij} - 1} \left[1 - \sum_{j=1}^{p-1} y_{ij} \right]^{\phi \mu_{ip} - 1} \right\} \\ &= \log \Gamma(\phi) - \sum_{j=1}^p \log \Gamma(\phi \mu_{ij}) + \sum_{j=1}^p (\phi \mu_{ij} - 1) \log y_{ij}, \end{aligned}$$

sendo μ_{ij} dado em (2.2.2.3) e $y_{ip} = 1 - \sum_{j=1}^{p-1} y_{ij}$.

Portanto,

$$\begin{aligned} \ell(B, \phi) &= \sum_{i=1}^N \left\{ \log \Gamma(\phi) - \sum_{j=1}^{p-1} \log \Gamma(\phi \mu_{ij}) - \log \Gamma \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \right. \\ &\quad \left. + \sum_{j=1}^{p-1} (\phi \mu_{ij} - 1) \log y_{ij} + \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) - 1 \right] \left(1 - \sum_{j=1}^{p-1} \log y_{ij} \right) \right\}. \end{aligned} \quad (2.2.2.4)$$

Para a obtenção da função escore precisamos da primeira derivada da função log-verossimilhança, (2.2.2.4) com relação aos parâmetros desconhecidos. Como é apresentado no Apêndice 1, para $v = 1, \dots, k$ e $u = 1, \dots, p - 1$

$$\begin{aligned} \frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} &= g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \log y_{iu} - \psi(\phi \mu_{iu}) - \left[\log y_{i(u+1)} - \psi(\phi \mu_{i(u+1)}) \right] g(\eta_{i(u+1)}) \right. \\ &\quad - \sum_{j=u+2}^{p-1} \left[\log y_{ij} - \psi(\phi \mu_{ij}) \right] g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \left(\log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \right. \\ &\quad \left. \left. - \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\} \end{aligned} \quad (2.2.2.5)$$

e

$$\frac{\partial \ell(B; \phi)}{\partial \phi} = \sum_{i=1}^N \left\{ \psi(\phi) + \sum_{j=1}^p [\psi(\phi \mu_{ij}) - \log y_{ij}] \mu_{ij} \right\} = \sum_{i=1}^N d_i, \quad (2.2.2.6)$$

onde

$$d_i = \psi(\phi) + \sum_{j=1}^p [\psi(\phi\mu_{ij}) - \log y_{ij}] \mu_{ij}. \quad (2.2.2.7)$$

As derivadas apresentadas em (2.2.2.5) e (2.2.2.6) podem ser expressas matricialmente. Seja X a matriz formada pelas covariáveis do modelo, com dimensão $N \times k$ e $x_v^\top = [x_{1v} \ x_{2v} \ \dots \ x_{Nv}]$ o v -ésimo vetor coluna de X . Defina a matriz C de dimensão $N \times (p-1)$ com $c_u = [c_{1u} \ c_{2u} \ \dots \ c_{Nu}]$, o u -ésimo vetor coluna, dado por

$$\begin{aligned} c_{iu} = & g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \log y_{iu} - \psi(\phi\mu_{iu}) - \left[\log y_{i(u+1)} - \psi(\phi\mu_{i(u+1)}) \right] g(\eta_{i(u+1)}) \right. \\ & - \sum_{j=u+2}^{p-1} \left[\log y_{ij} - \psi(\phi\mu_{ij}) \right] g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \left(\log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \right. \\ & \left. \left. - \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\}. \end{aligned}$$

Seja $D = (d_1 \ d_2 \ \dots \ d_N)$, com d_i definido em (2.2.2.7), e $\mathbf{1} = (1, \dots, 1)^\top$ o vetor de uns de dimensão $N \times 1$. Então a equação (2.2.2.5) pode ser reescrita como

$$\frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} = \phi x_v^\top c_u, \quad (2.2.2.8)$$

logo, de (2.2.2.8) e (2.2.2.6)

$$\frac{\partial \ell(B, \phi)}{\partial B} = U_B(B, \phi) = \phi X^\top C$$

e

$$\frac{\partial \ell(B, \phi)}{\partial \phi} = U_\phi(B, \phi) = D \mathbf{1}.$$

Os estimadores de máxima verossimilhança de B e ϕ são obtidos pela solução do sistema

$$\begin{aligned} U_B(B, \phi) &= 0 \\ U_\phi(B, \phi) &= 0 \end{aligned}$$

e tais estimadores não possuem forma fechada. Assim, eles devem ser obtidos numericamente usando um algoritmo de otimização não-linear, como o algoritmo de Newton ou algum algoritmo quasi-Newton, como por exemplo, BFGS. Usando a idéia de Ferrari & Cribari-Neto (2004), no esquema iterativo usamos como valor inicial para B a estimativa dos mínimos quadrados ordinários, $(X^\top X)^{-1} X^\top Z$ com $Z = [z_1 \ z_2 \ \dots \ z_N]^\top$ onde

$$z_i = \left[h(y_{i1}) \quad h\left(\frac{y_{i2}}{1 - y_{i1}}\right) \quad h\left(\frac{y_{i3}}{1 - y_{i1} - y_{i2}}\right) \quad \dots \quad h\left(\frac{y_{i(p-1)}}{1 - y_{i1} - \dots - y_{i(p-2)}}\right) \right]$$

e $h(y) = \log[y/(1 - y)]$. Usamos também o valor inicial para ϕ dado pela expressão

$$\frac{1}{N(p-1)} \sum_{i=1}^N \sum_{j=1}^{p-1} \frac{\tilde{\mu}_{ij}(1 - \tilde{\mu}_{ij})}{\tilde{\sigma}_{ij}^2} - 1,$$

onde $\tilde{\mu}_{ij}$ é obtido aplicando $h^{-1}(\cdot)$ ao i -ésimo valor ajustado da regressão linear de Z em X , ou seja, $\tilde{\mu}_{ij} = h_j^{-1}(x_i^\top (X^\top X)^{-1} X^\top z_j)$, e $\tilde{\sigma}_{ij}^2 = \tilde{e}_j^\top \tilde{e}_j / [(N - k)\{h'(\tilde{\mu}_{ij})\}^2]$, sendo $\tilde{e}_j = z_j - (X^\top X)^{-1} X^\top z_j$ o vetor de resíduos de mínimos quadrados da regressão linear sob a variável resposta transformada.

A matriz de informação de Fisher é dada por

$$I(B; \phi) = \begin{pmatrix} I_{\beta\beta} & I_{\beta\phi} \\ I_{\phi\beta} & I_{\phi\phi} \end{pmatrix}, \quad (2.2.2.9)$$

onde

$$\begin{aligned} I_{\beta\beta} &= \text{Var}\left(\frac{\partial \ell(B; \phi)}{\partial B}\right) \\ &= E\left[\frac{\partial \ell(B; \phi)}{\partial B} \frac{\partial \ell(B; \phi)}{\partial B^\top}\right] - E\left[\frac{\partial \ell(B; \phi)}{\partial B}\right] E\left[\frac{\partial \ell(B; \phi)}{\partial B^\top}\right] \\ &= -E\left[\frac{\partial^2 \ell(B; \phi)}{\partial B \partial B^\top}\right] - E\left[\frac{\partial \ell(B; \phi)}{\partial B}\right] E\left[\frac{\partial \ell(B; \phi)}{\partial B^\top}\right] \end{aligned} \quad (2.2.2.10)$$

mas

$$E\left[\frac{\partial \ell(B; \phi)}{\partial B}\right] = 0$$

logo

$$I_{\beta\beta} = -E\left[\frac{\partial^2 \ell(B; \phi)}{\partial B \partial B^\top}\right].$$

Observamos, pelo Apêndice 1, que para $v, n \in \{1, 2, \dots, k\}$ e $u, m \in \{1, 2, \dots, p-1\}$, temos

$$-E\left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{mu} \partial \beta_{vu}}\right] = \sum_{i=1}^N X_{iv} X_{in} U_{mu,i}^{(1)},$$

onde $U_{mu,i}^{(1)}$ é definido de acordo com os casos abaixo.

Caso 1: Se $u = m$,

$$\begin{aligned}
U_{uu,i}^{(1)} = & \phi^2 g'(\eta_{iu})^2 \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]^2 \left\{ \psi'(\phi\mu_{iu}) + \psi'(\phi\mu_{i(u+1)}) g(\eta_{i(u+1)})^2 \right. \\
& + \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij}) g(\eta_{ij})^2 \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]^2 + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] \\
& \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right)^2 \right\}.
\end{aligned}$$

Caso 2: Se $u > m$,

$$\begin{aligned}
U_{mu,i}^{(1)} = & \phi^2 g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left[1 - \sum_{t=1}^{m-1} \mu_{it}\right] \left[-\psi'(\phi\mu_{iu}) g(\eta_{iu}) \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] \right. \\
& + \psi'(\phi\mu_{i(u+1)}) g(\eta_{i(u+1)})^2 \prod_{t=m+1}^u [1 - g(\eta_{it})] + \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij}) g(\eta_{ij})^2 \\
& \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] \left(1 - g(\eta_{i(m+1)}) \right. \\
& - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \left. \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right].
\end{aligned}$$

Caso 3: Se $u < m$,

$$\begin{aligned}
U_{mu,i}^{(1)} = & \phi^2 g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left[1 - \sum_{t=1}^{m-1} \mu_{it}\right] \left[-\psi'(\phi\mu_{im}) g(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \right. \\
& + \psi'(\phi\mu_{i(m+1)}) g(\eta_{i(m+1)})^2 \prod_{t=u+1}^m [1 - g(\eta_{it})] + \sum_{j=m+2}^{p-1} \psi'(\phi\mu_{ij}) g(\eta_{ij})^2 \\
& \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] \left(1 - g(\eta_{i(m+1)}) \right. \\
& - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \left. \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right].
\end{aligned}$$

Portanto,

$$I_{\beta\beta} = (I_{p-1} \otimes X)^\top U (I_{p-1} \otimes X),$$

onde I_{p-1} é uma matriz identidade de dimensão $(p-1) \times (p-1)$, o símbolo $\otimes$ representa o produto de Kronecker e U é uma matriz $N(p-1) \times N(p-1)$ particionada em $(p-1) \times (p-1)$ blocos diagonais de dimensão $N \times N$ da forma

$$U_{mu} = \text{diag} (U_{mu,1}^{(1)}, U_{mu,2}^{(1)}, \dots, U_{mu,N}^{(1)}),$$

com u e $m \in \{1, 2, \dots, p-1\}$. Note ainda que $(I_{p-1} \otimes X)$ tem dimensão $N(p-1) \times k(p-1)$, o que implica que a matriz $I_{\beta\beta}$ tem dimensão $k(p-1) \times k(p-1)$.

Observamos ainda, também pelo Apêndice 1, que para $v = 1, 2, \dots, k$ e $u = 1, 2, \dots, p-1$,

$$-E \left[\frac{\partial^2 \ell(B; \phi)}{\partial \phi \partial \beta_{vu}} \right] = \sum_{i=1}^N X_{iv} V_{iu}^{(1)},$$

onde

$$\begin{aligned} V_{iu}^{(1)} = & \phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[\psi'(\phi \mu_{iu}) \mu_{iu} - \psi'(\phi \mu_{i(u+1)}) \mu_{i(u+1)} g(\eta_{i(u+1)}) \right. \\ & - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \mu_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \\ & \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \end{aligned}$$

e

$$-E \left[\frac{\partial^2 \ell(B; \phi)}{\partial \phi^2} \right] = \sum_{i=1}^N \left\{ -\psi'(\phi) + \sum_{j=1}^p \psi'(\phi \mu_{ij}) \mu_{ij}^2 \right\}.$$

Logo

$$I_{\beta\phi} = I_{\phi\beta}^\top = \text{vec}(X^\top V^{(1)}),$$

onde $\text{vec}(\cdot)$ é o operador que transforma uma matriz em vetor coluna sobrepondo as colunas da matriz e $V^{(1)}$ é uma matriz de dimensão $N \times (p-1)$; então $I_{\beta\phi}$ tem dimensão $k \times (p-1)$, e $I_{\phi\phi}$ é um escalar, dado por

$$I_{\phi\phi} = \sum_{i=1}^N \left\{ -\psi'(\phi) + \sum_{j=1}^p \psi'(\phi \mu_{ij}) \mu_{ij}^2 \right\}.$$

Portanto a matriz de informação de Fisher $I(B; \phi)$ tem dimensão $(k(p-1)+1) \times (k(p-1)+1)$.

Usando a expressão padrão para a inversa de matrizes particionadas temos que a inversa da matriz de informação de Fisher (2.2.2.9) é

$$I(B; \phi)^{-1} = \begin{pmatrix} I_{\beta\beta} & I_{\beta\phi} \\ I_{\phi\beta} & I_{\phi\phi} \end{pmatrix}^{-1} = \begin{pmatrix} I^{\beta\beta} & I^{\beta\phi} \\ I^{\phi\beta} & I^{\phi\phi} \end{pmatrix},$$

onde, por Rao (1973),

$$I^{\beta\beta} = I_{\beta\beta}^{-1} + I_{\beta\beta}^{-1} I_{\beta\phi} \left[I_{\phi\phi} - I_{\phi\beta} I_{\beta\beta}^{-1} I_{\beta\phi} \right]^{-1} \begin{pmatrix} I_{\beta\beta}^{-1} & I_{\beta\phi} \end{pmatrix}^{\top},$$

$$I^{\beta\phi} = -I_{\beta\beta}^{-1} I_{\beta\phi} \left[I_{\phi\phi} - I_{\phi\beta} I_{\beta\beta}^{-1} I_{\beta\phi} \right]^{-1},$$

$$I^{\phi\beta} = - \left[I_{\phi\phi} - I_{\phi\beta} I_{\beta\beta}^{-1} I_{\beta\phi} \right]^{-1} \begin{pmatrix} I_{\beta\beta}^{-1} & I_{\beta\phi} \end{pmatrix}^{\top} \quad \text{e}$$

$$I^{\phi\phi} = \left[I_{\phi\phi} - I_{\phi\beta} I_{\beta\beta}^{-1} I_{\beta\phi} \right]^{-1}.$$

Diante das condições de regularidade para estimação de máxima verossimilhança discutidas em Cordeiro (1999), se o tamanho da amostra N é grande, tem-se

$$\begin{pmatrix} \text{vec}(\widehat{B}) \\ \widehat{\phi} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \text{vec}(B) \\ \phi \end{pmatrix}, I(B; \phi)^{-1} \right),$$

aproximadamente, onde $\widehat{B}$ e $\widehat{\phi}$ são os estimadores de máxima verossimilhança da matriz de parâmetros B e do parâmetro ϕ , respectivamente.

2.2.3. Modelo 2

Nesta seção definimos o segundo modelo de regressão Dirichlet considerado. Este modelo é obtido supondo que $\alpha_{ij} = g(\eta_{ij})$, ou seja, a média de cada Y_{ij} , com $i = 1, \dots, N$ e $j = 1, \dots, p$, pode ser escrita como

$$\mu_{ij} = \frac{g(\eta_{ij})}{g(\eta_{i1}) + g(\eta_{i2}) + \dots + g(\eta_{ip})},$$

onde η_{ij} é definido da mesma forma do primeiro modelo e $g(\cdot) = h^{-1}(\cdot)$ onde $h(\cdot)$ é uma função de ligação que assume valores reais em $(0, \infty)$, estritamente monotônica e duas vezes diferenciável; neste trabalho consideramos $g(\eta_{ij}) = e^{\eta_{ij}}$.

Sabemos que $\alpha_{ij} = g(\eta_{ij}) = e^{\eta_{ij}}$ onde $\eta_{ij} = x_i^\top \underline{\beta}_j$, com $\underline{\beta}_j = (\beta_{1j}, \beta_{2j}, \dots, \beta_{kj})^\top$ e $x_i = (x_{i1}, \dots, x_{ik})^\top$. Seja x_i^* tal que $x_{il}^* = x_{il} + 1$, onde $l = 1, 2, \dots, k$; temos ainda que $\alpha_{ij}^* = g(\eta_{ij}^*) = e^{\eta_{ij}^*}$, onde $\eta_{ij}^* = x_i^{*\top} \underline{\beta}_j$. Então

$$\frac{\alpha_{ij}^*}{\alpha_{ij}} = \frac{e^{\eta_{ij}^*}}{e^{\eta_{ij}}} = e^{x_i^{*\top} \underline{\beta}_j - x_i^\top \underline{\beta}_j} = e^{\beta_{lj}},$$

logo

$$\beta_{lj} = \log(\alpha_{ij}^*) - \log(\alpha_{ij}).$$

Então os parâmetros β_{lj} 's representam a variação no $\log \alpha_{ij}$ quando os x_{il} 's aumentam de uma unidade.

A vantagem deste modelo é que se permutarmos as componentes, Y_j , de Y as componentes estimadas, $\hat{\mu}_j$, também são permutadas; isso decorre da não dependência da componente $\hat{\mu}_j$ de componentes anteriores, diferentemente do que ocorre no modelo 1. E a desvantagem é que este modelo não modela diretamente a média, como se usa na prática.

Considerando a função log-verossimilhança dada em (2.2.2.4), temos que sua derivada com relação aos parâmetros desconhecidos é dada pela expressão abaixo; os detalhes da derivação podem ser visto no Apêndice 2.

$$\begin{aligned} \frac{\partial \ell(B)}{\partial \beta_{vu}} &= \sum_{i=1}^N X_{iv} \left\{ g'(\eta_{iu}) \left[\psi(\phi_i) + \log y_{iu} - \psi(\phi_i \mu_{iu}) \right] \right\} \\ &= \sum_{i=1}^N X_{iv} c_{iu}, \quad v = 1, \dots, k \quad \text{e} \quad u = 1, \dots, p \end{aligned} \quad (2.2.3.1)$$

onde

$$c_{iu} = g'(\eta_{iu}) \left[\psi(\phi_i) + \log y_{iu} - \psi(\phi_i \mu_{iu}) \right] \quad (2.2.3.2)$$

e

$$\phi_i = \sum_{j=1}^p g(\eta_{ij}).$$

A derivada apresentada em (2.2.3.1) pode ser expressa matricialmente, ou seja,

$$\frac{\partial \ell(B)}{\partial B} = U_B(B) = X^\top C,$$

onde X é a matriz formada pelas covariáveis do modelo, com dimensão $N \times k$ e a matriz C definida em (2.2.3.2) tem dimensão $N \times p$.

Os estimadores de máxima verossimilhança de B são obtidos pela solução do sistema

$$U_B(B) = 0$$

e, como já foi dito no estudo do primeiro modelo, estes estimadores não possuem forma fechada e para serem obtidos deve-se usar um algoritmo de otimização não-linear. Para o esquema iterativo, consideramos a matriz nula como valor inicial para B .

Neste presente modelo, a matriz de informação de Fisher de B é obtida fazendo cálculos análogos a (2.2.2.10). Pelo Apêndice 2, temos que para $v, n \in \{1, 2, \dots, k\}$ e $u, m \in \{1, 2, \dots, p\}$

$$-E \left[\frac{\partial^2 \ell(B)}{\partial \beta_{mu} \partial \beta_{vu}} \right] = \sum_{i=1}^N X_{iv} X_{in} U_{mu,i}^{(2)},$$

onde

$$U_{mu,i}^{(2)} = \begin{cases} -g'(\eta_{iu})^2 [\psi'(\phi_i) - \psi'(\phi_i \mu_{iu})] & \text{se } u = m \\ -g'(\eta_{iu}) g'(\eta_{im}) \psi'(\phi_i) & \text{se } u \neq m \end{cases}$$

Logo, a matriz de informação de Fisher de B é

$$I(B) = (I_p \otimes X)^\top U^* (I_p \otimes X),$$

onde I_p é uma matriz identidade de ordem p , o símbolo $\otimes$ representa o produto de Kronecker e U^* é uma matriz $Np \times Np$ particionada em $p \times p$ blocos diagonais de dimensão $N \times N$, U_{mu}^* , definidos por

$$U_{mu}^* = \text{diag} (U_{mu,1}^{(2)}, U_{mu,2}^{(2)}, \dots, U_{mu,N}^{(2)}),$$

com u e $m \in \{1, 2, \dots, p\}$. Ainda, $(I_p \otimes X)$ tem dimensão $Np \times kp$, o que implica que a matriz $I(B)$ tem dimensão $kp \times kp$.

Como já foi dito, se o tamanho da amostra N é grande, tem-se

$$\text{vec}(\hat{B}) \sim \mathcal{N}(\text{vec}(B), I(B)^{-1}),$$

aproximadamente, onde $\hat{B}$ é a matriz de estimadores de máxima verossimilhança da matriz de parâmetros B .

Capítulo 3

Estimador do Posto

3.1. Considerações Iniciais

Muitos modelos são especificados por meio do posto de uma certa matriz desconhecida para a qual uma estimativa consistente existe. Por exemplo, o problema clássico de identificação em modelos de equações simultâneas lineares envolve o posto de uma submatriz particular de forma reduzida. No que se refere a verossimilhança, o posto da matriz de informação relata identificabilidade de um vetor de parâmetros (Hsiao, 1986). Lewbel (1991) e Lewbel e Perraudin (1995) mostraram que vários resultados na teoria de consumo podem depender do verdadeiro posto de certas matrizes a serem estimadas.

Determinar o posto de uma matriz é uma tarefa ainda mais difícil se a matriz contém erros, que é sempre o caso em aplicações estatísticas baseadas em matrizes estimadas. O método comumente usado para avaliar o posto de uma matriz baseia-se na fatorização QR, ver Stewart (1984). Esta ferramenta, contudo, somente ajuda a tomar decisões e não constitui um teste. Gill e Lewbel (1992) introduzem um teste para o posto baseado na decomposição *Lower-Diagonal-Upper* (LDU). Mas a distribuição assintótica dada pela estatística de teste é incorreta, exceto para alguns casos especiais. Cragg e Donald (1996) mostram uma modificação apropriada. Seu teste tem a vantagem de possuir uma distribuição qui-quadrado. Anderson e Kunitomo (1994), Robin e Smith (2000) definem uma estatística de teste que é N vezes uma função do menor valor singular, onde esta estatística de teste é a da raiz característica (CRT). Esta classe de estatísticas (CRT) inclui muitas estatísticas de teste particulares como estatística da razão de verossimilhança, a estatística de score e estatística de Wald, ver Anderson e Kunitomo (1994). Segundo Robin e Smith (2000) a estatística CRT é distribuída assintoticamente como uma combinação linear de variáveis aleatórias independentes com distribuições qui-quadrado. A desvantagem desta estatística é que as matrizes pesos são desconhecidas e devem ser estimadas; na amostra, suas estimativas introduzem variabilidade, e, conseqüentemente, menos poder no procedimento do teste.

Sabemos que o posto de uma matriz é igual ao número de valores singulares diferentes de zero. Assim, um teste formal pode ser expresso como um teste do número de valores singulares nulos da matriz. Ratsimalahelo (2003) construiu um teste de posto baseado nos menores valores singulares de uma matriz estimada. Ele usou a teoria da perturbação de matrizes para estimar uma base satisfatória do núcleo da matriz e para determinar distribuições limite para os estimadores dos menores valores singulares da matriz. O teste permite determinar quantos valores singulares da estimativa de uma matriz são significativamente diferentes de zero. A vantagem deste teste sobre os testes padrões é a facilidade de seu cálculo.

Os coeficientes do modelo de regressão Dirichlet constituem uma matriz B . Se B não tem posto completo, ou seja, se algumas de suas colunas (ou linhas) podem ser escritas como combinações lineares de outras, então a quantidade de parâmetros do modelo a serem estimados é menor. Então nosso objetivo neste trabalho é estimar o posto b da matriz B

de parâmetros do modelo de regressão Dirichlet. Para isso vamos utilizar a estatística de teste $L(b)$, proposta por Ratsimalahelo (2003), associada à estratégia de teste seqüencial e aos critérios de informação BIC (*Bayesian Information Criteria*) e HQIC (*Hannan Quinn Information Criteria*).

3.1.1. Teste de Hipótese

Consideramos uma matriz não-observável B de dimensão $k \times q$, com posto verdadeiro desconhecido $b > 0$; sem perda de generalidade, assumimos que $k \geq q$.

Seja $\hat{B}$, estimador consistente de B tal que

$$\sqrt{N} \text{vec}(\hat{B} - B) \xrightarrow{d} \mathcal{N}(0, \Sigma), \quad (3.1.1.1)$$

onde a matriz de covariância Σ tem dimensão $kq \times kq$ e é não-nula.

Ratsimalahelo (2003) construiu um teste para o posto b de B , $r(B)$. Suponha que desejamos testar a hipótese nula

$$H_0 : r(B) = b$$

contra a hipótese alternativa

$$H_1 : r(B) > b.$$

O posto verdadeiro de B é desconhecido, mas $\hat{B}$ com probabilidade um tem posto completo. Assim, a hipótese nula baseada em um teste usando $\hat{B}$ é satisfeita somente assintoticamente em N . Então, o interesse se prende a testes estatísticos que nos habilitam a determinar o posto de B , dado o estimador $\hat{B}$.

Seja $\hat{B}$ a matriz perturbada

$$\hat{B} = B + \varepsilon C, \quad (3.1.1.2)$$

onde a matriz εC é a perturbação da matriz B . Assumimos que a matriz B , $k \times q$, tem posto b , mas que a matriz perturbação C , composta de valores pequenos, tem posto completo. Então, usando (3.1.1.1) e o teorema central do limite, os elementos da matriz perturbação têm média zero e variância de ordem N^{-1} . Isto implica que o termo dominante da matriz εC é de ordem $N^{-1/2}$, denotado por

$$\varepsilon C = \mathcal{O}_p(N^{-1/2}),$$

ou seja, εC é de ordem no máximo $N^{-1/2}$ em probabilidade, isto é, $\forall \delta > 0, \exists M_\delta$ tal que

$$P\left(\left|\frac{\varepsilon C}{N^{-1/2}}\right| \geq M_\delta\right) \leq \delta.$$

A análise de perturbação de matrizes estuda a sensibilidade dos autoelementos da matriz a uma perturbação em suas componentes. Existem bons desenvolvimentos matemáticos para perturbação de matrizes apresentados por Kato (1982), Golub e Van Loan (1996), Stewart e Sun (1990).

Uma boa técnica para determinar o posto de uma matriz é usar a decomposição da matriz em valores singulares e contar o número de valores singulares diferentes de zero. Esta técnica está descrita na próxima seção.

3.1.2. Decomposição em Valores Singulares e a Estatística de Teste

A decomposição em valores singulares de uma matriz foi proposta de forma independente por Beltrami em 1873 e Jordan em 1874. Sua generalização para espaços de dimensão infinita foi desenvolvida em um contexto de equações integrais por Smith (1907) e Weyl (1912).

Pelo teorema (Golub e Van Loan, 1996) de existência da decomposição em valores singulares, temos que se B é uma matriz real de dimensão $k \times q$, então, existem duas matrizes ortogonais

$$U = [u_1, u_2, \dots, u_k] \in \mathbb{R}^{k \times k} \quad e \quad V = [v_1, v_2, \dots, v_q] \in \mathbb{R}^{q \times q}$$

tais que

$$U^\top B V = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_\delta) = D,$$

onde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_\delta$ são os valores singulares de B e $\delta = \min(k, q)$.

A decomposição em valores singulares da matriz B de dimensão $k \times q$ é $B = U D V^\top$, com U e V ortogonais, onde U tem dimensão $k \times k$, D é uma matriz diagonal retangular $k \times q$ e V tem dimensão $q \times q$. Um algoritmo para decompor a matriz B em valores singulares se $k > q$ é

- (I) Encontre os autovalores da matriz $B^\top B$ e os organize em ordem decrescente;
- (II) Encontre o número de autovalores diferentes de zero da matriz $B^\top B$;
- (III) Encontre autovetores ortonormais v_i , com $i = 1, 2, \dots, q$, da matriz $B^\top B$, correspondentes aos seus respectivos autovalores, calculados em (II);
- (IV) Forme uma matriz diagonal $D \in \mathbb{R}^{k \times q}$ colocando na diagonal principal as raízes quadradas $\sqrt{\lambda_i^*}$ dos q autovalores da matriz $B^\top B$ em ordem decrescente;
- (V) Encontre os primeiros q vetores colunas de uma matriz $U \in \mathbb{R}^{k \times k}$ tal que

$$u_i = (\sqrt{\lambda_i^*})^{-1} B v_i, \quad \text{com } i = 1, 2, \dots, q;$$

- (VI) Complete a matriz U com $(k - q)$ vetores ortonormais usando o processo de ortonormalização de Gram-Schmidt.

Processo de ortonormalização de Gram-Schmidt: Se $(v_1, v_2, \dots, v_k)$ é uma k -upla de vetores linearmente independentes em um espaço vetorial euclidiano V , então, definindo $\tilde{v}_1 := v_1 / \|v_1\|$ e utilizando a fórmula de recorrência

$$\tilde{v}_{j+1} := \frac{v_{j+1} - \sum_{i=1}^j \langle v_{j+1}, \tilde{v}_i \rangle \tilde{v}_i}{\|v_{j+1} - \sum_{i=1}^j \langle v_{j+1}, \tilde{v}_i \rangle \tilde{v}_i\|},$$

para $j = 1, 2, \dots, k-1$, obtemos um sistema ortonormal $(\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_k)$. Os primeiros j vetores $\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_j$ geram o mesmo subespaço que o conjunto de vetores originais $v_1, v_2, \dots, v_j$. Se $(v_1, v_2, \dots, v_n)$ em particular era uma base de V , então $(\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_n)$ será uma base ortonormal de V . Os símbolos $\langle \cdot, \cdot \rangle$ e $\|\cdot\|$ são o produto interno entre dois vetores e a norma euclidiana de um vetor, respectivamente.

Podemos escrever ainda a decomposição em valores singulares de B na forma

$$B = UDV^\top = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} V_1^\top \\ V_2^\top \end{pmatrix}$$

onde U_1 , com dimensão $k \times b$, é uma base ortonormal do subespaço gerado pelas colunas de B , U_2 , com dimensão $k \times (k-b)$, é uma base ortonormal do complemento ortogonal do subespaço gerado pelas colunas de B , V_1 , com dimensão $q \times b$, é uma base ortonormal do subespaço gerado pelas linhas de B e V_2 , $q \times (q-b)$, é uma base ortonormal do espaço nulo de B . Por causa da ortogonalidade de U e V , ou seja, $U^\top U = I_{k \times k}$ e $V^\top V = I_{q \times q}$, a matriz D pode ser escrita como

$$D = \begin{pmatrix} U_1^\top \\ U_2^\top \end{pmatrix} B \begin{pmatrix} V_1 & V_2 \end{pmatrix} = \begin{pmatrix} U_1^\top B V_1 & U_1^\top B V_2 \\ U_2^\top B V_1 & U_2^\top B V_2 \end{pmatrix};$$

assim

$$D_1 = U_1^\top B V_1 = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_b),$$

$$D_2 = U_2^\top B V_2 = O_{k-b, q-b},$$

onde $O_{k-b, q-b}$ é uma matriz nula de dimensão $k-b \times q-b$, e as expressões

$$U_1^\top B V_2 = O_{b, q-b} \quad \text{e} \quad U_2^\top B V_1 = O_{k-b, b}$$

são válidas, pois $B V_2 = 0$ e $U_2^\top B = (B^\top U_2)^\top = 0$.

Seja $\hat{B}$ um estimador da matriz B ; então, sua decomposição em valores singulares é

$$\hat{B} = \hat{U} \hat{D} \hat{V}^\top = \begin{pmatrix} \hat{U}_1 & \hat{U}_2 \end{pmatrix} \begin{pmatrix} \hat{D}_1 & 0 \\ 0 & \hat{D}_2 \end{pmatrix} \begin{pmatrix} \hat{V}_1^\top \\ \hat{V}_2^\top \end{pmatrix}.$$

Particionamos as matrizes $\hat{U}$, $\hat{V}$, e $\hat{D}$ de acordo com a partição feita para U , V , e D , onde $\hat{U}_1$ tem b colunas e $\hat{U}_2$ tem $k-b$ colunas. Os espaços gerados por $\hat{U}_1$, $\hat{U}_2$, $\hat{V}_1$ e $\hat{V}_2$ são aproximações para os subespaços correspondentes a U_1 , U_2 , V_1 e V_2 . A matriz diagonal $\hat{D}$ pode ser escrita como

$$\hat{D} = \begin{pmatrix} \hat{U}_1^\top \\ \hat{U}_2^\top \end{pmatrix} \hat{B} \begin{pmatrix} \hat{V}_1 & \hat{V}_2 \end{pmatrix} = \begin{pmatrix} \hat{U}_1^\top \hat{B} \hat{V}_1 & \hat{U}_1^\top \hat{B} \hat{V}_2 \\ \hat{U}_2^\top \hat{B} \hat{V}_1 & \hat{U}_2^\top \hat{B} \hat{V}_2 \end{pmatrix},$$

que satisfaz as equações

$$\widehat{D}_1 = \widehat{U}_1^\top \widehat{B} \widehat{V}_1 = \text{diag}(\widehat{\lambda}_1, \widehat{\lambda}_2, \dots, \widehat{\lambda}_b) \quad \text{com} \quad \widehat{\lambda}_i \geq \widehat{\lambda}_{i+1}$$

$$\widehat{D}_2 = \widehat{U}_2^\top \widehat{B} \widehat{V}_2 = \text{diag}(\widehat{\lambda}_{b+1}, \widehat{\lambda}_{b+2}, \dots, \widehat{\lambda}_q)$$

e

$$\widehat{U}_2^\top \widehat{B} \widehat{V}_1 = 0 = \widehat{U}_1^\top \widehat{B} \widehat{V}_2$$

Definimos

$$L = U_2^\top \widehat{B} V_2,$$

que define uma matriz aleatória como função da matriz aleatória $\widehat{B}$. Se a hipótese nula é verdadeira então $D_2 = O_{k-b, q-b}$. Neste caso, o vetor $\sqrt{N} l$, $l = \text{vec} L$, é assintoticamente distribuído como uma normal com média zero e matriz de covariância Q , de dimensão $(k-b)(q-b) \times (k-b)(q-b)$, finita e positiva definida, dada por

$$Q = (V_2^\top \otimes U_2^\top) \Sigma (V_2 \otimes U_2).$$

Sob estas condições e sob a hipótese nula

$$H_0 : r(B) = b,$$

a estatística de teste, proposta por Ratsimalahelo (2003), dada por

$$L(b) = N \widehat{l}^\top \widehat{Q}^{-1} \widehat{l},$$

onde

$$\widehat{Q}^{-1} = (\widehat{V}_2^\top \otimes \widehat{U}_2^\top) \widehat{\Sigma}^{-1} (\widehat{V}_2 \otimes \widehat{U}_2) \quad \text{e} \quad \widehat{l} = \text{vec}(\widehat{U}_2^\top \widehat{B} \widehat{V}_2),$$

converge em distribuição para uma qui-quadrado com $(k-b)(q-b)$ graus de liberdade, ou seja,

$$L(b) \xrightarrow{d} \chi_{(k-b)(q-b)}^2.$$

Três procedimentos foram considerados para estimar o posto de B : o teste seqüencial e os critérios de informação BIC e HQIC. Estes procedimentos estão descritos nas próximas seções juntamente com condições para consistência forte dos estimadores baseados nos mesmos.

3.2. Procedimento de Teste Seqüencial

Sabemos que o procedimento chamado de teste seqüencial não conduz a uma estimativa consistente do verdadeiro valor do posto da matriz a não ser que algum ajuste seja feito no nível de significância, ver Cragg e Donald (1997); Robin e Smith (2000). Ratsimalahelo (2003) propõe um nível de significância apropriado para a obtenção de consistência forte

do estimador do posto da matriz B usando o procedimento de teste seqüencial, supondo consistência forte de $\hat{B}$.

A estatística de teste $L(b)$ tem distribuição assintótica qui-quadrado com $(k-b)(q-b)$ graus de liberdade, ou ainda, $\chi^2_{(k-b)(q-b)}$. Seja $\gamma_b C_N > 0$ o valor crítico da estatística de teste $L(b)$ para o nível de significância escolhido. O valor γ_b representa o quantil correspondente da distribuição $\chi^2_{(k-b)(q-b)}$ e C_N é uma seqüência predeterminada de números. Para um estimador fortemente consistente do posto da matriz B , Ratsimalahelo (2003) assume que a função C_N satisfaz as três condições

- (i) $C_N > 0$
 - (ii) $C_N/N \rightarrow 0$
 - (iii) $C_N/\log \log(N) \rightarrow \infty$
- (3.2.1)

Há várias possíveis escolhas de C_N satisfazendo (i)-(iii) de (3.2.1); para o nosso caso escolhemos $C_N = \sqrt{\log(N)}$.

O teste seqüencial é feito da seguinte forma: Começamos com $b = 0$, daí verificamos se $L(b) \leq \gamma_b C_N$, se isso acontecer então o estimador de b será 0, ou seja, $\hat{b} = 0$, senão verificaremos a desigualdade quando $b = 1$. E esse processo é repetido até que $L(b) \leq \gamma_b C_N$, onde $b \in \{0, 1, 2, \dots, q-1\}$, se a desigualdade anterior não for satisfeita, então, consideramos $\hat{b} = q$.

3.3. Critérios de Informação BIC e HQIC

Agora consideramos a aproximação do critério de informação para estimativa do posto da matriz.

A função critério de informação é definida da seguinte forma:

$$IC(b, C_N) = L(b) + \varphi(b) C_N,$$

com $b = 0, 1, 2, \dots, q$, onde a função $\varphi(b)$ é estritamente crescente e a constante C_N é o termo de penalidade que satisfaz as condições dadas em (3.2.1). Quando $b = q$ consideramos $L(b) = 0$.

O estimador $\hat{b}$ do posto b é o valor que minimiza a função critério, ou seja,

$$\hat{b} = \arg \min IC(b, C_N),$$

com $b = 0, 1, 2, \dots, q$. Ratsimalahelo (2003) mostra que $\hat{b}$ é estimador fortemente consistente de b , ou seja, $\hat{b}$ converge quase certamente para b , ($\hat{b} \xrightarrow{q.c.} b$).

Na literatura estatística, algumas escolhas fixas de C_N têm sido sugeridas tais como o critério em que $C_N = 2$ por Akaike(1974), chamado de AIC; $C_N = c \log \log(N)$ para algum $c > 2$ por Hannan e Quinn (1979), chamado de HQIC; $C_N = \log(N)$ por Schwarz (1978), chamado BIC. O procedimento AIC não satisfaz a condição $C_N/\log \log(N) \rightarrow \infty$, em conseqüência não podemos garantir que o estimador baseado no AIC é consistente.

Para o cálculo da função critério consideramos neste trabalho os critérios de informação BIC, HQIC e $\varphi(b) = b$, com $b = 0, 1, 2, \dots, q$.

Capítulo 4

Avaliação numérica

Neste capítulo apresentamos os resultados de simulação para as estimativas do posto da matriz B dos parâmetros do modelo de regressão Dirichlet. Todo procedimento de cálculo foi programado na linguagem de programação 0x e o programa de simulação está apresentado no apêndice 5.

Para simulação utilizamos o método de Monte Carlo. Consideramos 10.000 réplicas de Monte Carlo e para os tamanhos de amostra usamos $N = 50, 100, 150$ e 200 . Os números de regressores da matriz X considerados foram $k = 3, 4, 5, 6$ e os números de componentes dos vetores de observações foram $p = 2, 3, 4$ para os modelos 1 e 2, lembrando que, na primeira parametrização, a matriz B tem dimensão $k \times (p - 1)$ e na segunda tem dimensão $k \times p$, onde em todos os casos assumimos, sem perda de generalidade, que $k \geq (p - 1)$ no modelo 1 e $k \geq p$ no modelo 2.

Para os valores verdadeiros dos parâmetros da matriz B consideramos a base canônica; por exemplo, se $k = 3, p = 4$, no modelo 1, e o posto verdadeiro da matriz B for b , temos

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ se } b = 1; \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ se } b = 2; \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ se } b = 3;$$

consideramos $b = 0, 1, 2$ e 3 ; para $b = 0$ a matriz B é nula, e para o parâmetro de precisão verdadeiro usamos $\phi = 30$ no modelo 1. A função de ligação que usamos no modelo 1 foi a logit, ou seja, $g(\eta_{ij}) = e^{\eta_{ij}} / (1 + e^{\eta_{ij}})$, $j = 1, 2, \dots, p - 1$, enquanto no modelo 2 usamos $g(\eta_{ij}) = e^{\eta_{ij}}$, $j = 1, 2, \dots, p$.

Para cada réplica de Monte Carlo geramos uma amostra aleatória da variável resposta $Y = (y_{i1}, y_{i2}, \dots, y_{ip})$, com $i = 1, 2, \dots, N$, obtida através da parametrização (2.2.2.3) da distribuição Dirichlet. Os valores da covariável X foram obtidos aleatoriamente da distribuição uniforme no intervalo $(0, 1)$, $\mathcal{U}(0, 1)$; ressaltamos que os valores de X permanecem constantes em todo o experimento, ou seja, X é gerado apenas uma vez antes do início do método de Monte Carlo. No primeiro modelo, dados os valores da matriz resposta e da matriz de regressores, avaliamos a regressão auxiliar

$$(X^\top X)^{-1} X^\top Z,$$

onde Z é a matriz da variável resposta transformada definida no capítulo 2. Tomamos as estimativas dos parâmetros obtidas da regressão auxiliar como valores iniciais para o processo de maximização da função de log-verossimilhança (2.2.2.4). A função de log-verossimilhança é maximizada através do método BFGS, discutido no Apêndice 3, e os valores do ponto de máximo desta função são as estimativas $\hat{B}$ e $\hat{\phi}$ de máxima verossimilhança dos parâmetros. Para o modelo 2 tomamos a matriz nula como valor inicial no processo de maximização da

função de log-verossimilhança, onde também foi usado o método BFGS, e o valor do ponto de máximo desta função é a estimativa $\hat{B}$ de máxima verossimilhança.

Em seguida usamos a decomposição em valores singulares de $\hat{B}$, descrita no capítulo 3, e preenchemos a matriz de informação de Fischer para calcular a estatística de teste, descrita na seção (3.1.3), que é dada por

$$L(b) = N \hat{l}^\top \hat{Q}^{-1} \hat{l},$$

com

$$\hat{Q}^{-1} = (\hat{V}_2^\top \otimes \hat{U}_2^\top) \hat{\Sigma}^{-1} (\hat{V}_2 \otimes \hat{U}_2),$$

onde $\hat{\Sigma}^{-1} = (I^{\beta\beta})^{-1}$, quando o primeiro modelo é usado, sendo $I^{\beta\beta}$ definido na seção 2.2.2 e $\hat{\Sigma}^{-1} = I(B)$, definido na seção 2.2.3, se usamos o segundo modelo. Como já foi dito no capítulo 3, a estatística de teste $L(b)$ converge em distribuição para uma qui-quadrado com $(k-b)(q-b)$ graus de liberdade; nos modelos 1 e 2 temos que $q = p-1$ e p , respectivamente, e o nível de significância usado em ambos os casos foi de 5%.

Para estimar o posto b da matriz B utilizamos três procedimentos: o teste seqüencial e os critérios de informação BIC e HQIC. Através desses procedimentos estimamos b , usando a estatística de teste $L(b)$, para cada réplica de Monte Carlo. Depois calculamos, para cada tamanho de amostra e para cada procedimento, a proporção em que cada posto estimado é $0, 1, 2, \dots, p-1$ para o modelo 1 e $0, 1, 2, \dots, p$ para o modelo 2. O viés e o erro quadrático médio do estimador $\hat{b}$ também foram calculados para a análise dos resultados.

Nas próximas seções estão descritos os resultados numéricos para as duas parametrizações consideradas neste trabalho.

4.1. Resultados Numéricos para o Modelo 1

Ao longo desta seção, apresentamos os resultados da simulação feita para a estimação do posto b referente ao primeiro modelo considerado, que está descrito na seção 2.2.2.

Os resultados da simulação estão expostos em tabelas. Em cada tabela, para diferentes tamanhos de amostra N e dimensões da matriz B , apresentamos: o valor verdadeiro do posto da matriz B , a proporção de vezes em que o teste seqüencial e os critérios de informação BIC e HQIC acertam o verdadeiro posto de B , o viés e o erro quadrático médio (EQM) do estimador $\hat{b}$.

Nas tabelas numeradas de 4.1 a 4.6 consideramos $b = 0$; isto significa que a matriz B é nula, ou seja, o número de valores singulares diferentes de zero da matriz B é nulo. Notamos, através destas tabelas, que os vieses e os EQM's das estimativas de b são próximos de zero quando o teste usado para estimação de b é o seqüencial. Por exemplo, na tabela 4.5 observamos que o viés e o EQM de $\hat{b}$ são, respectivamente, 0.0007 e 0.0013, se $N = 50$ e B tem dimensão 3×3 .

Quando consideramos $b = 1$, $k = 3, 4, 5, 6$ e $p - 1 = 1$, em todas as 10.000 réplicas de Monte Carlo o teste seqüencial e os critérios de informação BIC e HQIC apontaram que o posto estimado $\hat{b}$ é igual a um, isto pode ser visto nas tabelas 4.7 e 4.8.

Observamos que, na maioria das vezes, se aumentamos o tamanho da amostra N , o viés e o EQM de $\hat{b}$ diminuem, como esperado, pois sabemos que $\hat{b}$ é assintoticamente não-viesado e consistente. Por exemplo, para $N = 50, 100, 150$ e 200 , os EQM's das estimativas de b são $0.3122, 0.2177, 0.1971$ e 0.1862 , respectivamente, como se pode observar na tabela 4.9, quando B tem dimensão 3×2 , $b = 1$ e o teste usado para estimação de b é o seqüencial. Outro exemplo pode ser visto na tabela 4.15, se $k = 4$, $b = 2$ e $p - 1 = 3$, para $N = 50, 100, 150$ e 200 ; os vieses das estimativas de b são $0.6303, 0.4972, 0.4623$ e 0.4304 , respectivamente, quando o critério de informação usado para estimação de b é o BIC. À medida que acrescentamos variáveis explicativas no modelo, os resultados da estimação de b , usando os três testes já mencionados, são piores na maioria dos casos, ou seja, os vieses de $\hat{b}$ aumentam. Por exemplo, nas tabelas 4.11 e 4.12 temos que se $k = 3, 4, 5$ e 6 , $b = 1$ e $N = 200$ os vieses das estimativas de b são $0.3287, 0.5144, 0.5673$ e 0.6391 , respectivamente, quando o teste usado é o seqüencial e $1.0051, 1.3321, 1.5493$ e 1.6964 para o critério de informação HQIC.

Para todas as dimensões da matriz B consideradas, percebemos que quando aumentamos o número de componentes p da matriz Y , o teste seqüencial e os critérios de informação BIC e HQIC tornam-se cada vez mais imprecisos. Isto pode ser visto, por exemplo, nas tabelas 4.10 e 4.12 onde $p = 3$ e 4 , respectivamente, e em ambos os casos, $b = 1$; observamos para $N = 100$ que, se $p = 3$ o viés de $\hat{b}$ é 0.7512 e se $p = 4$ o viés é 1.4408 , ou seja, as estimativas do posto da matriz B tornam-se mais viesadas quando p cresce.

Consideramos vários casos em que a matriz B , de dimensão $k \times p - 1$, tem posto completo, $b = p - 1$, ou seja, casos em que a matriz B tem $p - 1$ valores singulares diferentes de zero. Estes casos estão apresentados nas tabelas 4.7, 4.8, 4.13, 4.14, 4.17 e 4.18.

Nos resultados de simulação percebemos que se a matriz B tem posto completo, o critério de informação HQIC tem desempenho ligeiramente melhor que o BIC e o teste seqüencial. Isso pode ser visto, por exemplo, na tabela 4.13 quando B tem dimensão 3×2 , $b = 2$ e $N = 50$; o critério HQIC acerta em $97,87\%$ das 10.000 réplicas o posto da matriz B , enquanto o critério BIC acerta em $96,27\%$ e o teste seqüencial em $84,02\%$. Como já foi dito, se aumentamos o valor de p , por exemplo, quando B tem dimensão 3×3 e $N = 50$, os testes acertam menos, mas o critério HQIC continua um pouco melhor que o BIC e o seqüencial; o critério HQIC acerta em $74,85\%$ das 10.000 réplicas o posto da matriz B , enquanto o critério BIC acerta em $68,72\%$ e o teste seqüencial em $54,93\%$, como pode ser visto na tabela 4.17. Quando a matriz B não tem posto completo, o teste seqüencial tem um desempenho bem melhor que os critérios de informação BIC e HQIC. Por exemplo, na tabela 4.16 se $k = 6$, $p - 1 = 3$, $b = 2$ e $N = 200$, o teste seqüencial acerta em $61,98\%$ das 10.000 réplicas o posto da matriz B , enquanto o critério BIC acerta em $25,53\%$ e o HQIC em $14,96\%$. Outro exemplo pode ser visto na tabela 4.12 quando $k = 6$, $p - 1 = 3$, $b = 1$ e $N = 200$; o teste seqüencial acerta em $50,81\%$ das vezes o posto da matriz B , enquanto o critério BIC acerta em $3,85\%$ e o HQIC em 0.92% .

Tabela 4.1. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 3, 4$ e $p - 1 = 1$.

$b = 0$	Dimensão: 3×1				Dimensão: 4×1			
	Seqüencial				Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,9945	0,9983	0,9992	0,9996	0,9953	0,9991	0,9992	0,9998
1	0,0055	0,0017	0,0008	0,0004	0,0047	0,0009	0,0008	0,0002
Vies($\hat{b}$)	0,0055	0,0017	0,0008	0,0004	0,0047	0,0009	0,0008	0,0002
EQM($\hat{b}$)	0,0055	0,0017	0,0008	0,0004	0,0047	0,0009	0,0008	0,0002
	BIC				BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,6971	0,7810	0,8122	0,8428	0,5388	0,6444	0,6971	0,7345
1	0,3029	0,2190	0,1878	0,1572	0,4612	0,3556	0,3029	0,2655
Vies($\hat{b}$)	0,3029	0,2190	0,1878	0,1572	0,4612	0,3556	0,3029	0,2655
EQM($\hat{b}$)	0,3029	0,2190	0,1878	0,1572	0,4612	0,3556	0,3029	0,2655
	HQIC				HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,5399	0,6044	0,6310	0,6537	0,3660	0,4350	0,4679	0,4908
1	0,4601	0,3956	0,3690	0,3463	0,6340	0,5650	0,5321	0,5092
Vies($\hat{b}$)	0,4601	0,3956	0,3690	0,3463	0,6340	0,5650	0,5321	0,5092
EQM($\hat{b}$)	0,4601	0,3956	0,3690	0,3463	0,6340	0,5650	0,5321	0,5092

Tabela 4.2. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 5$, 6 e $p - 1 = 1$.

$b = 0$	Dimensão: 5×1				Dimensão: 6×1			
	Seqüencial				Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,9966	0,9994	0,9995	0,9998	0,9950	0,9993	0,9993	0,9999
1	0,0034	0,0006	0,0005	0,0002	0,0050	0,0007	0,0007	0,0001
Vies($\hat{b}$)	0,0034	0,0006	0,0005	0,0002	0,0050	0,0007	0,0007	0,0001
EQM($\hat{b}$)	0,0034	0,0006	0,0005	0,0002	0,0050	0,0007	0,0007	0,0001
	BIC				BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,3866	0,5048	0,5644	0,6009	0,2496	0,3642	0,4252	0,4739
1	0,6134	0,4952	0,4356	0,3991	0,7504	0,6358	0,5748	0,5261
Vies($\hat{b}$)	0,6134	0,4952	0,4356	0,3991	0,7504	0,6358	0,5748	0,5261
EQM($\hat{b}$)	0,6134	0,4952	0,4356	0,3991	0,7504	0,6358	0,5748	0,5261
	HQIC				HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,2296	0,2871	0,3218	0,3371	0,1254	0,1756	0,2048	0,2305
1	0,7704	0,7129	0,6782	0,6629	0,8746	0,8244	0,7952	0,7695
Vies($\hat{b}$)	0,7704	0,7129	0,6782	0,6629	0,8746	0,8244	0,7952	0,7695
EQM($\hat{b}$)	0,7704	0,7129	0,6782	0,6629	0,8746	0,8244	0,7952	0,7695

Tabela 4.3. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 3, 4$ e $p - 1 = 2$.

$b = 0$		Dimensão: 3×2				Dimensão: 4×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,9988	1,0000	1,0000	1,0000	0,9988	0,9999	1,0000	0,9999
1		0,0009	0,0000	0,0000	0,0000	0,0010	0,0001	0,0000	0,0001
2		0,0003	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
Vies($\hat{b}$)		0,0015	0,0000	0,0000	0,0000	0,0014	0,0001	0,0000	0,0001
EQM($\hat{b}$)		0,0021	0,0000	0,0000	0,0000	0,0018	0,0001	0,0000	0,0001
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,4446	0,5588	0,6290	0,6659	0,2634	0,3737	0,4444	0,4913
1		0,4772	0,3959	0,3407	0,3114	0,5555	0,5190	0,4835	0,4483
2		0,0782	0,0453	0,0303	0,0227	0,1811	0,1073	0,0721	0,0604
Vies($\hat{b}$)		0,6336	0,4865	0,4013	0,3568	0,9177	0,7336	0,6277	0,5691
EQM($\hat{b}$)		0,7900	0,5771	0,4619	0,4022	1,2799	0,9482	0,7719	0,6899
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,2576	0,3263	0,3604	0,3926	0,1208	0,1698	0,1965	0,2152
1		0,5738	0,5485	0,5378	0,5123	0,5524	0,5778	0,5924	0,5878
2		0,1686	0,1252	0,1018	0,0951	0,3268	0,2524	0,2111	0,1970
Vies($\hat{b}$)		0,9110	0,7989	0,7414	0,7025	1,2060	1,0826	1,0146	0,9818
EQM($\hat{b}$)		1,2482	1,0493	0,9450	0,8927	1,8596	1,5874	1,4368	1,3758

Tabela 4.4. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 5$, 6 e $p - 1 = 2$.

$b = 0$		Dimensão: 5×2				Dimensão: 6×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,9992	1,0000	1,0000	1,0000	0,9987	1,0000	1,0000	1,0000
1		0,0006	0,0000	0,0000	0,0000	0,0011	0,0000	0,0000	0,0000
2		0,0002	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
Vies($\hat{b}$)		0,0010	0,0000	0,0000	0,0000	0,0015	0,0000	0,0000	0,0000
EQM($\hat{b}$)		0,0014	0,0000	0,0000	0,0000	0,0019	0,0000	0,0000	0,0000
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,1349	0,2247	0,2831	0,3336	0,0605	0,1194	0,1666	0,2063
1		0,5522	0,5878	0,5736	0,5467	0,4733	0,5625	0,5813	0,5764
2		0,3129	0,1875	0,1433	0,1197	0,4662	0,3181	0,2521	0,2173
Vies($\hat{b}$)		1,1780	0,9628	0,8602	0,7861	1,4057	1,1987	1,0855	1,0110
EQM($\hat{b}$)		1,8038	1,3378	1,1468	1,0255	2,3381	1,8349	1,5897	1,4456
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0485	0,0759	0,0904	0,1057	0,0159	0,0265	0,0380	0,0432
1		0,4610	0,5272	0,5546	0,5600	0,3231	0,4144	0,4477	0,4689
2		0,4905	0,3969	0,3550	0,3343	0,6610	0,5591	0,5143	0,4879
Vies($\hat{b}$)		1,4420	1,3210	1,2646	1,2286	1,6451	1,5326	1,4763	1,4447
EQM($\hat{b}$)		2,4230	2,1148	1,9746	1,8972	2,9671	2,6508	2,5049	2,4205

Tabela 4.5. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 3, 4$ e $p - 1 = 3$.

$b = 0$	Dimensão: 3×3				Dimensão: 4×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,9995	0,9999	1,0000	1,0000	0,9998	1,0000	0,9999	0,9999
1	0,0004	0,0001	0,0000	0,0000	0,0002	0,0000	0,0001	0,0001
2	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
3	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Vies($\hat{b}$)	0,0007	0,0001	0,0000	0,0000	0,0002	0,0000	0,0001	0,0001
EQM($\hat{b}$)	0,0013	0,0001	0,0000	0,0000	0,0002	0,0000	0,0001	0,0001
				BIC				
$\hat{b}$								
0	0,2749	0,3804	0,4498	0,4932	0,1242	0,2078	0,2636	0,3003
1	0,5297	0,4998	0,4572	0,4341	0,4784	0,5201	0,5320	0,5196
2	0,1739	0,1126	0,0869	0,0686	0,3401	0,2438	0,1869	0,1654
3	0,0215	0,0072	0,0061	0,0041	0,0573	0,0283	0,0175	0,0147
Vies($\hat{b}$)	0,9420	0,7466	0,6493	0,5836	1,3305	1,0926	0,9583	0,8945
EQM($\hat{b}$)	1,4188	1,0150	0,8597	0,7454	2,3545	1,7500	1,4371	1,3135
				HQIC				
$\hat{b}$								
0	0,1208	0,1689	0,1929	0,2105	0,0394	0,0600	0,0776	0,0836
1	0,5114	0,5392	0,5483	0,5502	0,3548	0,4244	0,4625	0,4638
2	0,3083	0,2546	0,2284	0,2119	0,4723	0,4132	0,3780	0,3807
3	0,0595	0,0373	0,0304	0,0274	0,1335	0,1024	0,0819	0,0719
Vies($\hat{b}$)	1,3065	1,1603	1,0963	1,0562	1,6999	1,5580	1,4642	1,4409
EQM($\hat{b}$)	2,2801	1,8933	1,7355	1,6444	3,4455	2,9988	2,7116	2,6337

Tabela 4.6. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 5, 6$ e $p - 1 = 3$.

$b = 0$	Dimensão: 5×3					Dimensão: 6×3				
	Sequencial					Sequencial				
	N = 50	N = 100	N = 150	N = 200		N = 50	N = 100	N = 150	N = 200	
$\hat{b}$										
0	0,9998	0,9999	0,9999	1,0000		0,9997	1,0000	1,0000	1,0000	
1	0,0002	0,0001	0,0001	0,0000		0,0002	0,0000	0,0000	0,0000	
2	0,0000	0,0000	0,0000	0,0000		0,0001	0,0000	0,0000	0,0000	
3	0,0000	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
Vies($\hat{b}$)	0,0002	0,0001	0,0001	0,0000		0,0004	0,0000	0,0000	0,0000	
EQM($\hat{b}$)	0,0002	0,0001	0,0001	0,0000		0,0006	0,0000	0,0000	0,0000	
BIC						BIC				
$\hat{b}$										
0	0,0465	0,0988	0,1323	0,1751		0,0144	0,0387	0,0633	0,0806	
1	0,3481	0,4535	0,4915	0,4995		0,2072	0,3347	0,3890	0,4219	
2	0,4724	0,3729	0,3254	0,2903		0,5268	0,4787	0,4439	0,4116	
3	0,1330	0,0748	0,0508	0,0351		0,2516	0,1479	0,1038	0,0859	
Vies($\hat{b}$)	1,6919	1,4237	1,2947	1,1854		2,0156	1,7358	1,5882	1,5028	
EQM($\hat{b}$)	3,4347	2,6183	2,2503	1,9766		4,5788	3,5806	3,0988	2,8414	
HQIC						HQIC				
$\hat{b}$										
0	0,0106	0,0176	0,0219	0,0302		0,0015	0,0043	0,0077	0,0086	
1	0,1892	0,2729	0,2924	0,3225		0,0827	0,1398	0,1650	0,1870	
2	0,5293	0,5053	0,5140	0,4938		0,4766	0,5224	0,5380	0,5366	
3	0,2709	0,2042	0,1717	0,1535		0,4392	0,3335	0,2893	0,2678	
Vies($\hat{b}$)	2,0605	1,8961	1,8355	1,7706		2,3535	2,1851	2,1089	2,0636	
EQM($\hat{b}$)	4,7445	4,1319	3,8937	3,6792		5,9419	5,2309	4,9207	4,7436	

Tabela 4.7. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 3, 4$ e $p - 1 = 1$.

$b = 1$	Dimensão: 3×1				Dimensão: 4×1			
	Seqüencial				Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$V_{ies}(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$EQM(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	BIC				BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$V_{ies}(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$EQM(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	HQIC				HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$V_{ies}(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$EQM(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Tabela 4.8. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 5$, 6 e $p - 1 = 1$.

$b = 1$	Dimensão: 5×1				Dimensão: 6×1			
	Seqüencial				Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$V_{ies}(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$EQM(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	BIC				BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$V_{ies}(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$EQM(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	HQIC				HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
$V_{ies}(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
$EQM(\hat{b})$	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Tabela 4.9. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 3, 4$ e $p - 1 = 2$.

$b = 1$		Dimensão: 3×2				Dimensão: 4×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,6878	0,7823	0,8029	0,8138	0,5681	0,6903	0,7351	0,7455
2		0,3122	0,2177	0,1971	0,1862	0,4319	0,3097	0,2649	0,2545
Vies($\hat{b}$)		0,3122	0,2177	0,1971	0,1862	0,4319	0,3097	0,2649	0,2545
EQM($\hat{b}$)		0,3122	0,2177	0,1971	0,1862	0,4319	0,3097	0,2649	0,2545
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,4379	0,5273	0,5509	0,5582	0,2652	0,3470	0,3845	0,3919
2		0,5621	0,4727	0,4491	0,4418	0,7348	0,6530	0,6155	0,6081
Vies($\hat{b}$)		0,5621	0,4727	0,4491	0,4418	0,7348	0,6530	0,6155	0,6081
EQM($\hat{b}$)		0,5621	0,4727	0,4491	0,4418	0,7348	0,6530	0,6155	0,6081
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,3568	0,4181	0,4287	0,4247	0,1894	0,2457	0,2566	0,2673
2		0,6432	0,5819	0,5713	0,5753	0,8106	0,7543	0,7434	0,7327
Vies($\hat{b}$)		0,6432	0,5819	0,5713	0,5753	0,8106	0,7543	0,7434	0,7327
EQM($\hat{b}$)		0,6432	0,5819	0,5713	0,5753	0,8106	0,7543	0,7434	0,7327

Tabela 4.10. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 5, 6$ e $p - 1 = 2$.

$b = 1$		Dimensão: 5×2				Dimensão: 6×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,5288	0,6581	0,6877	0,7135	0,4552	0,6102	0,6721	0,7086
2		0,4712	0,3419	0,3123	0,2865	0,5448	0,3898	0,3279	0,2914
Vies($\hat{b}$)		0,4712	0,3419	0,3123	0,2865	0,5448	0,3898	0,3279	0,2914
EQM($\hat{b}$)		0,4712	0,3419	0,3123	0,2865	0,5448	0,3898	0,3279	0,2914
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,1772	0,2488	0,2624	0,2976	0,1070	0,1624	0,1922	0,2122
2		0,8228	0,7512	0,7376	0,7024	0,8930	0,8376	0,8078	0,7878
Vies($\hat{b}$)		0,8228	0,7512	0,7376	0,7024	0,8930	0,8376	0,8078	0,7878
EQM($\hat{b}$)		0,8228	0,7512	0,7376	0,7024	0,8930	0,8376	0,8078	0,7878
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,1131	0,1530	0,1559	0,1733	0,0604	0,0888	0,0961	0,1071
2		0,8869	0,8470	0,8441	0,8267	0,9396	0,9112	0,9039	0,8929
Vies($\hat{b}$)		0,8869	0,8470	0,8441	0,8267	0,9396	0,9112	0,9039	0,8929
EQM($\hat{b}$)		0,8869	0,8470	0,8441	0,8267	0,9396	0,9112	0,9039	0,8929

Tabela 4.11. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 3$, 4 e $p - 1 = 3$.

$b = 1$	Dimensão: 3×3				Dimensão: 4×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,4201	0,5956	0,6590	0,7026	0,2925	0,4431	0,5392	0,5608
2	0,4794	0,3607	0,3033	0,2661	0,5237	0,4473	0,3792	0,3640
3	0,1005	0,0437	0,0377	0,0313	0,1838	0,1096	0,0816	0,0752
Vies($\hat{b}$)	0,6804	0,4481	0,3787	0,3287	0,8913	0,6665	0,5424	0,5144
EQM($\hat{b}$)	0,8814	0,5355	0,4541	0,3913	1,2589	0,8857	0,7056	0,6648
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,1538	0,2265	0,2638	0,2837	0,0600	0,0953	0,1316	0,1381
2	0,6494	0,6583	0,6278	0,6224	0,5307	0,5902	0,5820	0,5894
3	0,1968	0,1152	0,1084	0,0939	0,4093	0,3145	0,2864	0,2725
Vies($\hat{b}$)	1,0430	0,8887	0,8446	0,8102	1,3493	1,2192	1,1548	1,1344
EQM($\hat{b}$)	1,4366	1,1191	1,0614	0,9980	2,1679	1,8482	1,7276	1,6794
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,1009	0,1390	0,1579	0,1626	0,0325	0,0471	0,0633	0,0613
2	0,6422	0,6859	0,6689	0,6697	0,4724	0,5333	0,5354	0,5453
3	0,2569	0,1751	0,1732	0,1677	0,4951	0,4196	0,4013	0,3934
Vies($\hat{b}$)	1,1560	1,0361	1,0153	1,0051	1,4626	1,3725	1,3380	1,3321
EQM($\hat{b}$)	1,6698	1,3863	1,3617	1,3405	2,4528	2,2117	2,1406	2,1189

Tabela 4.12. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 5$, 6 e $p - 1 = 3$.

b = 1	Dimensão: 5×3				Dimensão: 6×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,2148	0,4144	0,4810	0,5369	0,1625	0,3362	0,4309	0,5081
2	0,5063	0,4192	0,3907	0,3589	0,4665	0,4304	0,3836	0,3447
3	0,2789	0,1664	0,1283	0,1042	0,3710	0,2334	0,1855	0,1472
Vies($\hat{b}$)	1,0641	0,7520	0,6473	0,5673	1,2085	0,8972	0,7546	0,6391
EQM($\hat{b}$)	1,6219	1,0848	0,9039	0,7757	1,9505	1,3640	1,1256	0,9335
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0212	0,0538	0,0651	0,0805	0,0079	0,0222	0,0341	0,0385
2	0,3810	0,4516	0,4826	0,4844	0,2586	0,3367	0,3717	0,4031
3	0,5978	0,4946	0,4523	0,4351	0,7335	0,6411	0,5942	0,5584
Vies($\hat{b}$)	1,5766	1,4408	1,3872	1,3546	1,7256	1,6189	1,5601	1,5199
EQM($\hat{b}$)	2,7722	2,4300	2,2918	2,2248	3,1926	2,9011	2,7485	2,6367
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0095	0,0218	0,0217	0,0259	0,0030	0,0071	0,0104	0,0092
2	0,2975	0,3624	0,3891	0,3989	0,1836	0,2427	0,2635	0,2852
3	0,6930	0,6158	0,5892	0,5752	0,8134	0,7502	0,7261	0,7056
Vies($\hat{b}$)	1,6835	1,5940	1,5675	1,5493	1,8104	1,7431	1,7157	1,6964
EQM($\hat{b}$)	3,0695	2,8256	2,7459	2,6997	3,4372	3,2435	3,1679	3,1076

Tabela 4.13. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 3$, 4 e $p - 1 = 2$.

b = 2		Dimensão: 3 × 2				Dimensão: 4 × 2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,1598	0,0604	0,0138	0,0015	0,1397	0,0960	0,0297	0,0110
2		0,8402	0,9396	0,9862	0,9985	0,8603	0,9040	0,9703	0,9890
Vies($\hat{b}$)		-0,1598	-0,0604	-0,0138	-0,0015	-0,1397	-0,0960	-0,0297	-0,0110
EQM($\hat{b}$)		0,1598	0,0604	0,0138	0,0015	0,1397	0,0960	0,0297	0,0110
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0373	0,0044	0,0003	0,0000	0,0179	0,0050	0,0004	0,0003
2		0,9627	0,9956	0,9997	1,0000	0,9821	0,9950	0,9996	0,9997
Vies($\hat{b}$)		-0,0373	-0,0044	-0,0003	0,0000	-0,0179	-0,0050	-0,0004	-0,0003
EQM($\hat{b}$)		0,0373	0,0044	0,0003	0,0000	0,0179	0,0050	0,0004	0,0003
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0213	0,0009	0,0001	0,0000	0,0094	0,0018	0,0001	0,0000
2		0,9787	0,9991	0,9999	1,0000	0,9906	0,9982	0,9999	1,0000
Vies($\hat{b}$)		-0,0213	-0,0009	-0,0001	0,0000	-0,0094	-0,0018	-0,0001	0,0000
EQM($\hat{b}$)		0,0213	0,0009	0,0001	0,0000	0,0094	0,0018	0,0001	0,0000

Tabela 4.14. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 5$, 6 e $p - 1 = 2$.

$b = 2$		Dimensão: 5×2				Dimensão: 6×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,1258	0,0703	0,0284	0,0080	0,1368	0,0984	0,0502	0,0183
2		0,8742	0,9297	0,9716	0,9920	0,8632	0,9016	0,9498	0,9817
Vies($\hat{b}$)		-0,1258	-0,0703	-0,0284	-0,0080	-0,1368	-0,0984	-0,0502	-0,0183
EQM($\hat{b}$)		0,1258	0,0703	0,0284	0,0080	0,1368	0,0984	0,0502	0,0183
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0100	0,0007	0,0000	0,0000	0,0057	0,0013	0,0002	0,0000
2		0,9900	0,9993	1,0000	1,0000	0,9943	0,9987	0,9998	1,0000
Vies($\hat{b}$)		-0,0100	-0,0007	0,0000	0,0000	-0,0057	-0,0013	-0,0002	0,0000
EQM($\hat{b}$)		0,0100	0,0007	0,0000	0,0000	0,0057	0,0013	0,0002	0,0000
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0056	0,0003	0,0000	0,0000	0,0025	0,0003	0,0000	0,0000
2		0,9944	0,9997	1,0000	1,0000	0,9975	0,9997	1,0000	1,0000
Vies($\hat{b}$)		-0,0056	-0,0003	0,0000	0,0000	-0,0025	-0,0003	0,0000	0,0000
EQM($\hat{b}$)		0,0056	0,0003	0,0000	0,0000	0,0025	0,0003	0,0000	0,0000

Tabela 4.15. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 3$, 4 e $p - 1 = 3$.

$b = 2$	Dimensão: 3×3				Dimensão: 4×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0949	0,0606	0,0239	0,0046	0,0240	0,0116	0,0020	0,0003
2	0,6498	0,7328	0,7918	0,8135	0,5372	0,7018	0,7522	0,7815
3	0,2553	0,2066	0,1843	0,1819	0,4388	0,2866	0,2458	0,2182
Vies($\hat{b}$)	0,1604	0,1460	0,1604	0,1773	0,4148	0,2750	0,2438	0,2179
EQM($\hat{b}$)	0,3502	0,2672	0,2082	0,1865	0,4628	0,2982	0,2478	0,2185
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0127	0,0026	0,0003	0,0000	0,0011	0,0003	0,0000	0,0000
2	0,6233	0,6856	0,7160	0,7250	0,3675	0,5022	0,5377	0,5696
3	0,3640	0,3118	0,2837	0,2750	0,6314	0,4975	0,4623	0,4304
Vies($\hat{b}$)	0,3513	0,3092	0,2834	0,2750	0,6303	0,4972	0,4623	0,4304
EQM($\hat{b}$)	0,3767	0,3144	0,2840	0,2750	0,6325	0,4978	0,4623	0,4304
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0061	0,0008	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000
2	0,5666	0,6078	0,6268	0,6309	0,3052	0,4045	0,4234	0,4451
3	0,4273	0,3914	0,3732	0,3691	0,6945	0,5955	0,5766	0,5549
Vies($\hat{b}$)	0,4212	0,3906	0,3732	0,3691	0,6942	0,5955	0,5766	0,5549
EQM($\hat{b}$)	0,4334	0,3922	0,3732	0,3691	0,6948	0,5955	0,5766	0,5549

Tabela 4.16. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 5$, 6 e $p - 1 = 3$.

$b = 2$	Dimensão: 5×3					Dimensão: 6×3				
	Sequencial					Sequencial				
	N = 50	N = 100	N = 150	N = 200		N = 50	N = 100	N = 150	N = 200	
$\hat{b}$										
0	0,0000	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
1	0,0210	0,0064	0,0017	0,0002		0,0175	0,0083	0,0049	0,0003	
2	0,4210	0,5483	0,6245	0,6965		0,3239	0,4676	0,5824	0,6198	
3	0,5580	0,4453	0,3738	0,3033		0,6586	0,5241	0,4127	0,3799	
Vies($\hat{b}$)	0,5370	0,4389	0,3721	0,3031		0,6411	0,5158	0,4078	0,3796	
EQM($\hat{b}$)	0,5790	0,4517	0,3755	0,3035		0,6761	0,5324	0,4176	0,3802	
	BIC					BIC				
$\hat{b}$										
0	0,0000	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
1	0,0003	0,0001	0,0000	0,0000		0,0004	0,0000	0,0000	0,0000	
2	0,2081	0,2881	0,3310	0,3816		0,1152	0,1834	0,2396	0,2553	
3	0,7916	0,7118	0,6690	0,6184		0,8844	0,8166	0,7604	0,7447	
Vies($\hat{b}$)	0,7913	0,7117	0,6690	0,6184		0,8840	0,8166	0,7604	0,7447	
EQM($\hat{b}$)	0,7919	0,7119	0,6690	0,6184		0,8848	0,8166	0,7604	0,7447	
	HQIC					HQIC				
$\hat{b}$										
0	0,0000	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
1	0,0000	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
2	0,1516	0,2068	0,2305	0,2606		0,0775	0,1157	0,1462	0,1496	
3	0,8484	0,7932	0,7695	0,7394		0,9225	0,8843	0,8538	0,8504	
Vies($\hat{b}$)	0,8484	0,7932	0,7695	0,7394		0,9225	0,8843	0,8538	0,8504	
EQM($\hat{b}$)	0,8484	0,7932	0,7695	0,7394		0,9225	0,8843	0,8538	0,8504	

Tabela 4.17. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 3$, $k = 3$, 4 e $p - 1 = 3$.

$b = 3$	Dimensão: 3×3				Dimensão: 4×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0447	0,0102	0,0013	0,0000	0,0246	0,0119	0,0028	0,0000
2	0,4060	0,2001	0,0974	0,0375	0,3149	0,2191	0,1104	0,0431
3	0,5493	0,7897	0,9013	0,9625	0,6605	0,7690	0,8868	0,9569
Vies($\hat{b}$)	-0,4954	-0,2205	-0,1000	-0,0375	-0,3641	-0,2429	-0,1160	-0,0431
EQM($\hat{b}$)	0,5848	0,2409	0,1026	0,0375	0,4133	0,2667	0,1216	0,0431
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0054	0,0002	0,0000	0,0000	0,0004	0,0000	0,0000	0,0000
2	0,3074	0,0960	0,0372	0,0115	0,1370	0,0475	0,0125	0,0032
3	0,6872	0,9038	0,9628	0,9885	0,8626	0,9525	0,9875	0,9968
Vies($\hat{b}$)	-0,3182	-0,0964	-0,0372	-0,0115	-0,1378	-0,0475	-0,0125	-0,0032
EQM($\hat{b}$)	0,3290	0,0968	0,0372	0,0115	0,1386	0,0475	0,0125	0,0032
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0026	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,2489	0,0562	0,0173	0,0043	0,0944	0,0231	0,0047	0,0009
3	0,7485	0,9438	0,9827	0,9957	0,9056	0,9769	0,9953	0,9991
Vies($\hat{b}$)	-0,2541	-0,0562	-0,0173	-0,0043	-0,0944	-0,0231	-0,0047	-0,0009
EQM($\hat{b}$)	0,2593	0,0562	0,0173	0,0043	0,0944	0,0231	0,0047	0,0009

Tabela 4.18. Proporções de valores assumidos por $\hat{b}$ em 10,000 experimentos quando $b = 3$, $k = 5$, 6 e $p - 1 = 3$.

b = 3		Dimensão: 5 × 3				Dimensão: 6 × 3			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0246	0,0095	0,0027	0,0004	0,0153	0,0124	0,0044	0,0017
2		0,2876	0,2046	0,1264	0,0726	0,2508	0,2041	0,1446	0,0803
3		0,6878	0,7859	0,8709	0,9270	0,7339	0,7835	0,8510	0,9180
Vies($\hat{b}$)		-0,3368	-0,2236	-0,1318	-0,0734	-0,2814	-0,2289	-0,1534	-0,0837
EQM($\hat{b}$)		0,3860	0,2426	0,1372	0,0742	0,3120	0,2537	0,1622	0,0871
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0007	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
2		0,0897	0,0316	0,0087	0,0024	0,0485	0,0157	0,0055	0,0014
3		0,9096	0,9684	0,9913	0,9976	0,9513	0,9843	0,9945	0,9986
Vies($\hat{b}$)		-0,0911	-0,0316	-0,0087	-0,0024	-0,0489	-0,0157	-0,0055	-0,0014
EQM($\hat{b}$)		0,0925	0,0316	0,0087	0,0024	0,0493	0,0157	0,0055	0,0014
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0001	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
2		0,0565	0,0151	0,0045	0,0008	0,0263	0,0078	0,0019	0,0002
3		0,9434	0,9849	0,9955	0,9992	0,9736	0,9922	0,9981	0,9998
Vies($\hat{b}$)		-0,0567	-0,0151	-0,0045	-0,0008	-0,0265	-0,0078	-0,0019	-0,0002
EQM($\hat{b}$)		0,0569	0,0151	0,0045	0,0008	0,0267	0,0078	0,0019	0,0002

4.2. Resultados Numéricos para o Modelo 2

Apresentaremos a seguir os resultados da simulação feita para a estimação do posto da matriz B usando a segunda parametrização, definida na seção 2.2.3.

Assim como na seção anterior, para cada valor de $k \in \{3, 4, 5, 6\}$ e $p \in \{2, 3, 4\}$ calculamos a proporção de vezes em que o teste seqüencial e os critérios de informação BIC e HQIC acertam o verdadeiro posto de B , o viés e o EQM do estimador $\hat{b}$.

Os casos em que a matriz B não possui valores singulares diferentes de zero, ou seja, $b = 0$, estão expostos nas tabelas numeradas de 4.19 a 4.24. Nestes casos percebemos que os vieses e os EQM's das estimativas de b são próximos ou iguais a zero quando o teste seqüencial é usado para a estimação de b . Por exemplo, se $k = 6$, $p = 2$ e $N = 200$ o viés e o EQM de $\hat{b}$ são iguais a 0.0002 como pode ser visto na tabela 4.20; quando $k = 5$, $p = 3$ e $N = 150$, o viés e o EQM de $\hat{b}$ são praticamente nulos, ver tabela 4.22.

Como já visto nos resultados da simulação feita para a estimação de b no modelo 1, na maioria dos casos, $\hat{b}$ é assintoticamente não viesado e consistente. O mesmo acontece para os resultados referentes ao modelo 2. Ou seja, quando a segunda parametrização é usada na estimação de b , na maior parte das vezes, os vieses e os EQM's das estimativas de b diminuem quando o tamanho da amostra N cresce. Por exemplo, para $N = 50, 100, 150$ e 200 , os EQM's das estimativas de b são, respectivamente, 0.8838, 0.8559, 0.8385 e 0.8204, como se pode observar na tabela 4.34, quando $k = 5$, $p = 3$, $b = 2$ e o critério de informação HQIC é usado. Outro exemplo pode ser visto na tabela 4.30, se $k = 6$, $p = 4$, $b = 1$, para $N = 50, 100, 150$ e 200 ; os vieses das estimativas de b são, respectivamente, 1.5825, 1.3515, 1.1852 e 1.0606, quando o teste usado na estimação de b é o seqüencial. Quando acrescentamos variáveis explicativas no modelo, os resultados da estimação de b , usando os três testes já mencionados, ficam piores na maioria dos casos, ou seja, os vieses de $\hat{b}$ aumentam. Por exemplo, nas tabelas 4.27 e 4.28 temos que se $k = 3, 4, 5$ e 6 , os vieses das estimativas de b são 0.4707, 0.6416, 0.7776 e 0.8676, respectivamente, quando o teste usado é o seqüencial, $b = 1$, $p = 3$ e $N = 200$. Quando, nas tabelas 4.25 e 4.26, $b = 1$, $p = 2$, $N = 50$ e o critério usado para estimação de b é o HQIC, se $k = 3, 4, 5$ e 6 , os vieses das estimativas de b são 0.8229, 0.9278, 0.9713 e 0.9922, respectivamente.

Para todos os valores de k , p e b considerados, percebemos que quando aumentamos o número de componentes p da matriz Y , os critérios de informação BIC e HQIC acertam cada vez menos o verdadeiro posto da matriz B ; o mesmo acontece para o teste seqüencial, exceto quando $b = 0$. Nas tabelas 4.31 e 4.33 onde $p = 2$ e 3 , respectivamente, e em ambos os casos, $b = 2$ e $k = 3$, observamos para $N = 150$ que, se $p = 2$ o critério de informação HQIC acerta em 99.91% das 10.000 réplicas o posto da matriz B , e se $p = 3$ o mesmo critério de informação acerta em 58.22%. Outro exemplo pode ser visto através das tabelas 4.32, 4.34 e 4.36 onde $p = 2, 3$ e 4 , respectivamente, e nesses casos, $b = 2$ e $k = 6$, observamos para $N = 100$ que, se $p = 2$, o critério de informação BIC acerta em 99.62% das vezes o posto da matriz B ; se $p = 3$ o acerto é de 14.26%, e se $p = 4$ o acerto é de 3.33%.

Nas tabelas 4.31, 4.32, 4.37, 4.38, 4.41 e 4.42 estão expostos os resultados de simulação feita para a estimação de b quando a matriz B , de dimensão $k \times p$, tem posto completo, ou seja, $b = p$.

Nos resultados de simulação, observamos que os desempenhos do teste seqüencial e dos critérios de informação BIC e HQIC foram análogo ao visto na seção anterior. Se a matriz B tem posto completo, o critério de informação HQIC tem desempenho ligeiramente melhor que o BIC e o teste seqüencial. Isso pode ser visto, por exemplo, na tabela 4.31 quando B tem dimensão 3×2 , $b = 2$ e $N = 150$; o critério HQIC acerta em 99.91% das 10.000 réplicas o posto da matriz B , enquanto o critério BIC acerta em 99.82% e o teste seqüencial em 97.51%. Na tabela 4.42 temos que, se $k = 6$, $p = 4$, $b = 4$ e $N = 100$, o critério HQIC acerta em 97.94% das vezes, o BIC acerta em 95.81% e o teste seqüencial em 71.77%. Quando a matriz B não tem posto completo, o teste seqüencial tem um desempenho bem melhor que os critérios de informação BIC e HQIC. Isto pode ser visto, por exemplo, nas tabelas 4.29, 4.35, 4.39 e 4.40. Na tabela 4.29, se $k = 4$, $p = 4$, $b = 1$ e $N = 200$, o teste seqüencial acerta em 41.79% das 10.000 réplicas o posto da matriz B , enquanto o critério BIC acerta em 3.18% e o HQIC em 0.81%. Quando $k = 4$, $p = 4$, $b = 2$ e $N = 200$; o teste seqüencial acerta em 67.17% das vezes o posto da matriz B , enquanto o critério BIC acerta em 26.55% e o HQIC em 15.37%, ver tabela 4.35. Outro exemplo pode ser visto na tabela 4.24, se $k = 6$, $p = 4$, $b = 0$ e $N = 100$, o teste seqüencial acerta em 99.98% das vezes, o critério BIC acerta em 0.67% e o HQIC em 0.03%.

Tabela 4.19. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 3$, 4 e $p = 2$.

$b = 0$		Dimensão: 3×2				Dimensão: 4×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,9893	0,9981	0,9994	0,9996	0,9840	0,9986	0,9992	0,9998
1		0,0077	0,0015	0,0003	0,0004	0,0120	0,0012	0,0006	0,0002
2		0,0030	0,0004	0,0003	0,0000	0,0040	0,0002	0,0002	0,0000
Vies($\hat{b}$)		0,0137	0,0023	0,0009	0,0004	0,0200	0,0016	0,0010	0,0002
EQM($\hat{b}$)		0,0197	0,0031	0,0015	0,0004	0,0280	0,0020	0,0014	0,0002
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,4407	0,5659	0,6381	0,6773	0,2358	0,3733	0,4471	0,5064
1		0,4220	0,3640	0,3160	0,2870	0,4516	0,4474	0,4299	0,3925
2		0,1373	0,0701	0,0459	0,0357	0,3126	0,1793	0,1230	0,1011
Vies($\hat{b}$)		0,6966	0,5042	0,4078	0,3584	1,0768	0,8060	0,6759	0,5947
EQM($\hat{b}$)		0,9712	0,6444	0,4996	0,4298	1,7020	1,1646	0,9219	0,7969
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,2661	0,3365	0,3828	0,4146	0,1083	0,1702	0,2050	0,2260
1		0,4856	0,4894	0,4673	0,4535	0,4176	0,4731	0,4938	0,4895
2		0,2483	0,1741	0,1499	0,1319	0,4741	0,3567	0,3012	0,2845
Vies($\hat{b}$)		0,9822	0,8376	0,7671	0,7173	1,3658	1,1865	1,0962	1,0585
EQM($\hat{b}$)		1,4788	1,1858	1,0669	0,9811	2,3140	1,8999	1,6986	1,6275

Tabela 4.20. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 5, 6$ e $p = 2$.

$b = 0$		Dimensão: 5×2				Dimensão: 6×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,9704	0,9970	0,9990	0,9998	0,9460	0,9963	0,9997	0,9998
1		0,0184	0,0025	0,0009	0,0001	0,0330	0,0028	0,0003	0,0002
2		0,0112	0,0005	0,0001	0,0001	0,0210	0,0009	0,0000	0,0000
Vies($\hat{b}$)		0,0408	0,0035	0,0011	0,0003	0,0750	0,0046	0,0003	0,0002
EQM($\hat{b}$)		0,0632	0,0045	0,0013	0,0005	0,1170	0,0064	0,0003	0,0002
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,1137	0,2208	0,2995	0,3543	0,0377	0,1136	0,1770	0,2196
1		0,3688	0,4536	0,4491	0,4444	0,2458	0,3799	0,4221	0,4337
2		0,5175	0,3256	0,2514	0,2013	0,7165	0,5065	0,4009	0,3467
Vies($\hat{b}$)		1,4038	1,1048	0,9519	0,8470	1,6788	1,3929	1,2239	1,1271
EQM($\hat{b}$)		2,4388	1,7560	1,4547	1,2496	3,1118	2,4059	2,0257	1,8205
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0390	0,0756	0,1040	0,1201	0,0093	0,0280	0,0407	0,0538
1		0,2792	0,3743	0,4068	0,4347	0,1460	0,2501	0,3040	0,3211
2		0,6818	0,5501	0,4892	0,4452	0,8447	0,7219	0,6553	0,6251
Vies($\hat{b}$)		1,6428	1,4745	1,3852	1,3251	1,8354	1,6939	1,6146	1,5713
EQM($\hat{b}$)		3,0064	2,5747	2,3636	2,2155	3,5248	3,1377	2,9252	2,8215

Tabela 4.21. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 3$, 4 e $p = 3$.

$b = 0$	Dimensão: 3×3				Dimensão: 4×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,9967	0,9997	0,9999	0,9999	0,9950	0,9999	1,0000	0,9998
1	0,0027	0,0002	0,0000	0,0001	0,0031	0,0000	0,0000	0,0002
2	0,0005	0,0001	0,0000	0,0000	0,0013	0,0001	0,0000	0,0000
3	0,0001	0,0000	0,0001	0,0000	0,0006	0,0000	0,0000	0,0000
Vies($\hat{b}$)	0,0040	0,0004	0,0003	0,0001	0,0075	0,0002	0,0000	0,0002
EQM($\hat{b}$)	0,0056	0,0006	0,0009	0,0001	0,0137	0,0004	0,0000	0,0002
BIC					BIC			
$\hat{b}$								
0	0,2484	0,3846	0,4547	0,4989	0,1083	0,1947	0,2616	0,3146
1	0,4999	0,4699	0,4368	0,4099	0,3915	0,4626	0,4875	0,4727
2	0,2190	0,1345	0,1017	0,0862	0,3997	0,3043	0,2245	0,1935
3	0,0327	0,0110	0,0068	0,0050	0,1005	0,0384	0,0264	0,0192
Vies($\hat{b}$)	1,0360	0,7719	0,6606	0,5973	1,4924	1,1864	1,0157	0,9173
EQM($\hat{b}$)	1,6702	1,1069	0,9048	0,7997	2,8948	2,0254	1,6231	1,4195
HQIC					HQIC			
$\hat{b}$								
0	0,1112	0,1701	0,2014	0,2237	0,0339	0,0587	0,0763	0,0879
1	0,4623	0,5072	0,5093	0,5078	0,2736	0,3583	0,4019	0,4140
2	0,3501	0,2792	0,2538	0,2378	0,5050	0,4668	0,4268	0,4150
3	0,0764	0,0435	0,0355	0,0307	0,1875	0,1162	0,0950	0,0831
Vies($\hat{b}$)	1,3917	1,1961	1,1234	1,0755	1,8461	1,6405	1,5405	1,4933
EQM($\hat{b}$)	2,5503	2,0155	1,8440	1,7353	3,9811	3,2713	2,9641	2,8219

Tabela 4.22. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 5$, 6 e $p = 3$.

$b = 0$	Dimensão: 5×3				Dimensão: 6×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,9894	0,9999	1,0000	0,9999	0,9811	0,9996	1,0000	1,0000
1	0,0072	0,0000	0,0000	0,0001	0,0134	0,0003	0,0000	0,0000
2	0,0025	0,0001	0,0000	0,0000	0,0044	0,0001	0,0000	0,0000
3	0,0009	0,0000	0,0000	0,0000	0,0011	0,0000	0,0000	0,0000
Vies($\hat{b}$)	0,0149	0,0002	0,0000	0,0001	0,0255	0,0005	0,0000	0,0000
EQM($\hat{b}$)	0,0253	0,0004	0,0000	0,0001	0,0409	0,0007	0,0000	0,0000
BIC								
$\hat{b}$								
0	0,0360	0,0931	0,1402	0,1674	0,0077	0,0345	0,0573	0,0827
1	0,2241	0,3642	0,4049	0,4360	0,1074	0,2255	0,2962	0,3265
2	0,5351	0,4509	0,3949	0,3539	0,5304	0,5556	0,5271	0,4943
3	0,2048	0,0918	0,0600	0,0427	0,3545	0,1844	0,1194	0,0965
Vies($\hat{b}$)	1,9087	1,5414	1,3747	1,2719	2,2317	1,8899	1,7086	1,6046
EQM($\hat{b}$)	4,2077	2,9940	2,5245	2,2359	5,4195	4,1075	3,4792	3,1722
HQIC								
$\hat{b}$								
0	0,0067	0,0185	0,0248	0,0298	0,0006	0,0038	0,0063	0,0101
1	0,1108	0,2007	0,2367	0,2600	0,0418	0,0849	0,1150	0,1339
2	0,5276	0,5642	0,5499	0,5511	0,4267	0,5366	0,5685	0,5733
3	0,3549	0,2166	0,1886	0,1591	0,5309	0,3747	0,3102	0,2827
Vies($\hat{b}$)	2,2307	1,9789	1,9023	1,8395	2,4879	2,2822	2,1826	2,1286
EQM($\hat{b}$)	5,4153	4,4069	4,1337	3,8963	6,5267	5,6036	5,1808	4,9714

Tabela 4.23. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 4, 5$ e $p = 4$.

$b = 0$	Dimensão: 4×4				Dimensão: 5×4			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,9962	0,9996	1,0000	1,0000	0,9947	0,9999	1,0000	1,0000
1	0,0026	0,0003	0,0000	0,0000	0,0038	0,0001	0,0000	0,0000
2	0,0007	0,0001	0,0000	0,0000	0,0011	0,0000	0,0000	0,0000
3	0,0002	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
4	0,0003	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
Vies($\hat{b}$)	0,0058	0,0005	0,0000	0,0000	0,0074	0,0001	0,0000	0,0000
EQM($\hat{b}$)	0,0120	0,0007	0,0000	0,0000	0,0132	0,0001	0,0000	0,0000
BIC					BIC			
$\hat{b}$								
0	0,0384	0,0923	0,1346	0,1650	0,0116	0,0282	0,0477	0,0684
1	0,2815	0,3974	0,4517	0,4700	0,1171	0,2357	0,3054	0,3384
2	0,4712	0,4060	0,3437	0,3107	0,4717	0,5157	0,4976	0,4790
3	0,1812	0,0954	0,0645	0,0516	0,3237	0,1916	0,1332	0,1047
4	0,0277	0,0089	0,0055	0,0027	0,0759	0,0288	0,0161	0,0095
Vies($\hat{b}$)	1,8783	1,5312	1,3546	1,2570	2,3352	1,9571	1,7646	1,6485
EQM($\hat{b}$)	4,2403	3,0224	2,4950	2,2204	6,1316	4,4837	3,7522	3,3487
HQIC					HQIC			
$\hat{b}$								
0	0,0070	0,0186	0,0221	0,0320	0,0012	0,0027	0,0046	0,0069
1	0,1427	0,2159	0,2611	0,2756	0,0386	0,0827	0,1106	0,1226
2	0,4837	0,5064	0,5029	0,4910	0,3539	0,4608	0,4942	0,5231
3	0,3013	0,2238	0,1868	0,1775	0,4556	0,3647	0,3192	0,2905
4	0,0653	0,0353	0,0271	0,0239	0,1507	0,0891	0,0714	0,0569
Vies($\hat{b}$)	2,2752	2,0413	1,9357	1,8857	2,7160	2,4548	2,3422	2,2679
EQM($\hat{b}$)	5,8340	4,8205	4,3875	4,2195	7,9658	6,6338	6,1026	5,7399

Tabela 4.24. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 0$, $k = 6$ e $p = 4$.

b = 0	Dimensão: 6×4			
	Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,9908	0,9998	1,0000	1,0000
1	0,0060	0,0001	0,0000	0,0000
2	0,0024	0,0001	0,0000	0,0000
3	0,0003	0,0000	0,0000	0,0000
4	0,0005	0,0000	0,0000	0,0000
Vies($\hat{b}$)	0,0137	0,0003	0,0000	0,0000
EQM($\hat{b}$)	0,0263	0,0005	0,0000	0,0000
	BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0010	0,0067	0,0153	0,0236
1	0,0407	0,1089	0,1653	0,2033
2	0,3332	0,4941	0,5292	0,5409
3	0,4671	0,3291	0,2524	0,2075
4	0,1580	0,0612	0,0378	0,0247
Vies($\hat{b}$)	2,7404	2,3292	2,1321	2,0064
EQM($\hat{b}$)	8,1054	6,0264	5,1585	4,6296
	HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0003	0,0005	0,0013
1	0,0076	0,0236	0,0381	0,0468
2	0,1820	0,3049	0,3693	0,4056
3	0,5261	0,5035	0,4609	0,4356
4	0,2843	0,1677	0,1312	0,1107
Vies($\hat{b}$)	3,0871	2,8147	2,6842	2,6076
EQM($\hat{b}$)	10,0193	8,4579	7,7626	7,3608

Tabela 4.25. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 3$, 4 e $p = 2$.

$b = 1$		Dimensão: 3×2				Dimensão: 4×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0026	0,0000	0,0000	0,0000	0,0052	0,0000	0,0000	0,0000
1		0,4418	0,5220	0,5641	0,5948	0,3506	0,3953	0,4508	0,4780
2		0,5556	0,4780	0,4359	0,4052	0,6442	0,6047	0,5492	0,5220
Vies($\hat{b}$)		0,5530	0,4780	0,4359	0,4052	0,6390	0,6047	0,5492	0,5220
EQM($\hat{b}$)		0,5582	0,4780	0,4359	0,4052	0,6494	0,6047	0,5492	0,5220
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,2282	0,3046	0,3363	0,3560	0,1063	0,1479	0,1763	0,1935
2		0,7718	0,6954	0,6637	0,6440	0,8937	0,8521	0,8237	0,8065
Vies($\hat{b}$)		0,7718	0,6954	0,6637	0,6440	0,8937	0,8521	0,8237	0,8065
EQM($\hat{b}$)		0,7718	0,6954	0,6637	0,6440	0,8937	0,8521	0,8237	0,8065
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,1771	0,2337	0,2449	0,2555	0,0722	0,0955	0,1137	0,1221
2		0,8229	0,7663	0,7551	0,7445	0,9278	0,9045	0,8863	0,8779
Vies($\hat{b}$)		0,8229	0,7663	0,7551	0,7445	0,9278	0,9045	0,8863	0,8779
EQM($\hat{b}$)		0,8229	0,7663	0,7551	0,7445	0,9278	0,9045	0,8863	0,8779

Tabela 4.26. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 5$, 6 e $p = 2$.

$b = 1$		Dimensão: 5×2				Dimensão: 6×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0093	0,0000	0,0000	0,0000	0,0117	0,0000	0,0000	0,0000
1		0,2988	0,3361	0,3662	0,3997	0,2584	0,2804	0,3120	0,3426
2		0,6919	0,6639	0,6338	0,6003	0,7299	0,7196	0,6880	0,6574
Vies($\hat{b}$)		0,6826	0,6639	0,6338	0,6003	0,7182	0,7196	0,6880	0,6574
EQM($\hat{b}$)		0,7012	0,6639	0,6338	0,6003	0,7416	0,7196	0,6880	0,6574
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0521	0,0773	0,1001	0,1090	0,0171	0,0345	0,0433	0,0538
2		0,9479	0,9227	0,8999	0,8910	0,9829	0,9655	0,9567	0,9462
Vies($\hat{b}$)		0,9479	0,9227	0,8999	0,8910	0,9829	0,9655	0,9567	0,9462
EQM($\hat{b}$)		0,9479	0,9227	0,8999	0,8910	0,9829	0,9655	0,9567	0,9462
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0287	0,0465	0,0533	0,0551	0,0078	0,0164	0,0173	0,0235
2		0,9713	0,9535	0,9467	0,9449	0,9922	0,9836	0,9827	0,9765
Vies($\hat{b}$)		0,9713	0,9535	0,9467	0,9449	0,9922	0,9836	0,9827	0,9765
EQM($\hat{b}$)		0,9713	0,9535	0,9467	0,9449	0,9922	0,9836	0,9827	0,9765

Tabela 4.27. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 3$, 4 e $p = 3$.

$b = 1$	Dimensão: 3×3				Dimensão: 4×3			
	Seqüencial				Seqüencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0002	0,0000	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000
1	0,3708	0,4928	0,5549	0,5903	0,2792	0,3558	0,4384	0,4770
2	0,4819	0,4263	0,3810	0,3487	0,4699	0,4714	0,4274	0,4044
3	0,1471	0,0809	0,0641	0,0610	0,2506	0,1728	0,1342	0,1186
Vies($\hat{b}$)	0,7759	0,5881	0,5092	0,4707	0,9708	0,8170	0,6958	0,6416
EQM($\hat{b}$)	1,0705	0,7499	0,6374	0,5927	1,4726	1,1626	0,9642	0,8788
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,1116	0,1737	0,2054	0,2127	0,0366	0,0637	0,0779	0,0965
2	0,6206	0,6563	0,6507	0,6476	0,4470	0,5317	0,5561	0,5694
3	0,2678	0,1700	0,1439	0,1397	0,5164	0,4046	0,3660	0,3341
Vies($\hat{b}$)	1,1562	0,9963	0,9385	0,9270	1,4798	1,3409	1,2881	1,2376
EQM($\hat{b}$)	1,6918	1,3363	1,2263	1,2064	2,5126	2,1501	2,0201	1,9058
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0714	0,1070	0,1157	0,1157	0,0177	0,0306	0,0333	0,0389
2	0,5990	0,6536	0,6652	0,6606	0,3831	0,4657	0,4850	0,5047
3	0,3296	0,2394	0,2191	0,2237	0,5992	0,5037	0,4817	0,4564
Vies($\hat{b}$)	1,2582	1,1324	1,1034	1,1080	1,5815	1,4731	1,4484	1,4175
EQM($\hat{b}$)	1,9174	1,6112	1,5416	1,5554	2,7799	2,4805	2,4118	2,3303

Tabela 4.28. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 5$, 6 e $p = 3$.

$b = 1$	Dimensão: 5×3				Dimensão: 6×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0015	0,0000	0,0000	0,0000	0,0032	0,0000	0,0000	0,0000
1	0,2297	0,2985	0,3444	0,3946	0,2038	0,2485	0,3049	0,3527
2	0,4488	0,4673	0,4513	0,4332	0,4051	0,4507	0,4549	0,4270
3	0,3200	0,2342	0,2043	0,1722	0,3879	0,3008	0,2402	0,2203
Vies($\hat{b}$)	1,0873	0,9357	0,8599	0,7776	1,1777	1,0523	0,9353	0,8676
EQM($\hat{b}$)	1,7303	1,4041	1,2685	1,1220	1,9599	1,6539	1,4157	1,3082
BIC					BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0116	0,0230	0,0291	0,0394	0,0028	0,0072	0,0117	0,0157
2	0,3080	0,3917	0,4270	0,4564	0,1933	0,2725	0,3262	0,3386
3	0,6804	0,5853	0,5439	0,5042	0,8039	0,7203	0,6621	0,6457
Vies($\hat{b}$)	1,6688	1,5623	1,5148	1,4648	1,8011	1,7131	1,6504	1,6300
EQM($\hat{b}$)	3,0296	2,7329	2,6026	2,4732	3,4089	3,1537	2,9746	2,9214
HQIC					HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0045	0,0065	0,0095	0,0123	0,0004	0,0014	0,0028	0,0028
2	0,2319	0,3011	0,3207	0,3407	0,1299	0,1807	0,2159	0,2209
3	0,7636	0,6924	0,6698	0,6470	0,8697	0,8179	0,7813	0,7763
Vies($\hat{b}$)	1,7591	1,6859	1,6603	1,6347	1,8693	1,8165	1,7785	1,7735
EQM($\hat{b}$)	3,2863	3,0707	2,9999	2,9287	3,6087	3,4523	3,3411	3,3261

Tabela 4.29. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 4$, 5 e $p = 4$.

b = 1	Dimensão: 4×4				Dimensão: 5×4			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0004	0,0000	0,0000	0,0000	0,0018	0,0000	0,0000	0,0000
1	0,2148	0,2983	0,3864	0,4179	0,1834	0,2572	0,3061	0,3619
2	0,3940	0,4330	0,4060	0,4021	0,3341	0,3892	0,3930	0,3887
3	0,3248	0,2389	0,1891	0,1628	0,3613	0,2781	0,2462	0,2116
4	0,0660	0,0298	0,0185	0,0172	0,1194	0,0755	0,0547	0,0378
Vies($\hat{b}$)	1,2412	1,0002	0,8397	0,7793	1,4131	1,1719	1,0495	0,9253
EQM($\hat{b}$)	2,2876	1,6568	1,3289	1,2081	2,8557	2,1811	1,8701	1,5753
BIC					BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0090	0,0218	0,0290	0,0318	0,0018	0,0056	0,0085	0,0121
2	0,2104	0,3081	0,3522	0,3794	0,0962	0,1712	0,2015	0,2319
3	0,6162	0,5754	0,5448	0,5304	0,5283	0,5412	0,5482	0,5512
4	0,1644	0,0947	0,0740	0,0584	0,3737	0,2820	0,2418	0,2048
Vies($\hat{b}$)	1,9360	1,7430	1,6638	1,6154	2,2739	2,0996	2,0233	1,9487
EQM($\hat{b}$)	4,1548	3,4620	3,1974	3,0266	5,5727	4,8740	4,5705	4,2799
HQIC					HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0035	0,0077	0,0087	0,0081	0,0002	0,0011	0,0014	0,0018
2	0,1453	0,2173	0,2461	0,2604	0,0534	0,0920	0,1108	0,1205
3	0,6326	0,6263	0,6136	0,6250	0,4864	0,5192	0,5422	0,5656
4	0,2186	0,1487	0,1316	0,1065	0,4600	0,3877	0,3456	0,3121
Vies($\hat{b}$)	2,0663	1,9160	1,8681	1,8299	2,4062	2,2935	2,2320	2,1880
EQM($\hat{b}$)	4,6431	4,0608	3,8849	3,7189	6,1390	5,6581	5,3900	5,1918

Tabela 4.30. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 1$, $k = 6$ e $p = 4$.

b = 1	Dimensão: 6×4			
	Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0024	0,0000	0,0000	0,0000
1	0,1651	0,2114	0,2649	0,3265
2	0,2791	0,3513	0,3726	0,3585
3	0,3544	0,3117	0,2749	0,2429
4	0,1990	0,1256	0,0876	0,0721
Vies($\hat{b}$)	1,5825	1,3515	1,1852	1,0606
EQM($\hat{b}$)	3,4901	2,7285	2,2606	1,9790
	BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0006	0,0012	0,0027	0,0036
2	0,0397	0,0782	0,1113	0,1306
3	0,3875	0,4583	0,4856	0,4956
4	0,5722	0,4623	0,4004	0,3702
Vies($\hat{b}$)	2,5313	2,3817	2,2837	2,2324
EQM($\hat{b}$)	6,7395	6,0721	5,6573	5,4448
	HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0001	0,0001	0,0002	0,0003
2	0,0160	0,0313	0,0420	0,0565
3	0,3145	0,3893	0,4269	0,4367
4	0,6694	0,5793	0,5309	0,5065
Vies($\hat{b}$)	2,6532	2,5478	2,4885	2,4494
EQM($\hat{b}$)	7,2986	6,8022	6,5277	6,3618

Tabela 4.31. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 3$, 4 e $p = 2$.

$b = 2$		Dimensão: 3×2				Dimensão: 4×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0515	0,0000	0,0000	0,0000	0,0658	0,0005	0,0000	0,0000
1		0,2264	0,0830	0,0249	0,0041	0,2431	0,1524	0,0515	0,0246
2		0,7221	0,9170	0,9751	0,9959	0,6911	0,8471	0,9485	0,9754
Vies($\hat{b}$)		-0,3294	-0,0830	-0,0249	-0,0041	-0,3747	-0,1534	-0,0515	-0,0246
EQM($\hat{b}$)		0,4324	0,0830	0,0249	0,0041	0,5063	0,1544	0,0515	0,0246
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0665	0,0105	0,0018	0,0001	0,0391	0,0137	0,0014	0,0002
2		0,9333	0,9895	0,9982	0,9999	0,9609	0,9863	0,9986	0,9998
Vies($\hat{b}$)		-0,0669	-0,0105	-0,0018	-0,0001	-0,0391	-0,0137	-0,0014	-0,0002
EQM($\hat{b}$)		0,0673	0,0105	0,0018	0,0001	0,0391	0,0137	0,0014	0,0002
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0432	0,0063	0,0009	0,0000	0,0203	0,0067	0,0007	0,0001
2		0,9568	0,9937	0,9991	1,0000	0,9797	0,9933	0,9993	0,9999
Vies($\hat{b}$)		-0,0432	-0,0063	-0,0009	0,0000	-0,0203	-0,0067	-0,0007	-0,0001
EQM($\hat{b}$)		0,0432	0,0063	0,0009	0,0000	0,0203	0,0067	0,0007	0,0001

Tabela 4.32. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 5$, 6 e $p = 2$.

$b = 2$		Dimensão: 5×2				Dimensão: 6×2			
		Seqüencial				Seqüencial			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0757	0,0004	0,0000	0,0000	0,1010	0,0040	0,0000	0,0000
1		0,2398	0,1523	0,0737	0,0250	0,2691	0,2070	0,1046	0,0429
2		0,6845	0,8473	0,9263	0,9750	0,6299	0,7890	0,8954	0,9571
Vies($\hat{b}$)		-0,3912	-0,1531	-0,0737	-0,0250	-0,4711	-0,2150	-0,1046	-0,0429
EQM($\hat{b}$)		0,5426	0,1539	0,0737	0,0250	0,6731	0,2230	0,1046	0,0429
		BIC				BIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0193	0,0031	0,0005	0,0002	0,0124	0,0038	0,0008	0,0001
2		0,9807	0,9969	0,9995	0,9998	0,9876	0,9962	0,9992	0,9999
Vies($\hat{b}$)		-0,0193	-0,0031	-0,0005	-0,0002	-0,0124	-0,0038	-0,0008	-0,0001
EQM($\hat{b}$)		0,0193	0,0031	0,0005	0,0002	0,0124	0,0038	0,0008	0,0001
		HQIC				HQIC			
$\hat{b}$		N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
0		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1		0,0089	0,0009	0,0003	0,0001	0,0059	0,0006	0,0001	0,0000
2		0,9911	0,9991	0,9997	0,9999	0,9941	0,9994	0,9999	1,0000
Vies($\hat{b}$)		-0,0089	-0,0009	-0,0003	-0,0001	-0,0059	-0,0006	-0,0001	0,0000
EQM($\hat{b}$)		0,0089	0,0009	0,0003	0,0001	0,0059	0,0006	0,0001	0,0000

Tabela 4.33. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 3$, 4 e $p = 3$.

$b = 2$	Dimensão: 3×3				Dimensão: 4×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0007	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,1355	0,0595	0,0260	0,0066	0,0561	0,0127	0,0027	0,0005
2	0,5690	0,6911	0,7413	0,7628	0,4849	0,6336	0,6808	0,7207
3	0,2948	0,2494	0,2327	0,2306	0,4590	0,3537	0,3165	0,2788
Vies($\hat{b}$)	0,1579	0,1899	0,2067	0,2240	0,4029	0,3410	0,3138	0,2783
EQM($\hat{b}$)	0,4331	0,3089	0,2587	0,2372	0,5151	0,3664	0,3192	0,2783
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0135	0,0021	0,0003	0,0000	0,0004	0,0000	0,0000	0,0000
2	0,5617	0,6396	0,6715	0,6752	0,2894	0,3999	0,4460	0,4704
3	0,4248	0,3583	0,3282	0,3248	0,7102	0,6001	0,5540	0,5296
Vies($\hat{b}$)	0,4113	0,3562	0,3279	0,3248	0,7098	0,6001	0,5540	0,5296
EQM($\hat{b}$)	0,4383	0,3604	0,3285	0,3248	0,7106	0,6001	0,5540	0,5296
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0059	0,0006	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
2	0,5065	0,5610	0,5822	0,5775	0,2298	0,3060	0,3384	0,3555
3	0,4876	0,4384	0,4178	0,4225	0,7700	0,6940	0,6616	0,6445
Vies($\hat{b}$)	0,4817	0,4378	0,4178	0,4225	0,7698	0,6940	0,6616	0,6445
EQM($\hat{b}$)	0,4935	0,4390	0,4178	0,4225	0,7702	0,6940	0,6616	0,6445

Tabela 4.34. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 5$, 6 e $p = 3$.

$b = 2$	Dimensão: 5×3				Dimensão: 6×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0003	0,0000	0,0000	0,0000	0,0070	0,0000	0,0000	0,0000
1	0,0725	0,0123	0,0026	0,0005	0,1800	0,1067	0,0599	0,0335
2	0,4072	0,5034	0,5673	0,6267	0,3484	0,4272	0,4795	0,5298
3	0,5200	0,4843	0,4301	0,3728	0,4646	0,4661	0,4606	0,4367
Vies($\hat{b}$)	0,4469	0,4720	0,4275	0,3723	0,2706	0,3594	0,4007	0,4032
EQM($\hat{b}$)	0,5937	0,4966	0,4327	0,3733	0,6726	0,5728	0,5205	0,4702
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0002	0,0000	0,0000	0,0000	0,0008	0,0002	0,0001	0,0000
2	0,1636	0,2126	0,2424	0,2897	0,1162	0,1426	0,1642	0,1736
3	0,8362	0,7874	0,7576	0,7103	0,8830	0,8572	0,8357	0,8264
Vies($\hat{b}$)	0,8360	0,7874	0,7576	0,7103	0,8822	0,8570	0,8356	0,8264
EQM($\hat{b}$)	0,8364	0,7874	0,7576	0,7103	0,8838	0,8574	0,8358	0,8264
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0002	0,0000	0,0000	0,0000	0,0002	0,0000	0,0000	0,0000
2	0,1162	0,1441	0,1615	0,1796	0,0677	0,0811	0,0929	0,0974
3	0,8836	0,8559	0,8385	0,8204	0,9321	0,9189	0,9071	0,9026
Vies($\hat{b}$)	0,8834	0,8559	0,8385	0,8204	0,9319	0,9189	0,9071	0,9026
EQM($\hat{b}$)	0,8838	0,8559	0,8385	0,8204	0,9323	0,9189	0,9071	0,9026

Tabela 4.35. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 4$, 5 e $p = 4$.

$b = 2$	Dimensão: 4×4				Dimensão: 5×4			
	Seqüencial				Seqüencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0008	0,0000	0,0000	0,0000
1	0,0373	0,0149	0,0032	0,0009	0,1160	0,0588	0,0426	0,0236
2	0,3805	0,5462	0,6190	0,6717	0,2768	0,3623	0,4179	0,4845
3	0,4705	0,3848	0,3364	0,2948	0,4166	0,4249	0,4147	0,3844
4	0,1117	0,0541	0,0414	0,0326	0,1898	0,1540	0,1248	0,1075
Vies($\hat{b}$)	0,6566	0,4781	0,4160	0,3591	0,6786	0,6741	0,6217	0,5758
EQM($\hat{b}$)	0,9546	0,6161	0,5052	0,4261	1,2950	1,0997	0,9565	0,8380
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0002	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
2	0,1201	0,1976	0,2347	0,2655	0,0531	0,0707	0,0908	0,1044
3	0,6591	0,6766	0,6633	0,6486	0,4846	0,5366	0,5664	0,5746
4	0,2206	0,1258	0,1020	0,0859	0,4622	0,3927	0,3428	0,3210
Vies($\hat{b}$)	1,1001	0,9282	0,8673	0,8204	1,4089	1,3220	1,2520	1,2166
EQM($\hat{b}$)	1,5417	1,1798	1,0713	0,9922	2,3335	2,1074	1,9376	1,8586
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
2	0,0751	0,1175	0,1387	0,1537	0,0238	0,0349	0,0417	0,0405
3	0,6447	0,6964	0,6978	0,6966	0,4261	0,4704	0,5061	0,5144
4	0,2802	0,1861	0,1635	0,1497	0,5500	0,4947	0,4522	0,4451
Vies($\hat{b}$)	1,2051	1,0686	1,0248	0,9960	1,5260	1,4598	1,4105	1,4046
EQM($\hat{b}$)	1,7655	1,4408	1,3518	1,2954	2,6262	2,4492	2,3149	2,2948

Tabela 4.36. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 2$, $k = 6$ e $p = 4$.

b = 2	Dimensão: 6×4			
	Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0017	0,0000	0,0000	0,0000
1	0,1289	0,0819	0,0514	0,0354
2	0,2606	0,3293	0,4017	0,4402
3	0,3673	0,3933	0,3893	0,3706
4	0,2415	0,1955	0,1576	0,1538
Vies($\hat{b}$)	0,7180	0,7024	0,6531	0,6428
EQM($\hat{b}$)	1,4690	1,2572	1,0711	1,0212
	BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000
2	0,0185	0,0333	0,0411	0,0455
3	0,3371	0,4145	0,4566	0,4677
4	0,6444	0,5522	0,5023	0,4868
Vies($\hat{b}$)	1,6259	1,5189	1,4612	1,4413
EQM($\hat{b}$)	2,9147	2,6233	2,4658	2,4149
	HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000
2	0,0066	0,0111	0,0139	0,0127
3	0,2597	0,3261	0,3553	0,3588
4	0,7337	0,6628	0,6308	0,6285
Vies($\hat{b}$)	1,7271	1,6517	1,6169	1,6158
EQM($\hat{b}$)	3,1945	2,9773	2,8785	2,8728

Tabela 4.37. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 3$, $k = 3$, 4 e $p = 3$.

$b = 3$	Dimensão: 3×3				Dimensão: 4×3			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0047	0,0000	0,0000	0,0000	0,0063	0,0000	0,0000	0,0000
1	0,0582	0,0034	0,0001	0,0000	0,0843	0,0119	0,0013	0,0001
2	0,2293	0,0743	0,0183	0,0018	0,2583	0,1379	0,0495	0,0175
3	0,7078	0,9223	0,9816	0,9982	0,6511	0,8502	0,9492	0,9824
Vies($\hat{b}$)	-0,3598	-0,0811	-0,0185	-0,0018	-0,4458	-0,1617	-0,0521	-0,0177
EQM($\hat{b}$)	0,5044	0,0879	0,0187	0,0018	0,6522	0,1855	0,0547	0,0179
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0026	0,0001	0,0000	0,0000	0,0003	0,0000	0,0000	0,0000
2	0,1770	0,0343	0,0058	0,0003	0,1102	0,0308	0,0056	0,0011
3	0,8204	0,9656	0,9942	0,9997	0,8895	0,9692	0,9944	0,9989
Vies($\hat{b}$)	-0,1822	-0,0345	-0,0058	-0,0003	-0,1108	-0,0308	-0,0056	-0,0011
EQM($\hat{b}$)	0,1874	0,0347	0,0058	0,0003	0,1114	0,0308	0,0056	0,0011
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0009	0,0000	0,0000	0,0000	0,0001	0,0000	0,0000	0,0000
2	0,1349	0,0176	0,0020	0,0002	0,0713	0,0168	0,0022	0,0001
3	0,8642	0,9824	0,9980	0,9998	0,9286	0,9832	0,9978	0,9999
Vies($\hat{b}$)	-0,1367	-0,0176	-0,0020	-0,0002	-0,0715	-0,0168	-0,0022	-0,0001
EQM($\hat{b}$)	0,1385	0,0176	0,0020	0,0002	0,0717	0,0168	0,0022	0,0001

Tabela 4.38. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 3$, $k = 5$, 6 e $p = 3$.

$b = 3$	Dimensão: 5×3					Dimensão: 6×3				
	Sequencial					Sequencial				
	N = 50	N = 100	N = 150	N = 200		N = 50	N = 100	N = 150	N = 200	
$\hat{b}$										
0	0,0159	0,0000	0,0000	0,0000		0,0199	0,0000	0,0000	0,0000	
1	0,1226	0,0260	0,0032	0,0003		0,1507	0,0424	0,0072	0,0008	
2	0,2821	0,1802	0,0870	0,0341		0,2854	0,2155	0,1317	0,0612	
3	0,5794	0,7938	0,9098	0,9656		0,5440	0,7421	0,8611	0,9380	
Vies($\hat{b}$)	-0,5750	-0,2322	-0,0934	-0,0347		-0,6465	-0,3003	-0,1461	-0,0628	
EQM($\hat{b}$)	0,9156	0,2842	0,0998	0,0353		1,0673	0,3851	0,1605	0,0644	
	BIC					BIC				
$\hat{b}$										
0	0,0000	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
1	0,0004	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
2	0,0865	0,0163	0,0026	0,0008		0,0524	0,0136	0,0036	0,0003	
3	0,9131	0,9837	0,9974	0,9992		0,9476	0,9864	0,9964	0,9997	
Vies($\hat{b}$)	-0,0873	-0,0163	-0,0026	-0,0008		-0,0524	-0,0136	-0,0036	-0,0003	
EQM($\hat{b}$)	0,0881	0,0163	0,0026	0,0008		0,0524	0,0136	0,0036	0,0003	
	HQIC					HQIC				
$\hat{b}$										
0	0,0000	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
1	0,0001	0,0000	0,0000	0,0000		0,0000	0,0000	0,0000	0,0000	
2	0,0537	0,0073	0,0009	0,0002		0,0268	0,0052	0,0012	0,0000	
3	0,9462	0,9927	0,9991	0,9998		0,9732	0,9948	0,9988	1,0000	
Vies($\hat{b}$)	-0,0539	-0,0073	-0,0009	-0,0002		-0,0268	-0,0052	-0,0012	0,0000	
EQM($\hat{b}$)	0,0541	0,0073	0,0009	0,0002		0,0268	0,0052	0,0012	0,0000	

Tabela 4.39. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 3$, $k = 4, 5$ e $p = 4$.

$b = 3$	Dimensão: 4×4				Dimensão: 5×4			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0005	0,0000	0,0000	0,0000	0,0017	0,0000	0,0000	0,0000
1	0,0451	0,0059	0,0011	0,0000	0,0715	0,0154	0,0026	0,0002
2	0,1834	0,1122	0,0547	0,0268	0,1996	0,1302	0,0743	0,0402
3	0,5103	0,6560	0,7177	0,7570	0,4338	0,5335	0,6172	0,6689
4	0,2607	0,2259	0,2265	0,2162	0,2934	0,3209	0,3059	0,2907
Vies($\hat{b}$)	-0,0144	0,1019	0,1696	0,1894	-0,0543	0,1599	0,2264	0,2501
EQM($\hat{b}$)	0,6290	0,3617	0,2856	0,2430	0,7943	0,5127	0,3906	0,3317
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,0310	0,0090	0,0012	0,0008	0,0160	0,0037	0,0004	0,0000
3	0,5696	0,6500	0,6696	0,6863	0,4009	0,4136	0,4455	0,4687
4	0,3994	0,3410	0,3292	0,3129	0,5831	0,5827	0,5541	0,5313
Vies($\hat{b}$)	0,3684	0,3320	0,3280	0,3121	0,5671	0,5790	0,5537	0,5313
EQM($\hat{b}$)	0,4304	0,3500	0,3304	0,3137	0,5991	0,5864	0,5545	0,5313
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,0158	0,0037	0,0002	0,0000	0,0068	0,0010	0,0000	0,0000
3	0,5204	0,5784	0,5848	0,5907	0,3309	0,3221	0,3356	0,3465
4	0,4638	0,4179	0,4150	0,4093	0,6623	0,6769	0,6644	0,6535
Vies($\hat{b}$)	0,4480	0,4142	0,4148	0,4093	0,6555	0,6759	0,6644	0,6535
EQM($\hat{b}$)	0,4796	0,4216	0,4152	0,4093	0,6691	0,6779	0,6644	0,6535

Tabela 4.40. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 3$, $k = 6$ e $p = 4$.

b = 3	Dimensão: 6×4			
	Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0027	0,0000	0,0000	0,0000
1	0,0935	0,0239	0,0050	0,0010
2	0,1940	0,1408	0,0938	0,0570
3	0,3625	0,4654	0,5617	0,5926
4	0,3473	0,3699	0,3395	0,3494
Vies($\hat{b}$)	-0,0418	0,1813	0,2357	0,2904
EQM($\hat{b}$)	0,9396	0,6063	0,4533	0,4104
	BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0001	0,0000	0,0000	0,0000
2	0,0063	0,0017	0,0005	0,0000
3	0,2535	0,2736	0,3117	0,3156
4	0,7401	0,7247	0,6878	0,6844
Vies($\hat{b}$)	0,7336	0,7230	0,6873	0,6844
EQM($\hat{b}$)	0,7468	0,7264	0,6883	0,6844
	HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000
2	0,0015	0,0004	0,0000	0,0000
3	0,1792	0,1866	0,2094	0,2092
4	0,8193	0,8130	0,7906	0,7908
Vies($\hat{b}$)	0,8178	0,8126	0,7906	0,7908
EQM($\hat{b}$)	0,8208	0,8134	0,7906	0,7908

Tabela 4.41. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 4$, $k = 4$, 5 e $p = 4$.

$b = 4$	Dimensão: 4×4				Dimensão: 5×4			
	Sequencial				Sequencial			
	N = 50	N = 100	N = 150	N = 200	N = 50	N = 100	N = 150	N = 200
$\hat{b}$								
0	0,0018	0,0000	0,0000	0,0000	0,0046	0,0000	0,0000	0,0000
1	0,0266	0,0013	0,0000	0,0000	0,0482	0,0023	0,0002	0,0000
2	0,0949	0,0153	0,0009	0,0001	0,1248	0,0251	0,0042	0,0003
3	0,2952	0,1352	0,0387	0,0104	0,3046	0,1731	0,0909	0,0289
4	0,5815	0,8482	0,9604	0,9895	0,5178	0,7995	0,9047	0,9708
Vies($\hat{b}$)	-0,5720	-0,1697	-0,0405	-0,0106	-0,7172	-0,2302	-0,0999	-0,0295
EQM($\hat{b}$)	0,9430	0,2081	0,0423	0,0108	1,3112	0,2942	0,1095	0,0301
	BIC				BIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,0089	0,0004	0,0000	0,0000	0,0035	0,0000	0,0000	0,0000
3	0,2708	0,0825	0,0144	0,0033	0,2111	0,0431	0,0117	0,0019
4	0,7203	0,9171	0,9856	0,9967	0,7854	0,9569	0,9883	0,9981
Vies($\hat{b}$)	-0,2886	-0,0833	-0,0144	-0,0033	-0,2181	-0,0431	-0,0117	-0,0019
EQM($\hat{b}$)	0,3064	0,0841	0,0144	0,0033	0,2251	0,0431	0,0117	0,0019
	HQIC				HQIC			
$\hat{b}$								
0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,0034	0,0001	0,0000	0,0000	0,0007	0,0000	0,0000	0,0000
3	0,2206	0,0552	0,0072	0,0011	0,1536	0,0235	0,0039	0,0004
4	0,7760	0,9447	0,9928	0,9989	0,8457	0,9765	0,9961	0,9996
Vies($\hat{b}$)	-0,2274	-0,0554	-0,0072	-0,0011	-0,1550	-0,0235	-0,0039	-0,0004
EQM($\hat{b}$)	0,2342	0,0556	0,0072	0,0011	0,1564	0,0235	0,0039	0,0004

Tabela 4.42. Proporções de valores assumidos por $\hat{b}$ em 10.000 experimentos quando $b = 4$, $k = 6$ e $p = 4$.

b = 4	Dimensão: 6×4			
	Seqüencial			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0063	0,0000	0,0000	0,0000
1	0,0730	0,0069	0,0003	0,0000
2	0,1633	0,0506	0,0099	0,0021
3	0,2983	0,2248	0,1372	0,0663
4	0,4591	0,7177	0,8526	0,9316
Vies($\hat{b}$)	-0,8691	-0,3467	-0,1579	-0,0705
EQM($\hat{b}$)	1,7093	0,4893	0,1795	0,0747
	BIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000
2	0,0026	0,0001	0,0000	0,0000
3	0,1534	0,0418	0,0103	0,0023
4	0,8440	0,9581	0,9897	0,9977
Vies($\hat{b}$)	-0,1586	-0,0420	-0,0103	-0,0023
EQM($\hat{b}$)	0,1638	0,0422	0,0103	0,0023
	HQIC			
$\hat{b}$	N = 50	N = 100	N = 150	N = 200
0	0,0000	0,0000	0,0000	0,0000
1	0,0000	0,0000	0,0000	0,0000
2	0,0007	0,0000	0,0000	0,0000
3	0,0976	0,0206	0,0037	0,0005
4	0,9017	0,9794	0,9963	0,9995
Vies($\hat{b}$)	-0,0990	-0,0206	-0,0037	-0,0005
EQM($\hat{b}$)	0,1004	0,0206	0,0037	0,0005

Capítulo 5

Aplicação

Ao longo deste capítulo, apresentaremos duas aplicações onde nosso interesse é estimar o posto da matriz de parâmetros do modelo de regressão Dirichlet, utilizando dados reais. Para isso usamos a segunda parametrização definida na seção 2.2.3, devido à sua maior simplicidade, o teste sequencial e os critérios de informação BIC e HQIC descritos no capítulo 3.

Consideramos dois conjuntos de dados apresentados por Aitchison (2003), os quais denotaremos por Dados 1 e Dados 2, que estão disponíveis no apêndice 4.

5.1. Dados do pólen fossilizado em flores

Nesta seção consideramos um conjunto de dados que contém 30 observações. A variável dependente Y deste estudo é o vetor de proporções de pólen fossilizado de três tipos de flores: *Pinus*, *Abies* e *Quercus*. As variáveis explicativas binárias (dummy) X que modelam as proporções correspondem a três locais onde o pólen fossilizado das flores é colhido. Logo, temos que $k = p = 3$.

Os parâmetros do modelo de regressão são estimados através da maximização da função de log-verossimilhança utilizando o método quasi-Newton BFGS com primeiras derivadas numéricas; no esquema iterativo usamos como valor inicial para B a matriz nula. Na tabela 5.1 apresentamos as estimativas de máxima verossimilhança (EMV) desses parâmetros.

Tabela 5.1. EMV dos parâmetros da matriz B para os Dados 1.

$\hat{\beta}_{11}$	1,87141	$\hat{\beta}_{12}$	2,58843	$\hat{\beta}_{13}$	1,53982
$\hat{\beta}_{21}$	2,30180	$\hat{\beta}_{22}$	2,83115	$\hat{\beta}_{23}$	2,32410
$\hat{\beta}_{31}$	2,75677	$\hat{\beta}_{32}$	2,48318	$\hat{\beta}_{33}$	1,07470

Após a estimação da matriz B usamos a decomposição em valores singulares de $\hat{B}$ e obtemos a matriz de informação observada de Fisher para calcular a estatística de teste $L(b)$, descrita na seção 3.1.2, que converge em distribuição para uma qui-quadrado com $(3 - b)^2$ graus de liberdade; o nível de significância foi o mesmo usado na simulação, de 5%. Como já visto no capítulo 3, o valor crítico da estatística de teste $L(b)$ é $\gamma_b C_{30}$, onde γ_b representa o quantil da distribuição $\chi^2_{(3-b)(3-b)}$ e $C_{30} = \sqrt{\log 30}$. Na tabela 5.2 temos os valores da estatística de teste $L(b)$ para $b = 0, 1$ e 2 e seus respectivos valores críticos. Observamos que para todos os valores de b a estatística $L(b)$ é maior que $\gamma_b C_{30}$, então,

usando o teste seqüencial, temos $\hat{b} = 3$.

Tabela 5.2. Estatísticas de teste $L(b)$ e seus respectivos valores críticos $\gamma_b C_{30}$.

b	$L(b)$	$\gamma_b C_{30}$
0	291,9588	31,2025
1	83,8572	17,4976
2	9,5065	7,0845

Na tabela 5.3 temos, para cada critério de informação considerado na seção 3.3, os valores da função critério, em cada ponto $b = 0, 1, 2$ e 3 , lembrando que $L(3) = 0$. Percebemos que os menores valores que as funções critério BIC e HQIC assumem são 10.2036 e 7.3815, respectivamente, para $b = 3$, isto implica que $\hat{b} = 3$.

Tabela 5.3. Função critério para BIC e HQIC.

b	BIC(b, C_{30})	HQIC(b, C_{30})
0	291,9588	291,9588
1	87,2584	86,3177
2	16,3089	14,4275
3	10,2036	7,3815

Pelo teste seqüencial e pelos critérios de informação BIC e HQIC temos que $\hat{b} = 3$; isto significa que aceitamos que a matriz B tem posto completo. Baseado nisso não há como reduzir a quantidade de parâmetros da matriz B para estimação.

5.2. Dados da composição de leucócitos no sangue

O segundo conjunto de dados contém 40 observações, mas duas foram desconsideradas por problemas de arredondamento. Então, para nosso estudo, consideramos apenas 38 observações. A variável dependente Y é o vetor de proporções de três tipos de leucócitos em amostras de sangue de dez pacientes. Os três tipos de leucócitos são: leucócitos polimorfonucleares, linfócitos pequenos e mononucleares grandes. As variáveis explicativas (dummy) correspondem a quatro métodos diferentes que avaliam a composição de leucócitos nas amostras de sangue. Então agora temos que $k = 4$ e $p = 3$.

Para calcular as estimativas de máxima verossimilhança (EMV) de B utilizamos o mesmo procedimento de estimação descrito na seção anterior. Estas estimativas estão apresentadas na tabela 5.4.

Tabela 5.4. EMV dos parâmetros da matriz B para os Dados 2.

$\hat{\beta}_{11}$	2,58136	$\hat{\beta}_{12}$	1,17376	$\hat{\beta}_{13}$	1,35387
$\hat{\beta}_{21}$	1,61533	$\hat{\beta}_{22}$	1,99360	$\hat{\beta}_{23}$	0,27110
$\hat{\beta}_{31}$	1,71084	$\hat{\beta}_{32}$	1,39066	$\hat{\beta}_{33}$	0,38854
$\hat{\beta}_{41}$	2,05352	$\hat{\beta}_{42}$	1,62212	$\hat{\beta}_{43}$	0,66549

Após a estimação da matriz B calculamos a estatística de teste $L(b)$, que converge em distribuição para uma qui-quadrado com $(4 - b)(3 - b)$ graus de liberdade, e um nível de significância de 5%. Para $b = 0, 1$ e 2 apresentamos os valores da estatística de teste $L(b)$ e seus respectivos valores críticos, ver tabela 5.5. Notamos que $L(2) \leq \gamma_2 C_{38}$, então, usando o teste seqüencial, consideramos que $\hat{b} = 2$.

Tabela 5.5. Estatísticas de teste $L(b)$ e seus respectivos valores críticos $\gamma_b C_{38}$.

b	$L(b)$	$\gamma_b C_{38}$
0	237,4092	40,1019
1	52,7538	24,0153
2	0,8070	11,4272

Na tabela 5.6 temos para cada critério de informação considerado os valores das funções critério em cada ponto $b = 0, 1, 2$ e 3 . Percebemos que os menores valores que as funções critério assumem são $\text{BIC}(2, C_{38}) = 8.0821$ e $\text{HQIC}(2, C_{38}) = 5.9981$; isto implica que $\hat{b} = 2$.

Tabela 5.6. Função critério para BIC e HQIC.

b	$\text{BIC}(b, C_{38})$	$\text{HQIC}(b, C_{38})$
0	237,4092	237,4092
1	56,3913	55,3493
2	8,0821	5,9981
3	10,9128	7,7867

Usando o teste seqüencial e os critérios de informação BIC e HQIC aceitamos que $b = 2$, ou seja, que B não tem posto completo com isso não precisamos estimar os doze parâmetros da matriz B , mas apenas dez. Então temos uma nova matriz de parâmetros onde podemos escrever, por exemplo, a terceira coluna como combinação linear das duas primeiras usando dois novos parâmetros, k_1 e k_2 , ou seja,

$$B^* = \begin{pmatrix} \beta_{11}^* & \beta_{12}^* & k_1\beta_{11}^* + k_2\beta_{12}^* \\ \beta_{21}^* & \beta_{22}^* & k_1\beta_{21}^* + k_2\beta_{22}^* \\ \beta_{31}^* & \beta_{32}^* & k_1\beta_{31}^* + k_2\beta_{32}^* \\ \beta_{41}^* & \beta_{42}^* & k_1\beta_{41}^* + k_2\beta_{42}^* \end{pmatrix}.$$

Depois de estimar os dez parâmetros da matriz B^* obtemos

$$\begin{aligned}\widehat{k}_1 &= 0.742178 \\ \widehat{k}_2 &= -0.525821\end{aligned}$$

onde $\widehat{k}_1$ e $\widehat{k}_2$ são os estimadores de máxima verossimilhança de k_1 e k_2 , respectivamente. As estimativas da matriz B^* não diferem muito das estimativas de B . Por exemplo, $\widehat{\beta}_{13}^* = 1,29432$ enquanto $\widehat{\beta}_{13} = 1,35387$; $\widehat{\beta}_{43}^* = 0,67124$ e $\widehat{\beta}_{43} = 0,66549$; isso pode ser visto nas tabelas 5.4 e 5.7.

Tabela 5.7. EMV dos parâmetros da matriz B^* .

$\widehat{\beta}_{11}^*$	2,53893	$\widehat{\beta}_{12}^*$	1,12208	$\widehat{\beta}_{13}^*$	1,29432
$\widehat{\beta}_{21}^*$	1,56760	$\widehat{\beta}_{22}^*$	1,91829	$\widehat{\beta}_{23}^*$	0,15476
$\widehat{\beta}_{31}^*$	1,77440	$\widehat{\beta}_{32}^*$	1,49148	$\widehat{\beta}_{33}^*$	0,53267
$\widehat{\beta}_{41}^*$	2,05667	$\widehat{\beta}_{42}^*$	1,62638	$\widehat{\beta}_{43}^*$	0,67124

Denotaremos os modelos em que as matrizes de parâmetros B e B^* são usadas por Mod e Mod^* , respectivamente. Os valores dos coeficientes de correlação amostral ρ , com $-1 \leq \rho \leq 1$, entre as colunas de $\widehat{\eta} = X\widehat{B}$ e as de $g(Y)$, bem como entre as colunas de $\widehat{\eta}^* = X\widehat{B}^*$ e as de $g(Y)$ podem ser vistos na tabela 5.8, onde X é a matriz de variáveis explicativas e $g(\cdot)$ é a inversa da função de ligação usada no modelo 2. Observamos que os valores de $\rho(\widehat{\eta}, g(Y))$ e $\rho(\widehat{\eta}^*, g(Y))$ são parecidos, o que indica que o poder explicativo dos modelos Mod e Mod^* são semelhantes. Porém, o modelo Mod^* apresenta a vantagem de ser mais parcimonioso.

Tabela 5.8. Coeficientes de correlação entre $\widehat{\eta}$ e $g(Y)$ e entre $\widehat{\eta}^*$ e $g(Y)$.

	0,48700	-0,71965	0,42718
$\rho(\widehat{\eta}, g(Y))$	-0,50937	0,72380	-0,36196
	0,48780	-0,71882	0,43394
	0,51108	-0,75104	0,42908
$\rho(\widehat{\eta}^*, g(Y))$	-0,52414	0,75235	-0,40720
	0,53001	-0,77229	0,43293

Uma medida muito utilizada na seleção de modelos é o critério de informação $\mathcal{BIC}$ (Bayesian Information Criteria), proposto por Schwarz (1978), dado pela expressão

$$\mathcal{BIC} = -2 \log(\hat{\mathcal{L}}) + npar \log(N),$$

onde $\hat{\mathcal{L}}$ é a verossimilhança estimada e $npar$ é o número de parâmetros do modelo. A escolha do melhor modelo para a variável resposta é dada pelo menor $\mathcal{BIC}$.

Para o modelo Mod obtemos $\mathcal{BIC} = -108.404$ com $\log(\hat{\mathcal{L}}) = 76.0273$ enquanto que para o modelo Mod^* obtemos $\mathcal{BIC} = -115.244$, onde $\log(\hat{\mathcal{L}}) = 75.8101$. Portanto, o modelo que melhor se ajusta a variável resposta de acordo com o critério $\mathcal{BIC}$ é Mod^* .

A matriz B^* contém dois parâmetros a menos que a matriz B , além disso o modelo em que usamos B^* parece estar melhor ajustado que o modelo em que usamos a matriz B . Significa que Mod^* parece ser preferível.

Capítulo 6

Conclusões

Ao longo deste trabalho consideramos dois modelos de regressão Dirichlet e apresentamos algumas de suas propriedades. Obtemos as expressões da matriz de informação de Fisher referentes aos dois modelos, como pode ser visto nos apêndices 1 e 2. Apresentamos um teste proposto por Ratsimalahelo (2003) que permite determinar quantos valores singulares de uma matriz são significantemente diferentes de zero. Estimamos o posto b da matriz B de coeficientes destes modelos de regressão através da estatística de teste proposta por Ratsimalahelo (2003), utilizando o procedimento de teste seqüencial e os critérios de informação BIC e HQIC.

Através das simulações de Monte Carlo, observamos na maioria das vezes que o viés e o erro quadrático médio do estimador de b parecem decrescer quando N aumenta. Notamos que na maioria dos casos, as estimativas de b tornam-se mais viesadas quando acrescentamos variáveis explicativas nos modelos e/ou quando aumentamos o número de componentes de cada vetor de observações. Observamos ainda que, se a matriz B não tem posto completo, o procedimento de teste seqüencial tem um desempenho bem melhor, em termos de viés e erro quadrático médio, do que os critérios de informação BIC e HQIC; se a matriz B tem posto completo, o critério de informação HQIC tem desempenho ligeiramente melhor que o critério de informação BIC e o procedimento de teste seqüencial, nos modelos 1 e 2. Então, em geral, recomendamos o procedimento de teste seqüencial, para estimação do posto da matriz dos parâmetros dos modelos de regressão Dirichlet.

Como trabalhos futuros, que poderiam estender essa pesquisa, podemos sugerir a correção da estimativa do teste, o estudo do desempenho do estimador na presença de pontos de alavanca, a proposta de modelos alternativos para dados de proporções baseados na distribuição Dirichlet, o desenvolvimento de medidas de diagnóstico para estes modelos e a identificação da melhor combinação linear no caso do posto incompleto.

Apêndice 1

Apresentamos neste apêndice os cálculos das derivadas, de primeira e segunda ordem, da função de log-verossimilhança com relação aos parâmetros desconhecidos referentes ao modelo 1.

A função de log-verossimilhança para o modelo de regressão Dirichlet definido neste trabalho é da forma

$$\ell(B; \phi) = \sum_{i=1}^N \ell_i(\mu_i; \phi),$$

onde

$$\begin{aligned} \ell_i(\mu_i; \phi) = & \log \Gamma(\phi) - \sum_{j=1}^{p-1} \log \Gamma(\phi \mu_{ij}) - \log \Gamma\left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \\ & + \sum_{j=1}^{p-1} (\phi \mu_{ij} - 1) \log y_{ij} + \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) - 1\right] \log \left(1 - \sum_{j=1}^{p-1} y_{ij}\right). \end{aligned}$$

A derivada de primeira ordem da função de log-verossimilhança com relação aos parâmetros desconhecidos β_{vu} 's, com $v = 1, \dots, k$ e $u = 1, \dots, p-1$, é

$$\begin{aligned} \frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} &= \sum_{i=1}^N \frac{\partial \ell_i(\mu_i; \phi)}{\partial \beta_{vu}} \\ &= \sum_{i=1}^N \frac{\partial}{\partial \beta_{vu}} \left\{ \log \Gamma(\phi) - \sum_{j=1}^{p-1} \log \Gamma(\phi \mu_{ij}) - \log \Gamma\left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \right. \\ &\quad \left. + \sum_{j=1}^{p-1} (\phi \mu_{ij} - 1) \log y_{ij} + \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) - 1\right] \log \left(1 - \sum_{j=1}^{p-1} y_{ij}\right) \right\} \\ &= \sum_{i=1}^N \left\{ \left(- \sum_{j=1}^{p-1} \frac{\partial}{\partial \beta_{vu}} \log \Gamma(\phi \mu_{ij}) \right) - \left(\frac{\partial}{\partial \beta_{vu}} \log \Gamma\left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \right) \right. \\ &\quad \left. + \sum_{j=1}^{p-1} \frac{\partial}{\partial \beta_{vu}} \left((\phi \mu_{ij} - 1) \log y_{ij} \right) + \frac{\partial}{\partial \beta_{vu}} \left(\left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) - 1\right] \right. \right. \\ &\quad \left. \left. \log \left[1 - \sum_{j=1}^{p-1} y_{ij}\right] \right) \right\} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^N \left\{ - \sum_{j=1}^{p-1} \phi \psi(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \frac{\partial \eta_{iu}}{\partial \beta_{vu}} + \phi \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \frac{\partial \eta_{iu}}{\partial \beta_{vu}} \right. \\
&\quad \left. + \sum_{j=1}^{p-1} \phi \log y_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \frac{\partial \eta_{iu}}{\partial \beta_{vu}} - \phi \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \frac{\partial \eta_{iu}}{\partial \beta_{vu}} \right\}
\end{aligned}$$

mas

$$\eta_{ij} = X_{i1}\beta_{1j} + X_{i2}\beta_{2j} + X_{i3}\beta_{3j} + \dots + X_{ik}\beta_{kj},$$

logo

$$\frac{\partial \eta_{ij}}{\partial \beta_{vu}} = \begin{cases} X_{iv} & \text{se } j = u \\ 0 & \text{se } j \neq u, \end{cases} \quad (A1.1)$$

então

$$\begin{aligned}
\frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} &= \sum_{i=1}^N \left\{ - \sum_{j=1}^{p-1} \phi \psi(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{iu}} X_{iv} + \phi \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} X_{iv} \right. \\
&\quad \left. + \sum_{j=1}^{p-1} \phi \log y_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} X_{iv} - \phi \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} X_{iv} \right\} \\
&= \sum_{i=1}^N \phi X_{iv} \left\{ - \sum_{j=1}^{p-1} \psi(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{iu}} + \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right. \\
&\quad \left. + \sum_{j=1}^{p-1} \log y_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} - \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right\} \\
&= \sum_{i=1}^N \phi X_{iv} \sum_{j=1}^{p-1} \left\{ -\psi(\phi \mu_{ij}) + \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] + \log y_{ij} \right. \\
&\quad \left. - \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \right\} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \\
&= \sum_{i=1}^N \phi X_{iv} \sum_{j=1}^{p-1} A_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}}
\end{aligned}$$

onde

$$A_{ij} = -\psi(\phi \mu_{ij}) + \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] + \log y_{ij} - \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right].$$

Temos que

$$\mu_{ij} = \begin{cases} g(\eta_{ij}) & \text{se } j = 1, \\ g(\eta_{ij})[1 - \sum_{t=1}^{j-1} \mu_{it}] & \text{se } j = 2, 3, \dots, p-1, \end{cases} \quad (A1.2)$$

então

$$\frac{\partial \mu_{ij}}{\partial \eta_{iu}} = \begin{cases} 0 & \text{se } j < u, \\ g'(\eta_{iu}) & \text{se } j = u = 1, \\ g'(\eta_{iu})[1 - \sum_{t=1}^{u-1} \mu_{it}] & \text{se } j = u \neq 1, \\ -g'(\eta_{iu})g(\eta_{ij})[1 - \sum_{t=1}^{u-1} \mu_{it}] & \text{se } j = u+1, \\ -g'(\eta_{iu})g(\eta_{ij})[1 - \sum_{t=1}^{u-1} \mu_{it}] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] & \text{se } j \geq u+2. \end{cases} \quad (A1.3)$$

Usando (A1.3) temos que

$$\begin{aligned} \frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} &= \sum_{i=1}^N \phi X_{iv} \left\{ \sum_{j=1}^{p-1} A_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right\} \\ &= \sum_{i=1}^N \phi X_{iv} \left\{ \sum_{j=1}^{u-1} A_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} + A_{iu} \frac{\partial \mu_{iu}}{\partial \eta_{iu}} + A_{i(u+1)} \frac{\partial \mu_{i(u+1)}}{\partial \eta_{iu}} + \sum_{j=u+2}^{p-1} A_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right\} \\ &= \sum_{i=1}^N \phi X_{iv} \left\{ 0 + A_{iu} \frac{\partial \mu_{iu}}{\partial \eta_{iu}} + A_{i(u+1)} \frac{\partial \mu_{i(u+1)}}{\partial \eta_{iu}} + \sum_{j=u+2}^{p-1} A_{ij} \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right\} \\ &= \sum_{i=1}^N \phi X_{iv} \left\{ A_{iu} g'(\eta_{iu})[1 - \sum_{t=1}^{u-1} \mu_{it}] - A_{i(u+1)} g'(\eta_{iu})g(\eta_{ij})[1 - \sum_{t=1}^{u-1} \mu_{it}] \right. \\ &\quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{iu})g(\eta_{ij})[1 - \sum_{t=1}^{u-1} \mu_{it}] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right\} \end{aligned}$$

$$= \sum_{i=1}^N \phi X_{iv} g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\ \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right\}.$$

Portanto,

$$\frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} = \sum_{i=1}^N \phi X_{iv} g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\ \left. - \sum_{j=u+2}^{p-1} \left[A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right] \right\}.$$

Adicionalmente, a derivada de primeira ordem da função de log-verossimilhança com relação ao parâmetro desconhecido ϕ é

$$\frac{\partial \ell(B; \phi)}{\partial \phi} = \sum_{i=1}^N \left\{ \psi(\phi) + \sum_{j=1}^p [-\psi(\phi \mu_{ij}) \mu_{ij} + \log y_{ij} \mu_{ij}] \right\}. \quad (A1.4)$$

Para o cálculo da matriz de informação de Fisher precisamos da derivada de segunda ordem da função de log-verossimilhança com relação aos parâmetros desconhecidos, então, para $v, n \in (1, 2, \dots, k)$ e $u, m \in (1, 2, \dots, p-1)$, temos

$$\frac{\partial}{\partial \beta_{nm}} \left\{ \frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} \right\} = \frac{\partial}{\partial \beta_{nm}} \left\{ \sum_{i=1}^N \phi X_{iv} g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \right. \\ \left. \left. - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\} \\ = \sum_{i=1}^N \phi X_{iv} \frac{\partial}{\partial \beta_{nm}} \left\{ g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \right. \\ \left. \left. - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\} \\ = \sum_{i=1}^N \phi X_{iv} \text{ Der1},$$

onde

$$\begin{aligned}
\text{Der1} &= \frac{\partial}{\partial \beta_{nm}} \left\{ g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \right. \\
&\quad \left. \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right] \right\} \\
&= \frac{\partial}{\partial \beta_{nm}} \left(g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \right. \right. \right. \\
&\quad \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right) + g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \frac{\partial A_{iu}}{\partial \beta_{nm}} - \frac{\partial A_{i(u+1)}}{\partial \beta_{nm}} g(\eta_{i(u+1)}) \right. \\
&\quad - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} - \sum_{j=u+2}^{p-1} \left[\frac{\partial A_{ij}}{\partial \beta_{nm}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \\
&\quad \left. \left. + A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\} \\
&= \left(g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] - g'(\eta_{iu}) \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} \right) \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
&\quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \\
&\quad \left\{ \frac{\partial A_{iu}}{\partial \beta_{nm}} - \frac{\partial A_{i(u+1)}}{\partial \beta_{nm}} g(\eta_{i(u+1)}) - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} \right. \\
&\quad \left. - \sum_{j=u+2}^{p-1} \left[\left(\frac{\partial A_{ij}}{\partial \beta_{nm}} g(\eta_{ij}) + A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \right) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \right. \\
&\quad \left. \left. + A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\} \\
&= \left(g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] - g'(\eta_{iu}) \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} \right) \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
&\quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right]
\end{aligned}$$

$$\begin{aligned}
& \left\{ \frac{\partial A_{iu}}{\partial \beta_{nm}} - \frac{\partial A_{i(u+1)}}{\partial \beta_{nm}} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} \frac{\partial A_{ij}}{\partial \beta_{nm}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \\
& \quad - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} - \sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \\
& \quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\}.
\end{aligned}$$

Tem-se

$$\begin{aligned}
\frac{\partial A_{ij}}{\partial \beta_{nm}} &= \frac{\partial}{\partial \beta_{nm}} \left\{ -\psi(\phi \mu_{ij}) + \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] + \log y_{ij} - \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \right\} \\
&= -\phi \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} X_{in} - \phi \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{im}} X_{in},
\end{aligned}$$

mas

$$\begin{aligned}
\sum_{j=1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{im}} &= \sum_{j=1}^{m-1} \frac{\partial \mu_{ij}}{\partial \eta_{im}} + \frac{\partial \mu_{im}}{\partial \eta_{im}} + \sum_{j=m+1}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{im}} \\
&= 0 + \frac{\partial \mu_{im}}{\partial \eta_{im}} + \frac{\partial \mu_{i(m+1)}}{\partial \eta_{im}} \sum_{j=m+2}^{p-1} \frac{\partial \mu_{ij}}{\partial \eta_{im}} \\
&= g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] g(\eta_{i(m+1)}) \\
&\quad - \sum_{j=m+2}^{p-1} g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \\
&= g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right).
\end{aligned}$$

Logo

$$\begin{aligned}
\frac{\partial A_{ij}}{\partial \beta_{nm}} &= -\phi \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} X_{in} - \phi \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] X_{in} g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \\
&\quad \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \\
&= -\phi X_{in} \left[\psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \\
&\quad \left. \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right].
\end{aligned} \tag{A1.5}$$

Daí,

$$\begin{aligned}
\text{Der1} &= \left(g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] - g'(\eta_{iu}) \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} \right) \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
&\quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \\
&\quad \left\{ -\phi X_{in} \left[\psi'(\phi \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{im}} + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \right. \\
&\quad \left. \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \\
&\quad + \phi X_{in} \left[\psi'(\phi \mu_{i(u+1)}) \frac{\partial \mu_{i(u+1)}}{\partial \eta_{im}} + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \\
&\quad \left. \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right] g(\eta_{i(u+1)}) \\
&\quad + \sum_{j=u+2}^{p-1} \phi X_{in} \left[\psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \\
&\quad \left. \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right] g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]
\end{aligned}$$

$$\begin{aligned}
& - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} - \sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \\
& - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big\} \\
& = \left(g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} [1 - \sum_{t=1}^{u-1} \mu_{it}] - g'(\eta_{iu}) \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} \right) \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
& \quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + g'(\eta_{iu}) [1 - \sum_{t=1}^{u-1} \mu_{it}] \\
& \quad \left\{ \phi X_{in} \left[-\psi'(\phi \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{im}} - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \right. \\
& \quad \left. \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right. \\
& \quad \left. + \psi'(\phi \mu_{i(u+1)}) \frac{\partial \mu_{i(u+1)}}{\partial \eta_{im}} g(\eta_{i(u+1)}) + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \\
& \quad \left. \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) g(\eta_{i(u+1)}) \right. \\
& \quad \left. + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] g'(\eta_{im}) \right. \\
& \quad \left. \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right] - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} - \sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \\
& \quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\} \\
& = \left(g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} [1 - \sum_{t=1}^{u-1} \mu_{it}] - g'(\eta_{iu}) \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} \right) \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
& \quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + g'(\eta_{iu}) [1 - \sum_{t=1}^{u-1} \mu_{it}]
\end{aligned}$$

$$\begin{aligned}
& \left\{ \phi X_{in} \left[-\psi'(\phi\mu_{iu}) \frac{\partial\mu_{iu}}{\partial\eta_{im}} + \psi'(\phi\mu_{i(u+1)}) \frac{\partial\mu_{i(u+1)}}{\partial\eta_{im}} g(\eta_{i(u+1)}) + \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij}) \frac{\partial\mu_{ij}}{\partial\eta_{im}} \right. \right. \\
& g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) \right. \\
& \left. \left. - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Bigg] \\
& - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial\eta_{i(u+1)}}{\partial\beta_{nm}} - \sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{ij}) \frac{\partial\eta_{ij}}{\partial\beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \sum_{j=u+2}^{p-1} A_{ij} \\
& \left. g(\eta_{ij}) \frac{\partial}{\partial\beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\}.
\end{aligned}$$

Para simplificarmos a equação acima devemos considerar três casos, a saber: $u = m$, $u > m$ e $u < m$.

Caso 1: Se $u = m$

Neste caso temos que

$$\sum_{t=1}^{m-1} \frac{\partial\mu_{it}}{\partial\beta_{nm}} = 0$$

$$\sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{ij}) \frac{\partial\eta_{ij}}{\partial\beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] = 0$$

e

$$\frac{\partial}{\partial\beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) = 0$$

pois $u + 1 > u = m$, $u + 2 > u = m$, $\dots$, $j - 1 > u = m$, ou seja,

$$\frac{\partial\eta_{i(u+1)}}{\partial\beta_{nu}} = \frac{\partial\eta_{i(u+2)}}{\partial\beta_{nu}} = \dots = \frac{\partial\eta_{i(j-1)}}{\partial\beta_{nu}} = 0,$$

então

$$\begin{aligned}
\text{Der1} &= g''(\eta_{iu})X_{in}\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left(A_{iu} - A_{i(u+1)}g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij}g(\eta_{ij}) \right. \\
&\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi X_{in}g'(\eta_{iu})\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left\{ -\psi'(\phi\mu_{iu})\frac{\partial\mu_{iu}}{\partial\eta_{iu}} \right. \\
&\quad \left. + \psi'(\phi\mu_{i(u+1)})\frac{\partial\mu_{i(u+1)}}{\partial\eta_{iu}}g(\eta_{i(u+1)}) + \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij})\frac{\partial\mu_{ij}}{\partial\eta_{iu}}g(\eta_{ij}) \right. \\
&\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{iu})\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]\psi'\left[\phi\left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \left(1 - g(\eta_{i(u+1)}) \right. \right. \\
&\quad \left. \left. - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right)^2 \right\} \\
&= g''(\eta_{iu})X_{in}\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left(A_{iu} - A_{i(u+1)}g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij}g(\eta_{ij}) \right. \\
&\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi X_{in}g'(\eta_{iu})\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left\{ -\psi'(\phi\mu_{iu})g'(\eta_{iu})\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \right. \\
&\quad \left. - \psi'(\phi\mu_{i(u+1)})g'(\eta_{iu})\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]g(\eta_{i(u+1)})^2 - \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij})g'(\eta_{iu})\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \right. \\
&\quad \left. g(\eta_{ij})^2 \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]^2 - g'(\eta_{iu})\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]\psi'\left[\phi\left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \left(1 - g(\eta_{i(u+1)}) \right. \right. \\
&\quad \left. \left. - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right)^2 \right\} \\
&= g''(\eta_{iu})X_{in}\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left(A_{iu} - A_{i(u+1)}g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij}g(\eta_{ij}) \right. \\
&\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi X_{in}g'(\eta_{iu})^2\left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]^2 \left\{ -\psi'(\phi\mu_{iu}) - \psi'(\phi\mu_{i(u+1)}) \right. \\
&\quad \left. g(\eta_{i(u+1)})^2 - \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij})g(\eta_{ij})^2 \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]^2 - \psi'\left[\phi\left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \right\}
\end{aligned}$$

$$\left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]\right)^2 \Big\}.$$

Portanto

$$\begin{aligned} \text{Der1} = & g''(\eta_{iu}) X_{in} \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \right. \\ & \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi X_{in} g'(\eta_{iu})^2 \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]^2 \left\{ -\psi'(\phi \mu_{iu}) - \psi'(\phi \mu_{i(u+1)}) \right. \\ & \left. g(\eta_{i(u+1)})^2 - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]^2 - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] \right. \\ & \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]\right)^2 \right\}, \end{aligned}$$

então,

$$\begin{aligned} \frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nm} \partial \beta_{vu}} = & \sum_{i=1}^N \phi X_{iv} X_{in} \left\{ g''(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \right. \right. \\ & \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi g'(\eta_{iu})^2 \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]^2 \left[-\psi'(\phi \mu_{iu}) - \psi'(\phi \mu_{i(u+1)}) \right. \\ & \left. g(\eta_{i(u+1)})^2 - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]^2 - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] \right. \\ & \left. \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]\right)^2 \right] \right\}. \end{aligned}$$

Por (2.2.1.5) sabemos que $E[\log y_{ij}] = \psi(\phi \mu_{ij}) - \psi(\phi)$, então,

$$\begin{aligned} E[A_{ij}] = & E \left\{ -\psi(\phi \mu_{ij}) + \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] + \log y_{ij} - \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \right\} \\ = & -\psi(\phi \mu_{ij}) + \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] + E[\log y_{ij}] - E \left\{ \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \right\} \end{aligned}$$

$$= -\psi(\phi\mu_{ij}) + \psi\left[\phi\left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] + \psi(\phi\mu_{ij}) - \psi(\phi) - \psi\left[\phi\left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] + \psi(\phi) \quad (A1.6)$$

$$= 0.$$

Daí,

$$\begin{aligned} -\mathbb{E}\left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nm} \partial \beta_{vu}}\right] &= \sum_{i=1}^N X_{iv} X_{in} \phi^2 g'(\eta_{iu})^2 \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]^2 \\ &\quad \left\{ \psi'(\phi\mu_{iu}) + \psi'(\phi\mu_{i(u+1)}) g(\eta_{i(u+1)})^2 + \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij}) g(\eta_{ij})^2 \right. \\ &\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]^2 + \psi'\left[\phi\left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \right. \\ &\quad \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]\right)^2 \right\}. \end{aligned}$$

Então,

$$-\mathbb{E}\left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nu} \partial \beta_{vu}}\right] = \sum_{i=1}^N X_{iv} X_{in} U_{uu,i}^{(1)},$$

onde

$$\begin{aligned} U_{uu,i}^{(1)} &= \phi^2 g'(\eta_{iu})^2 \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right]^2 \left\{ \psi'(\phi\mu_{iu}) + \psi'(\phi\mu_{i(u+1)}) g(\eta_{i(u+1)})^2 \right. \\ &\quad + \sum_{j=u+2}^{p-1} \psi'(\phi\mu_{ij}) g(\eta_{ij})^2 \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]^2 + \psi'\left[\phi\left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right)\right] \\ &\quad \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})]\right)^2 \right\}. \end{aligned}$$

Caso 2: Se $u > m$

Neste caso também temos que

$$\frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) = 0,$$

pois $u+1 > u > m$, $u+2 > u > m$, $\dots$, $j-1 > u > m$, ou seja,

$$\frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} = \frac{\partial \eta_{i(u+2)}}{\partial \beta_{nm}} = \dots = \frac{\partial \eta_{i(j-1)}}{\partial \beta_{nm}} = 0,$$

então,

$$\begin{aligned}
\text{Der1} = & -g'(\eta_{iu}) \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \\
& + \phi X_{in} g'(\eta_{iu}) [1 - \sum_{t=1}^{u-1} \mu_{it}] \left\{ -\psi'(\phi \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{im}} + \psi'(\phi \mu_{i(u+1)}) \frac{\partial \mu_{i(u+1)}}{\partial \eta_{im}} g(\eta_{i(u+1)}) \right. \\
& + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) [1 - \sum_{t=1}^{m-1} \mu_{it}] \\
& \left. \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right. \right. \\
& \left. \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\},
\end{aligned}$$

mas

$$\begin{aligned}
\sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} &= \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \eta_{im}} \frac{\partial \eta_{im}}{\partial \beta_{nm}} \\
&= X_{in} \left\{ \sum_{t=1}^{m-1} \frac{\partial \mu_{it}}{\partial \eta_{im}} + \frac{\partial \mu_{im}}{\partial \eta_{im}} + \frac{\partial \mu_{i(m+1)}}{\partial \eta_{im}} + \sum_{t=m+2}^{u-1} \frac{\partial \mu_{it}}{\partial \eta_{im}} \right\} \\
&= X_{in} \left\{ 0 + \frac{\partial \mu_{im}}{\partial \eta_{im}} + \frac{\partial \mu_{i(m+1)}}{\partial \eta_{im}} + \sum_{t=m+2}^{u-1} \frac{\partial \mu_{it}}{\partial \eta_{im}} \right\} \\
&= X_{in} \left\{ g'(\eta_{im}) [1 - \sum_{l=1}^{m-1} \mu_{il}] - g'(\eta_{im}) g(\eta_{i(m+1)}) [1 - \sum_{l=1}^{m-1} \mu_{il}] \right. \\
&\quad \left. - \sum_{t=m+2}^{u-1} g'(\eta_{im}) g(\eta_{it}) [1 - \sum_{l=1}^{m-1} \mu_{il}] \prod_{l=m+1}^{t-1} [1 - g(\eta_{il})] \right\}
\end{aligned}$$

$$\begin{aligned}
&= X_{in} g'(\eta_{im}) [1 - \sum_{l=1}^{m-1} \mu_{il}] \left\{ 1 - g(\eta_{i(m+1)}) - \sum_{t=m+2}^{u-1} g(\eta_{it}) \prod_{l=m+1}^{t-1} [1 - g(\eta_{il})] \right\} \\
&= X_{in} g'(\eta_{im}) [1 - \sum_{l=1}^{m-1} \mu_{il}] \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})],
\end{aligned}$$

logo

$$\begin{aligned}
\text{Der1} &= -g'(\eta_{iu}) g'(\eta_{im}) [1 - \sum_{l=1}^{m-1} \mu_{il}] \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] X_{in} \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
&\quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi X_{in} g'(\eta_{iu}) [1 - \sum_{t=1}^{u-1} \mu_{it}] \left\{ -\psi'(\phi \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{im}} \right. \\
&\quad \left. + \psi'(\phi \mu_{i(u+1)}) \frac{\partial \mu_{i(u+1)}}{\partial \eta_{im}} g(\eta_{i(u+1)}) + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \\
&\quad \left. - g'(\eta_{im}) [1 - \sum_{t=1}^{m-1} \mu_{it}] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \right. \right. \\
&\quad \left. \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\}
\end{aligned}$$

$$\begin{aligned}
&= -g'(\eta_{iu}) g'(\eta_{im}) [1 - \sum_{l=1}^{m-1} \mu_{il}] \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] X_{in} \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
&\quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi X_{in} g'(\eta_{iu}) [1 - \sum_{t=1}^{u-1} \mu_{it}] \left\{ \psi'(\phi \mu_{iu}) \right. \\
&\quad \left. g'(\eta_{im}) g(\eta_{iu}) [1 - \sum_{t=1}^{m-1} \mu_{it}] \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] - \psi'(\phi \mu_{i(u+1)}) g'(\eta_{im}) g(\eta_{i(u+1)})^2 \right. \\
&\quad \left. [1 - \sum_{t=1}^{m-1} \mu_{it}] \prod_{t=m+1}^u [1 - g(\eta_{it})] - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) g'(\eta_{im}) g(\eta_{ij})^2 [1 - \sum_{t=1}^{m-1} \mu_{it}] \right. \\
&\quad \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) [1 - \sum_{t=1}^{m-1} \mu_{it}] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \right\}
\end{aligned}$$

$$\begin{aligned}
& \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big\} \\
& = -g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{l=1}^{m-1} \mu_{il} \right] \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] X_{in} \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
& \quad \left. - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi X_{in} g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \\
& \quad \left\{ \psi'(\phi \mu_{iu}) g(\eta_{iu}) \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] - \psi'(\phi \mu_{i(u+1)}) g(\eta_{i(u+1)})^2 \prod_{t=m+1}^u [1 - g(\eta_{it})] \right. \\
& \quad \left. - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \right. \\
& \quad \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big\}.
\end{aligned}$$

Daí,

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \beta_{nu} \partial \beta_{vu}} &= \sum_{i=1}^N \phi X_{iv} X_{in} \left\{ -g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{l=1}^{m-1} \mu_{il} \right] \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] \left(A_{iu} \right. \right. \\
& \quad \left. - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) + \phi g'(\eta_{iu}) \\
& \quad g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left[\psi'(\phi \mu_{iu}) g(\eta_{iu}) \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] \right. \\
& \quad \left. - \psi'(\phi \mu_{i(u+1)}) g(\eta_{i(u+1)})^2 \prod_{t=m+1}^u [1 - g(\eta_{it})] - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \right. \\
& \quad \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \right] \left(1 - g(\eta_{i(m+1)}) \right. \\
& \quad \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \Big\}
\end{aligned}$$

$$- \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \Big) \Big(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \Big) \Big] \Big\}.$$

Usando (A1.6) temos

$$\begin{aligned} -\mathbb{E} \left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nm} \partial \beta_{vu}} \right] &= \sum_{i=1}^N X_{iv} X_{in} \left\{ \phi^2 g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left[-\psi'(\phi \mu_{iu}) \right. \right. \\ &\quad g(\eta_{iu}) \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] + \psi'(\phi \mu_{i(u+1)}) g(\eta_{i(u+1)})^2 \prod_{t=m+1}^u [1 - g(\eta_{it})] \\ &\quad + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \\ &\quad + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big] \Big\}. \end{aligned}$$

Portanto

$$-\mathbb{E} \left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nm} \partial \beta_{vu}} \right] = \sum_{i=1}^N X_{iv} X_{in} U_{mu,i}^{(1)},$$

onde

$$\begin{aligned}
U_{mu,i}^{(1)} = & \phi^2 g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left[-\psi'(\phi \mu_{iu}) g(\eta_{iu}) \right. \\
& \prod_{t=m+1}^{u-1} [1 - g(\eta_{it})] + \psi'(\phi \mu_{i(u+1)}) g(\eta_{i(u+1)})^2 \prod_{t=m+1}^u [1 - g(\eta_{it})] \\
& + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \\
& + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right].
\end{aligned}$$

Caso 3: Se $u < m$

Aqui temos que

$$\begin{aligned}
\text{Der1} = & -g'(\eta_{iu}) \sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} \left(A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \\
& + g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \phi X_{in} \left[-\psi'(\phi \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{im}} + \psi'(\phi \mu_{i(u+1)}) \frac{\partial \mu_{i(u+1)}}{\partial \eta_{im}} g(\eta_{i(u+1)}) \right. \right. \\
& + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \\
& \left. \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right. \\
& \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} \\
& - \sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \\
& \left. \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\}
\end{aligned}$$

Como $u < m$, então, $u - 1 < m$, logo

$$\sum_{t=1}^{u-1} \frac{\partial \mu_{it}}{\partial \beta_{nm}} = 0.$$

Temos ainda que

$$\frac{\partial \mu_{iu}}{\partial \beta_{nm}} = 0.$$

Daí,

$$\begin{aligned} \text{Der1} = & g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \phi X_{in} \left[\psi'(\phi \mu_{i(u+1)}) \frac{\partial \mu_{i(u+1)}}{\partial \eta_{im}} g(\eta_{i(u+1)}) \right. \right. \\ & + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \\ & \left. \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right. \\ & \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] - A_{i(u+1)} g'(\eta_{i(u+1)}) \frac{\partial \eta_{i(u+1)}}{\partial \beta_{nm}} \\ & - \sum_{j=u+2}^{p-1} A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \sum_{j=u+2}^{p-1} A_{ij} g(\eta_{ij}) \\ & \left. \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\}. \end{aligned}$$

Se $m = u + 1$, então,

$$\begin{aligned} \text{Der1} = & g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \phi X_{in} \left[\psi'(\phi \mu_{im}) \frac{\partial \mu_{im}}{\partial \eta_{im}} g(\eta_{im}) + \sum_{j=m+1}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \right. \right. \\ & \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) \right. \\ & \left. - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \left. \right] \\ & - A_{im} g'(\eta_{im}) X_{in} - \sum_{j=m+1}^{p-1} A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=m}^{j-1} [1 - g(\eta_{it})] \right) \left. \right\}. \end{aligned}$$

Mas

$$\begin{aligned}
\sum_{j=m+1}^{p-1} A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=m}^{j-1} [1 - g(\eta_{it})] \right) &= A_{i(m+1)} g(\eta_{i(m+1)}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=m}^m [1 - g(\eta_{it})] \right) \\
&\quad + \sum_{j=m+2}^{p-1} A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=m}^{j-1} [1 - g(\eta_{it})] \right) \\
&= A_{i(m+1)} g(\eta_{i(m+1)}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=m}^m [1 - g(\eta_{it})] \right) + 0 \\
&= -g'(\eta_{im}) A_{i(m+1)} g(\eta_{i(m+1)}) X_{in}.
\end{aligned}$$

Portanto

$$\sum_{j=m+1}^{p-1} A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=m}^{j-1} [1 - g(\eta_{it})] \right) = -A_{i(m+1)} g'(\eta_{im}) g(\eta_{i(m+1)}) X_{in},$$

logo

$$\begin{aligned}
\text{Der1} &= g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] X_{in} \left\{ \phi \left[\psi'(\phi \mu_{im}) \frac{\partial \mu_{im}}{\partial \eta_{im}} g(\eta_{im}) + \sum_{j=m+1}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \right. \right. \\
&\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) \right. \right. \\
&\quad \left. \left. - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big] \\
&\quad \left. - A_{im} g'(\eta_{im}) + A_{i(m+1)} g'(\eta_{im}) g(\eta_{i(m+1)}) \right\} \\
&= g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] X_{in} \left\{ \phi \left[\psi'(\phi \mu_{im}) \frac{\partial \mu_{im}}{\partial \eta_{im}} g(\eta_{im}) + \psi'(\phi \mu_{i(m+1)}) \frac{\partial \mu_{i(m+1)}}{\partial \eta_{im}} \right. \right. \\
&\quad \left. g(\eta_{i(m+1)}) \prod_{t=u+1}^m [1 - g(\eta_{it})] + \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \\
&\quad \left. - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \right) \right\}
\end{aligned}$$

$$\begin{aligned}
& \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \Big) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big] - A_{im} g'(\eta_{im}) \\
& + A_{i(m+1)} g'(\eta_{im}) g(\eta_{i(m+1)}) \Big\} \\
& = g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] X_{in} \left\{ \phi \left[\psi'(\phi \mu_{im}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] g(\eta_{im}) - \psi'(\phi \mu_{i(m+1)}) \right. \right. \\
& g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] g(\eta_{i(m+1)})^2 \prod_{t=u+1}^m [1 - g(\eta_{it})] - \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \\
& g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \\
& \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] - A_{im} g'(\eta_{im}) + A_{i(m+1)} g'(\eta_{im}) g(\eta_{i(m+1)}) \Big\} \\
& = g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] X_{in} \left\{ \phi \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left[\psi'(\phi \mu_{im}) g(\eta_{im}) - \psi'(\phi \mu_{i(m+1)}) \right. \right. \\
& g(\eta_{i(m+1)})^2 \prod_{t=u+1}^m [1 - g(\eta_{it})] - \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \\
& - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \\
& \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] - A_{im} + A_{i(m+1)} g(\eta_{i(m+1)}) \Big\}.
\end{aligned}$$

Logo, usando (A1.6), temos que, se $m = u + 1$,

$$\begin{aligned}
-\mathbb{E} \left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nm} \partial \beta_{vu}} \right] &= \sum_{i=1}^N X_{iv} X_{in} \left\{ \phi^2 g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \\
& \left. \left[-\psi'(\phi \mu_{im}) g(\eta_{im}) + \psi'(\phi \mu_{i(m+1)}) g(\eta_{i(m+1)})^2 \prod_{t=u+1}^m [1 - g(\eta_{it})] \right] \right\}
\end{aligned}$$

$$\begin{aligned}
& + \sum_{j=m+2}^{p-1} \psi'(\phi\mu_{ij})g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \\
& \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big] \Big\}.
\end{aligned}$$

Mas se $m \geq u + 2$ temos

$$\begin{aligned}
\text{Der1} &= g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \sum_{j=u+2}^{p-1} \left[\phi X_{in} \psi'(\phi\mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \right. \\
& \quad \left. \left. - A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right. \\
& \quad \left. - \phi X_{in} g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \right. \right. \\
& \quad \left. \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right\} \\
&= g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \sum_{j=u+2}^{m-1} \left[\phi X_{in} \psi'(\phi\mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \right. \\
& \quad \left. \left. - A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right. \\
& \quad + \sum_{j=m}^m \left[\phi X_{in} \psi'(\phi\mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \\
& \quad \left. - A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \\
& \quad + \sum_{j=m+1}^{p-1} \left[\phi X_{in} \psi'(\phi\mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \\
& \quad \left. - A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right]
\end{aligned}$$

$$\begin{aligned}
& - \phi X_{in} g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it}\right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big\},
\end{aligned}$$

mas

$$\begin{aligned}
& \sum_{j=u+2}^{m-1} \left[\phi \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} X_{in} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - A_{ij} g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{nm}} \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right. \\
& \left. - A_{ij} g(\eta_{ij}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=m}^{j-1} [1 - g(\eta_{it})] \right) \right] = 0,
\end{aligned}$$

logo

$$\begin{aligned}
\text{Der1} &= g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left\{ \left[\phi X_{in} \psi'(\phi \mu_{im}) \frac{\partial \mu_{im}}{\partial \eta_{im}} g(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \right. \right. \\
& \left. - A_{im} g'(\eta_{im}) \frac{\partial \eta_{im}}{\partial \beta_{nm}} \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] - A_{im} g(\eta_{im}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \right) \right] \\
& + \sum_{j=m+1}^{p-1} \left[\phi X_{in} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - A_{ij} g(\eta_{ij}) \right. \\
& \left. \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] - \phi X_{in} g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it}\right] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij}\right) \right] \\
& \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Big\} \\
&= g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it}\right] \left\{ \left[\phi X_{in} \psi'(\phi \mu_{im}) \frac{\partial \mu_{im}}{\partial \eta_{im}} g(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] - A_{im} g'(\eta_{im}) X_{in} \right. \right. \\
& \left. \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] - A_{im} g(\eta_{im}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \right) \right] + \phi X_{in} \psi'(\phi \mu_{i(m+1)}) \\
& \frac{\partial \mu_{i(m+1)}}{\partial \eta_{im}} g(\eta_{i(m+1)}) \prod_{t=u+1}^m [1 - g(\eta_{it})] - A_{i(m+1)} g(\eta_{i(m+1)}) \frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^m [1 - g(\eta_{it})] \right)
\end{aligned}$$

$$\begin{aligned}
& + \phi X_{in} \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \phi X_{in} g'(\eta_{im}) [1 - \sum_{t=1}^{m-1} \mu_{it}] \\
& \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \\
& \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Bigg\},
\end{aligned}$$

mas

$$\frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \right) = 0$$

e

$$\frac{\partial}{\partial \beta_{nm}} \left(\prod_{t=u+1}^m [1 - g(\eta_{it})] \right) = -g'(\eta_{im}) X_{in} \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})].$$

Então

$$\begin{aligned}
\text{Der1} = & g'(\eta_{iu}) [1 - \sum_{t=1}^{u-1} \mu_{it}] \Bigg\{ \left[\phi X_{in} \psi'(\phi \mu_{im}) \frac{\partial \mu_{im}}{\partial \eta_{im}} g(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] - A_{im} g'(\eta_{im}) X_{in} \right. \\
& \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] + \phi X_{in} \psi'(\phi \mu_{i(m+1)}) \frac{\partial \mu_{i(m+1)}}{\partial \eta_{im}} g(\eta_{i(m+1)}) \prod_{t=u+1}^m [1 - g(\eta_{it})] \\
& + A_{i(m+1)} g(\eta_{i(m+1)}) g'(\eta_{im}) X_{in} \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] + \phi X_{in} \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) \frac{\partial \mu_{ij}}{\partial \eta_{im}} \\
& g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \phi X_{in} g'(\eta_{im}) [1 - \sum_{t=1}^{m-1} \mu_{it}] \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \\
& \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \right. \\
& \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \Bigg\}.
\end{aligned}$$

Daí,

$$\begin{aligned}
\text{Der1} = & X_{in} g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ \phi g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left[\psi'(\phi \mu_{im}) g(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \right. \right. \\
& - \psi'(\phi \mu_{i(m+1)}) g(\eta_{i(m+1)})^2 \prod_{t=u+1}^m [1 - g(\eta_{it})] - \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \\
& \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) \right. \\
& - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \left. \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \left. \right] \\
& - A_{im} g'(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] + A_{i(m+1)} g(\eta_{i(m+1)}) g'(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \left. \right\}.
\end{aligned}$$

Logo

$$\begin{aligned}
\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nu} \partial \beta_{vu}} = & \sum_{i=1}^N \phi^2 X_{iv} X_{in} g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left\{ g'(\eta_{im}) \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \left[\psi'(\phi \mu_{im}) g(\eta_{im}) \right. \right. \\
& \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] - \psi'(\phi \mu_{i(m+1)}) g(\eta_{i(m+1)})^2 \prod_{t=u+1}^m [1 - g(\eta_{it})] \\
& - \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \\
& - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \\
& \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \left. \right] - A_{im} g'(\eta_{im}) \\
& \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] + A_{i(m+1)} g(\eta_{i(m+1)}) g'(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] \left. \right\}.
\end{aligned}$$

Usando (A1.6) temos

$$\begin{aligned}
-\mathbb{E} \left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nm} \partial \beta_{vu}} \right] &= \sum_{i=1}^N X_{iv} X_{in} \left\{ \phi^2 g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \right. \\
&\quad \left[-\psi'(\phi \mu_{im}) g(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] + \psi'(\phi \mu_{i(m+1)}) g(\eta_{i(m+1)})^2 \right. \\
&\quad \prod_{t=u+1}^m [1 - g(\eta_{it})] + \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \\
&\quad \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \right. \\
&\quad \left. \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \left. \right\}
\end{aligned}$$

Portanto

$$-\mathbb{E} \left[\frac{\partial^2 \ell(B; \phi)}{\partial \beta_{nm} \partial \beta_{vu}} \right] = \sum_{i=1}^N X_{iv} X_{in} U_{mu,i}^{(1)},$$

onde

$$\begin{aligned}
U_{mu,i}^{(1)} &= \phi^2 g'(\eta_{iu}) g'(\eta_{im}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[1 - \sum_{t=1}^{m-1} \mu_{it} \right] \\
&\quad \left[-\psi'(\phi \mu_{im}) g(\eta_{im}) \prod_{t=u+1}^{m-1} [1 - g(\eta_{it})] + \psi'(\phi \mu_{i(m+1)}) g(\eta_{i(m+1)})^2 \right. \\
&\quad \prod_{t=u+1}^m [1 - g(\eta_{it})] + \sum_{j=m+2}^{p-1} \psi'(\phi \mu_{ij}) g(\eta_{ij})^2 \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \\
&\quad \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - g(\eta_{i(m+1)}) - \sum_{j=m+2}^{p-1} g(\eta_{ij}) \right. \\
&\quad \left. \prod_{t=m+1}^{j-1} [1 - g(\eta_{it})] \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \left. \right].
\end{aligned}$$

Usando (A1.4) temos que

$$\begin{aligned}
\frac{\partial}{\partial \phi} \left\{ \frac{\partial \ell(B; \phi)}{\partial \phi} \right\} &= \sum_{i=1}^N \frac{\partial}{\partial \phi} \left\{ \psi(\phi) + \sum_{j=1}^p [-\psi(\phi \mu_{ij}) \mu_{ij} + \log y_{ij} \mu_{ij}] \right\} \\
&= \sum_{i=1}^N \left\{ \psi'(\phi) - \sum_{j=1}^p \psi'(\phi \mu_{ij}) \mu_{ij}^2 \right\},
\end{aligned}$$

logo

$$-\mathbb{E} \left[\frac{\partial^2 \ell(\beta, \phi)}{\partial \phi^2} \right] = \sum_{i=1}^N \left\{ -\psi'(\phi) + \sum_{j=1}^p \psi'(\phi \mu_{ij}) \mu_{ij}^2 \right\}.$$

Sabemos que

$$\begin{aligned} \frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} = \sum_{i=1}^N \phi X_{iv} g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] & \left\{ A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \\ & \left. - \sum_{j=u+2}^{p-1} \left[A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right] \right\}, \end{aligned}$$

então

$$\begin{aligned} \frac{\partial}{\partial \phi} \left\{ \frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} \right\} &= \frac{\partial}{\partial \phi} \left\{ \sum_{i=1}^N \phi X_{iv} g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \right. \\ &\quad \left. \left. - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\} \\ &= \sum_{i=1}^N X_{iv} \left\{ \frac{\partial}{\partial \phi} \left[\phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) \right. \right. \\ &\quad \left. \left. - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\} + \phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \\ &\quad \left[\frac{\partial A_{iu}}{\partial \phi} - \frac{\partial A_{i(u+1)}}{\partial \phi} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} \left(\frac{\partial A_{ij}}{\partial \phi} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \Big\} \\ &= \sum_{i=1}^N X_{iv} \left\{ g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \right. \right. \right. \\ &\quad \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] + \phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[\frac{\partial A_{iu}}{\partial \phi} - \frac{\partial A_{i(u+1)}}{\partial \phi} g(\eta_{i(u+1)}) \right. \right. \\ &\quad \left. \left. - \sum_{j=u+2}^{p-1} \left(\frac{\partial A_{ij}}{\partial \phi} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\}, \end{aligned}$$

mas

$$\begin{aligned}
\frac{\partial A_{ij}}{\partial \phi} &= \frac{\partial}{\partial \phi} \left\{ -\psi(\phi \mu_{ij}) + \psi \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] + \log y_{ij} - \log \left[1 - \sum_{j=1}^{p-1} y_{ij} \right] \right\} \\
&= -\psi'(\phi \mu_{ij}) \mu_{ij} + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right),
\end{aligned}$$

logo

$$\begin{aligned}
\frac{\partial}{\partial \phi} \left\{ \frac{\partial \ell(B; \phi)}{\partial \beta_{vu}} \right\} &= \sum_{i=1}^N X_{iv} \left\{ g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \right. \right. \right. \\
&\quad \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] + \phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[-\psi'(\phi \mu_{iu}) \mu_{iu} \right. \\
&\quad + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) + \psi'(\phi \mu_{i(u+1)}) \mu_{i(u+1)} g(\eta_{i(u+1)}) \\
&\quad - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} -\psi'(\phi \mu_{ij}) \mu_{ij} \\
&\quad g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \sum_{j=u+2}^{p-1} g(\eta_{ij}) \\
&\quad \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right] \right\} \\
&= \sum_{i=1}^N X_{iv} \left\{ g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[A_{iu} - A_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} \left(A_{ij} g(\eta_{ij}) \right. \right. \right. \\
&\quad \left. \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] + \phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[-\psi'(\phi \mu_{iu}) \mu_{iu} \right. \\
&\quad + \psi'(\phi \mu_{i(u+1)}) \mu_{i(u+1)} g(\eta_{i(u+1)}) + \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \mu_{ij} g(\eta_{ij}) \\
&\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] + \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right. \\
&\quad \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\}
\end{aligned}$$

Então

$$\begin{aligned}
-\mathbb{E} \left[\frac{\partial^2 \ell(B; \phi)}{\partial \phi \partial \beta_{vu}} \right] &= \sum_{i=1}^N X_{iv} \left\{ \phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[\psi'(\phi \mu_{iu}) \mu_{iu} \right. \right. \\
&\quad \left. \left. - \psi'(\phi \mu_{i(u+1)}) \mu_{i(u+1)} g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \mu_{ij} g(\eta_{ij}) \right. \right. \\
&\quad \left. \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right. \\
&\quad \left. \left. \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right] \right\}.
\end{aligned}$$

Portanto

$$-\mathbb{E} \left[\frac{\partial^2 \ell(B; \phi)}{\partial \phi \partial \beta_{vu}} \right] = \sum_{i=1}^N X_{iv} V_{iu}^{(1)},$$

onde

$$\begin{aligned}
V_{iu}^{(1)} &= \phi g'(\eta_{iu}) \left[1 - \sum_{t=1}^{u-1} \mu_{it} \right] \left[\psi'(\phi \mu_{iu}) \mu_{iu} - \psi'(\phi \mu_{i(u+1)}) \mu_{i(u+1)} g(\eta_{i(u+1)}) \right. \\
&\quad \left. - \sum_{j=u+2}^{p-1} \psi'(\phi \mu_{ij}) \mu_{ij} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] - \psi' \left[\phi \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \right] \right. \\
&\quad \left. \left(1 - \sum_{j=1}^{p-1} \mu_{ij} \right) \left(1 - g(\eta_{i(u+1)}) - \sum_{j=u+2}^{p-1} g(\eta_{ij}) \prod_{t=u+1}^{j-1} [1 - g(\eta_{it})] \right) \right].
\end{aligned}$$

Apêndice 2

Neste Apêndice encontra-se os cálculos das derivadas, de primeira e segunda ordem, da função de log-verossimilhança com relação aos parâmetros desconhecidos referentes ao modelo 2. A função de log-verossimilhança é

$$\ell(B) = \sum_{i=1}^N \ell_i(\mu_i),$$

onde

$$\ell_i(\mu_i) = \log \Gamma(\phi_i) - \sum_{j=1}^p \log \Gamma(\phi_i \mu_{ij}) + \sum_{j=1}^p (\phi_i \mu_{ij} - 1) \log y_{ij},$$

então a derivada de primeira ordem com relação aos parâmetros β_{vu} 's, com $v = 1, \dots, k$ e $u = 1, \dots, p$, é

$$\begin{aligned} \frac{\partial \ell(B)}{\partial \beta_{vu}} &= \sum_{i=1}^N \frac{\partial}{\partial \beta_{vu}} \left[\log \Gamma(\phi_i) - \sum_{j=1}^p \log \Gamma(\phi_i \mu_{ij}) + \sum_{j=1}^p (\phi_i \mu_{ij} - 1) \log y_{ij} \right] \\ &= \sum_{i=1}^N \left\{ \frac{\partial}{\partial \beta_{vu}} [\log \Gamma(\phi_i)] + \sum_{j=1}^p \left[\frac{\partial}{\partial \beta_{vu}} (-\log \Gamma(\phi_i \mu_{ij})) \right. \right. \\ &\quad \left. \left. + \frac{\partial}{\partial \beta_{vu}} ((\phi_i \mu_{ij} - 1) \log y_{ij}) \right] \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p \left[-\psi(\phi_i \mu_{ij}) \frac{\partial}{\partial \beta_{vu}} [\phi_i \mu_{ij}] + \log y_{ij} \frac{\partial}{\partial \beta_{vu}} [\phi_i \mu_{ij}] \right] \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p \left[(\log y_{ij} - \psi(\phi_i \mu_{ij})) \frac{\partial}{\partial \beta_{vu}} [\phi_i \mu_{ij}] \right] \right\} \\ &= \sum_{i=1}^N \left\{ \psi(\phi_i) \frac{\partial \phi_i}{\partial \beta_{vu}} + \sum_{j=1}^p \left[A_{ij}^* \frac{\partial}{\partial \beta_{vu}} [\phi_i \mu_{ij}] \right] \right\} \end{aligned}$$

onde

$$A_{ij}^* = \log y_{ij} - \psi(\phi_i \mu_{ij})$$

e

$$\begin{aligned}
\frac{\partial}{\partial \beta_{vu}} [\phi_i \mu_{ij}] &= \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \beta_{vu}} \\
&= \frac{\partial \phi_i}{\partial \beta_{vu}} \mu_{ij} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} X_{iv}.
\end{aligned}$$

Sabemos que

$$\mu_{ij} = \frac{g(\eta_{ij})}{g(\eta_{i1}) + g(\eta_{i2}) + \dots + g(\eta_{ip})} = \frac{g(\eta_{ij})}{\phi_i} \quad (A2.2)$$

e

$$\begin{aligned}
\frac{\partial \phi_i}{\partial \beta_{vu}} &= \frac{\partial}{\partial \beta_{vu}} \sum_{j=1}^p g(\eta_{ij}) = \sum_{j=1}^p \frac{\partial}{\partial \beta_{vu}} g(\eta_{ij}) \\
&= \sum_{j=1}^p g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{vu}} \\
&= \sum_{\substack{j=1 \\ j \neq u}}^p g'(\eta_{ij}) \frac{\partial \eta_{ij}}{\partial \beta_{vu}} + g'(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{vu}} \\
&= 0 + g'(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{vu}} = g'(\eta_{iu}) X_{iv},
\end{aligned} \quad (A2.3)$$

então,

$$\frac{\partial}{\partial \beta_{vu}} [\phi_i \mu_{ij}] = g'(\eta_{iu}) X_{iv} \frac{g(\eta_{ij})}{\phi_i} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} X_{iv}.$$

Logo

$$\begin{aligned}
\frac{\partial \ell(B)}{\partial \beta_{vu}} &= \sum_{i=1}^N \left\{ \psi(\phi_i) g'(\eta_{iu}) X_{iv} + \sum_{j=1}^p A_{ij}^* \left[g'(\eta_{iu}) \frac{g(\eta_{ij})}{\phi_i} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right] X_{iv} \right\} \\
&= \sum_{i=1}^N X_{iv} \left\{ \psi(\phi_i) g'(\eta_{iu}) + \frac{g'(\eta_{iu})}{\phi_i} \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) + \phi_i \sum_{j=1}^p A_{ij}^* \frac{\partial \mu_{ij}}{\partial \eta_{iu}} \right\} \\
&= \sum_{i=1}^N X_{iv} \left\{ \psi(\phi_i) g'(\eta_{iu}) + \frac{g'(\eta_{iu})}{\phi_i} \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) + \phi_i \left[\sum_{\substack{j=1 \\ j \neq u}}^p A_{ij}^* \frac{\partial \mu_{ij}}{\partial \eta_{iu}} + A_{iu}^* \frac{\partial \mu_{iu}}{\partial \eta_{iu}} \right] \right\}.
\end{aligned}$$

Por (A2.2) temos que

$$\frac{\partial \mu_{ij}}{\partial \eta_{iu}} = \begin{cases} \frac{g'(\eta_{iu})[\phi_i - g(\eta_{iu})]}{\phi_i^2} & \text{se } j = u, \\ \frac{-g'(\eta_{iu})g(\eta_{ij})}{\phi_i^2} & \text{se } j \neq u, \end{cases} \quad (A2.4)$$

então,

$$\begin{aligned} \frac{\partial \ell(B)}{\partial \beta_{vu}} &= \sum_{i=1}^N X_{iv} \left\{ \psi(\phi_i) g'(\eta_{iu}) + \frac{g'(\eta_{iu})}{\phi_i} \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) + \phi_i \left[\sum_{\substack{j=1 \\ j \neq u}}^p -A_{ij}^* \frac{g'(\eta_{iu})g(\eta_{ij})}{\phi_i^2} \right. \right. \\ &\quad \left. \left. + A_{iu}^* \frac{g'(\eta_{iu})[\phi_i - g(\eta_{iu})]}{\phi_i^2} \right] \right\} \\ &= \sum_{i=1}^N X_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) - \sum_{\substack{j=1 \\ j \neq u}}^p A_{ij}^* g(\eta_{ij}) + A_{iu}^* [\phi_i - g(\eta_{iu})] \right\} \\ &= \sum_{i=1}^N X_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) - \sum_{\substack{j=1 \\ j \neq u}}^p A_{ij}^* g(\eta_{ij}) - A_{iu}^* g(\eta_{iu}) \right. \\ &\quad \left. + A_{iu}^* \phi_i \right\} \\ &= \sum_{i=1}^N X_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) - \sum_{j=1}^p A_{ij}^* g(\eta_{ij}) + A_{iu}^* \phi_i \right\} \\ &= \sum_{i=1}^N X_{iv} \frac{g'(\eta_{iu})}{\phi_i} \left\{ \phi_i \psi(\phi_i) + A_{iu}^* \phi_i \right\} \\ &= \sum_{i=1}^N X_{iv} \left\{ g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}^*] \right\}. \end{aligned}$$

Portanto

$$\frac{\partial \ell(B)}{\partial \beta_{vu}} = \sum_{i=1}^N X_{iv} \left\{ g'(\eta_{iu}) [\psi(\phi_i) + A_{iu}^*] \right\}.$$

A derivada de segunda ordem da função de log-verossimilhança com relação aos parâmetros desconhecidos β_{nm} 's foi calculada para que se obtenha a matriz de informação de Fisher; então, para $v, n \in (1, 2, \dots, k)$ e $u, m \in (1, 2, \dots, p)$ temos

$$\begin{aligned} \frac{\partial}{\partial \beta_{nm}} \left\{ \frac{\partial \ell(B)}{\partial \beta_{vu}} \right\} &= \frac{\partial}{\partial \beta_{nm}} \left\{ \sum_{i=1}^N X_{iv} \left[g'(\eta_{iu}) \left(\psi(\phi_i) + A_{iu}^* \right) \right] \right\} \\ &= \sum_{i=1}^N X_{iv} \left\{ \frac{\partial}{\partial \beta_{nm}} \left[g'(\eta_{iu}) \left(\psi(\phi_i) + A_{iu}^* \right) \right] \right\} \\ &= \sum_{i=1}^N X_{iv} \text{ Der2}, \end{aligned}$$

onde

$$\text{Der2} = \frac{\partial}{\partial \beta_{nm}} \left[g'(\eta_{iu}) \left(\psi(\phi_i) + A_{iu}^* \right) \right],$$

ou ainda,

$$\begin{aligned} \text{Der2} &= \frac{\partial}{\partial \beta_{nm}} [g'(\eta_{iu})] \left(\psi(\phi_i) + A_{iu}^* \right) + g'(\eta_{iu}) \frac{\partial}{\partial \beta_{nm}} \left(\psi(\phi_i) + A_{iu}^* \right) \\ &= g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left(\psi(\phi_i) + A_{iu}^* \right) + g'(\eta_{iu}) \left(\frac{\partial}{\partial \beta_{nm}} \psi(\phi_i) + \frac{\partial A_{iu}^*}{\partial \beta_{nm}} \right) \quad (A2.5) \\ &= g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left(\psi(\phi_i) + A_{iu}^* \right) + g'(\eta_{iu}) \left(\psi'(\phi_i) \frac{\partial \phi_i}{\partial \beta_{nm}} + \frac{\partial A_{iu}^*}{\partial \beta_{nm}} \right). \end{aligned}$$

Sabemos que

$$\begin{aligned} \frac{\partial A_{ij}^*}{\partial \beta_{nm}} &= \frac{\partial}{\partial \beta_{nm}} [\log y_{ij} - \psi(\phi_i \mu_{ij})] \\ &= -\psi'(\phi_i \mu_{ij}) \frac{\partial}{\partial \beta_{nm}} [\phi_i \mu_{ij}] \quad (A2.6) \\ &= -\psi'(\phi_i \mu_{ij}) X_{in} \left[g'(\eta_{im}) \frac{g(\eta_{ij})}{\phi_i} + \phi_i \frac{\partial \mu_{ij}}{\partial \eta_{im}} \right]. \end{aligned}$$

Substituindo (A2.3) e (A2.6) em (A2.5), temos

$$\begin{aligned}
\text{Der2} &= g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left(\psi(\phi_i) + A_{iu}^* \right) + g'(\eta_{iu}) \left(\psi'(\phi_i) g'(\eta_{im}) X_{in} - \psi'(\phi_i \mu_{iu}) X_{in} \right. \\
&\quad \left. \left[g'(\eta_{im}) \frac{g(\eta_{iu})}{\phi_i} + \phi_i \frac{\partial \mu_{iu}}{\partial \eta_{im}} \right] \right) \\
&= g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left(\psi(\phi_i) + A_{iu}^* \right) + g'(\eta_{iu}) g'(\eta_{im}) \psi'(\phi_i) X_{in} - g'(\eta_{iu}) g'(\eta_{im}) \psi'(\phi_i \mu_{iu}) \\
&\quad X_{in} \frac{g(\eta_{iu})}{\phi_i} - g'(\eta_{iu}) \psi'(\phi_i \mu_{iu}) X_{in} \phi_i \frac{\partial \mu_{iu}}{\partial \eta_{im}} \\
&= g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nm}} \left(\psi(\phi_i) + A_{iu}^* \right) + \frac{g'(\eta_{iu}) g'(\eta_{im}) X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right] \\
&\quad - \phi_i g'(\eta_{iu}) X_{in} \psi'(\phi_i \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{im}}.
\end{aligned} \tag{A2.7}$$

Para simplificarmos a equação acima devemos considerar dois casos, a saber: $u = m$ e $u \neq m$

Caso 1: Se $u = m$

Neste caso temos que

$$\begin{aligned}
\text{Der2} &= g''(\eta_{iu}) \frac{\partial \eta_{iu}}{\partial \beta_{nu}} \left(\psi(\phi_i) + A_{iu}^* \right) + \frac{g'(\eta_{iu}) g'(\eta_{iu}) X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right] \\
&\quad - \phi_i g'(\eta_{iu}) X_{in} \psi'(\phi_i \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{iu}} \\
&= g''(\eta_{iu}) X_{in} \left(\psi(\phi_i) + A_{iu}^* \right) + \frac{g'(\eta_{iu})^2 X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right] \\
&\quad - \phi_i g'(\eta_{iu}) X_{in} \psi'(\phi_i \mu_{iu}) \frac{g'(\eta_{iu}) [\phi_i - g(\eta_{iu})]}{\phi_i^2} \\
&= g''(\eta_{iu}) X_{in} \left(\psi(\phi_i) + A_{iu}^* \right) + \frac{g'(\eta_{iu})^2 X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right. \\
&\quad \left. - \psi'(\phi_i \mu_{iu}) [\phi_i - g(\eta_{iu})] \right]
\end{aligned}$$

$$\begin{aligned}
&= g''(\eta_{iu}) X_{in} \left(\psi(\phi_i) + A_{iu}^* \right) + \frac{g'(\eta_{iu})^2 X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right. \\
&\quad \left. - \psi'(\phi_i \mu_{iu}) \phi_i + \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right] \\
&= g''(\eta_{iu}) X_{in} \left(\psi(\phi_i) + A_{iu}^* \right) + \frac{g'(\eta_{iu})^2 X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) \phi_i \right] \\
&= g''(\eta_{iu}) X_{in} \left(\psi(\phi_i) + A_{iu}^* \right) + g'(\eta_{iu})^2 X_{in} \left[\psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) \right],
\end{aligned}$$

logo,

$$\frac{\partial^2 \ell(B)}{\partial \beta_{nm} \partial \beta_{vu}} = \sum_{i=1}^N X_{iv} X_{in} \left\{ g''(\eta_{iu}) \left(\psi(\phi_i) + A_{iu}^* \right) + g'(\eta_{iu})^2 \left[\psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) \right] \right\}.$$

Por (2.2.1.5) sabemos que $E[\log y_{ij}] = \psi(\phi_i \mu_{ij}) - \psi(\phi_i)$, então

$$\begin{aligned}
E[A_{ij}^*] &= E[\log y_{ij} - \psi(\phi_i \mu_{ij})] \\
&= E[\log y_{ij}] - \psi(\phi_i \mu_{ij}) \\
&= \psi(\phi_i \mu_{ij}) - \psi(\phi_i) - \psi(\phi_i \mu_{ij}) \\
&= -\psi(\phi_i).
\end{aligned} \tag{A2.8}$$

Usando (A2.8) temos

$$\begin{aligned}
-E \left[\frac{\partial^2 \ell(B)}{\partial \beta_{nm} \partial \beta_{vu}} \right] &= \sum_{i=1}^N X_{iv} X_{in} \left\{ g''(\eta_{iu}) \left(-\psi(\phi_i) - E[A_{iu}^*] \right) - g'(\eta_{iu})^2 \left[\psi'(\phi_i) \right. \right. \\
&\quad \left. \left. - \psi'(\phi_i \mu_{iu}) \right] \right\} \\
&= \sum_{i=1}^N X_{iv} X_{in} \left\{ g''(\eta_{iu}) \left(-\psi(\phi_i) + \psi(\phi_i) \right) - g'(\eta_{iu})^2 \left[\psi'(\phi_i) \right. \right. \\
&\quad \left. \left. - \psi'(\phi_i \mu_{iu}) \right] \right\} \\
&= \sum_{i=1}^N X_{iv} X_{in} \left\{ -g'(\eta_{iu})^2 \left[\psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) \right] \right\}.
\end{aligned}$$

Portanto

$$-\mathbb{E} \left[\frac{\partial^2 \ell(B)}{\partial \beta_{nu} \partial \beta_{vu}} \right] = \sum_{i=1}^N X_{iv} X_{in} U_{uu,i}^{(2)},$$

onde

$$U_{uu,i}^{(2)} = -g'(\eta_{iu})^2 \left[\psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) \right].$$

Caso 2: Se $u \neq m$

Neste caso temos que (A2.7) se reduz a

$$\begin{aligned} \text{Der2} &= \frac{g'(\eta_{iu}) g'(\eta_{im}) X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right] \\ &\quad - \phi_i g'(\eta_{iu}) X_{in} \psi'(\phi_i \mu_{iu}) \frac{\partial \mu_{iu}}{\partial \eta_{im}} \\ &= \frac{g'(\eta_{iu}) g'(\eta_{im}) X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right] \\ &\quad + \phi_i g'(\eta_{iu}) X_{in} \psi'(\phi_i \mu_{iu}) \frac{g'(\eta_{im}) g(\eta_{iu})}{\phi_i^2} \\ &= \frac{g'(\eta_{iu}) g'(\eta_{im}) X_{in}}{\phi_i} \left[\phi_i \psi'(\phi_i) - \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) + \psi'(\phi_i \mu_{iu}) g(\eta_{iu}) \right] \\ &= g'(\eta_{iu}) g'(\eta_{im}) X_{in} \psi'(\phi_i), \end{aligned}$$

logo

$$\frac{\partial^2 \ell(B)}{\partial \beta_{nm} \partial \beta_{vu}} = \sum_{i=1}^N X_{iv} X_{in} \left\{ g'(\eta_{iu}) g'(\eta_{im}) \psi'(\phi_i) \right\}.$$

Usando (A2.8) temos

$$-\mathbb{E} \left[\frac{\partial^2 \ell(B)}{\partial \beta_{nm} \partial \beta_{vu}} \right] = \sum_{i=1}^N X_{iv} X_{in} \left\{ -g'(\eta_{iu}) g'(\eta_{im}) \psi'(\phi_i) \right\}.$$

Portanto

$$-\mathrm{E}\left[\frac{\partial^2 \ell(B)}{\partial \beta_{nm} \partial \beta_{vu}}\right] = \sum_{i=1}^N X_{iv} X_{in} U_{mu,i}^{(2)},$$

onde

$$U_{mu,i}^{(2)} = -g'(\eta_{iu}) g'(\eta_{im}) \psi'(\phi_i).$$

Apêndice 3

Algoritmo BFGS

Um dos problemas mais freqüentes na inferência estatística é encontrar o estimador de máxima verossimilhança $\hat{\theta}$ de θ , que em geral não apresenta uma forma analítica fechada para seu cálculo. Para encontrar $\hat{\theta}$ devemos maximizar a log-verossimilhança e para isso precisamos usar técnicas de maximização. A literatura divide as técnicas de maximização em classes; por exemplo, existem técnicas que não fazem uso de derivadas e outras que requerem derivadas da função objetivo. Dentre os algoritmos que utilizam derivadas podemos encontrar duas classes: gradientes conjugados, que fazem uso da matriz hessiana, e variáveis métricas, quando uma aproximação da matriz hessiana é empregada. Na primeira classe os métodos mais conhecidos são Newton-Raphson, score e steepest ascent. Já dentre os algoritmos que empregam uma aproximação para a matriz hessiana, também chamados de métodos quasi-Newton, há duas técnicas principais: BFGS e DFP. No decorrer deste trabalho usamos a técnica de maximização BFGS.

O algoritmo de Newton-Raphson utiliza aproximações por série de Taylor até segunda ordem. Ou seja, numa vizinhança de um ponto $\theta(t)$, uma função pode ser aproximada como

$$w(\theta) = w(\theta(t)) + (\theta - \theta(t))^T \nabla w(\theta(t)) + \frac{1}{2}(\theta - \theta(t))^T H(\theta(t))(\theta - \theta(t)),$$

onde $\nabla w(\theta(t))$ é o gradiente de w avaliado em $\theta(t)$, $H(\theta(t))$ sendo a matriz hessiana avaliada em $\theta(t)$ e θ pertencente a uma vizinhança de $\theta(t)$. Desta expressão obtém-se

$$\nabla w(\theta) = \nabla w(\theta(t)) + H(\theta(t))(\theta - \theta(t)).$$

Então faz-se $\nabla w(\theta(t)) = 0$ para determinar o próximo ponto da iteração :

$$0 = \nabla w(\theta(t)) + H(\theta(t))(\theta - \theta(t)) \Rightarrow \theta = \theta(t) - H(\theta(t))^{-1} \nabla w(\theta(t)),$$

onde θ é uma aproximação para o ponto de máximo da função w .

Assim, aproximações para o ponto de máximo desta função são obtidos através da lei de recorrência

$$\theta(t+1) = \theta(t) - H(\theta(t))^{-1} \nabla w(\theta(t)),$$

que podemos generalizar por

$$\theta(t+1) = \theta(t) - s(t)H(\theta(t))^{-1} \nabla w(\theta(t)),$$

onde $s(t)$ é um escalar determinado por algum procedimento de busca linear a partir de $\theta(t)$ na direção $-H(\theta(t))^{-1} \nabla w(\theta(t))$ de forma que $w(\theta(t))$ cresça nesta direção. A velocidade de crescimento de w é caracterizada por

$$\nabla^T w(\theta(t))(\theta - \theta(t)) > 0 \Rightarrow -(\theta - \theta(t))^T H(\theta(t))(\theta - \theta(t)).$$

Assim, H deve ser negativa-definida para todo t . Se θ está longe do máximo, H pode não ser negativa-definida e o incremento $-s(t)H(\theta(t))^{-1}\nabla w(\theta(t))$ pode mover o ponto $\theta(t)$ para um ponto $\theta(t+1)$ onde o valor da função é menor, provocando a não-convergência do algoritmo. Uma opção é utilizar o algoritmo BFGS que não apresenta este problema.

O método BFGS tem esse nome por ter sido desenvolvido por Broyden, Fletcher, Goldfarb e Shanno. E o algoritmo deste método está implementado na plataforma `Opt` através da função `MaxBFGS`, maiores detalhes, ver Doornik (2001). O método segue o mesmo princípio do método Newton–Rapson, só que, ao invés de trabalhar com a matriz $-H^{-1}$, é utilizada uma matriz simétrica e positiva-definida $Q(t)$ tal que

$$\lim_{t \rightarrow \infty} Q(t) = -H^{-1}.$$

Comumente, toma-se como matriz inicial a identidade da mesma ordem, pois ela é positiva-definida e simétrica, assim conduzindo a aproximações $Q(t)$ positivas-definidas e simétricas.

A forma recursiva para obter tais matrizes é dada pela expressão

$$\begin{aligned} Q(t+1) = Q(t) &+ \frac{(\theta(t+1) - \theta(t))(\theta(t+1) - \theta(t))^{\top}}{(\theta(t+1) - \theta(t))^{\top}(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))} \\ &- \frac{[Q(t)(\nabla w(\theta(t+1)) - \nabla w(\theta(t))))^{\top}][Q(t)(\nabla w(\theta(t+1)) - \nabla w(\theta(t))))^{\top}]}{(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))^{\top}Q(t)(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))} \\ &+ (\nabla w(\theta(t+1)) - \nabla w(\theta(t)))^{\top}Q(t)(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))U U^{\top}, \end{aligned}$$

onde U é o vetor coluna definido por

$$\begin{aligned} U = &\frac{\theta(t+1) - \theta(t)}{(\theta(t+1) - \theta(t))^{\top}(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))} \\ &- \frac{Q(t)(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))}{(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))^{\top}Q(t)(\nabla w(\theta(t+1)) - \nabla w(\theta(t)))}. \end{aligned}$$

Assim, mesmo que θ esteja longe do ponto máximo, as matrizes $Q(t)$ garantem que os pontos se moverão na direção crescente. De forma análoga ao método de Newton–Raphson, o máximo é obtido pela recorrência

$$\theta(t+1) = \theta(t) - s(t)Q(\theta(t))\nabla w(\theta(t)).$$

Apêndice 4

Neste apêndice apresentamos os dois conjuntos de dados reais que utilizamos no capítulo 5. O primeiro conjunto de dados, que denotamos por Dados 1, é referente à proporção de pólen fossilizado de três tipos de flores colhidos em três locais diferentes que chamaremos de A, B e C. O segundo conjunto de dados, Dados 2, é referente à proporção de três tipos de leucócitos em diferentes amostras de sangue de dez pacientes determinadas por quatro métodos diferentes que denotaremos por A, B e C. Os três tipos de leucócitos são: P: leucócitos polimorfonucleares, S: linfócitos pequenos e L: mononucleares grandes.

Dados 1

Observação	Proporção		
	Pinus	Abies	Quercus
A1	0,297	0,583	0,120
A2	0,184	0,484	0,331
A3	0,478	0,436	0,086
A4	0,392	0,542	0,065
A5	0,201	0,506	0,293
A6	0,308	0,527	0,164
A7	0,239	0,603	0,158
A8	0,204	0,586	0,210
A9	0,185	0,581	0,234
A10	0,170	0,521	0,309
B1	0,240	0,363	0,397
B2	0,356	0,449	0,195
B3	0,183	0,444	0,373
B4	0,347	0,504	0,149
B5	0,365	0,458	0,177
B6	0,219	0,484	0,297
B7	0,221	0,448	0,332
B8	0,338	0,475	0,187
B9	0,246	0,446	0,308
B10	0,170	0,434	0,396
C1	0,416	0,440	0,144
C2	0,341	0,479	0,180
C3	0,440	0,416	0,143
C4	0,607	0,366	0,027
C5	0,667	0,299	0,034
C6	0,545	0,362	0,094
C7	0,486	0,408	0,107
C8	0,468	0,451	0,081
C9	0,707	0,253	0,040
C10	0,494	0,396	0,111

Datos 2

Nº de pacientes	Métodos	P	S	L
1	A	0,750	0,160	0,090
	B	0,350	0,620	0,030
	C	0,740	0,210	0,050
	D	0,690	0,270	0,040
2	A	0,660	0,240	0,100
	B	0,330	0,660	0,010
	C	0,440	0,530	0,030
	D	0,540	0,410	0,050
3	A	0,830	0,100	0,070
	C	0,730	0,220	0,050
	D	0,820	0,150	0,030
4	A	0,570	0,110	0,320
	B	0,230	0,570	0,200
	C	0,320	0,460	0,220
	D	0,350	0,400	0,250
5	A	0,610	0,110	0,280
	B	0,240	0,600	0,160
	C	0,325	0,495	0,180
	D	0,370	0,420	0,210
6	A	0,510	0,140	0,350
	B	0,380	0,480	0,140
	C	0,395	0,440	0,165
	D	0,420	0,400	0,180
7	A	0,560	0,180	0,260
	B	0,710	0,230	0,060
	C	0,700	0,140	0,160
	D	0,560	0,230	0,210
8	A	0,610	0,090	0,300
	B	0,280	0,540	0,180
	C	0,260	0,560	0,180
	D	0,440	0,360	0,200
9	B	0,320	0,610	0,070
	C	0,540	0,300	0,160
	D	0,430	0,430	0,140
10	A	0,740	0,180	0,080
	B	0,440	0,540	0,020
	C	0,660	0,310	0,030
	D	0,680	0,280	0,040

Apêndice 5

Programa de simulação

Neste apêndice apresentamos os programas de simulação utilizados nesta dissertação; os programas foram desenvolvidos na linguagem de programação `0x`. Temos dois programas, um referente ao modelo 1 e o outro ao modelo 2; dividimos cada um desses programas em três partes. Na primeira parte, que está apresentada na seção A5.1, encontram-se os programas da simulação de Monte Carlo para os modelos 1 e 2; a segunda parte, seção A5.2, contém as funções que permitem o cálculo das estimativas de máxima verossimilhança de B , o preenchimento da matriz de informação de Fisher e a estimação do posto b usando o teste seqüencial e os critérios de informação BIC e HQIC; na terceira parte, seção A5.3, encontram-se as diretrizes de ligação das partes `modelo1.ox` com `funcoes1.ox` e `modelo2.ox` com `funcoes2.ox`.

A5.1. Programas principais

- Modelo 1

```
/******  
* PROGRAMA: modelo1.ox  
* AUTORA: Tatiane Ferreira do Nascimento Melo  
* DATA: 27 de fevereiro de 2004  
* ULTIMA MODIFICACAO: 30 de agosto de 2004   Hora: 19:00  
*****/  
  
/* Arquivos principais */  
#include <oxstd.h>  
#include <oxfloat.h>  
#include <oxprob.h>  
#import <maximize>  
#import <solvenle>  
#include <funcoes1.h>  
#include <funcoes1.ox>  
  
// Variaveis Globais  
static decl soma_est;  
  
/* Parametros da simulacao */  
const decl REP = 10000; // numero de replicas Monte Carlo  
const decl MIN = 50;    // tamanho minimo da amostra  
const decl MAX = 200;   // tamanho maximo da amostra  
  
/* Corpo Principal */  
main()  
{  
    // Declaracao de variaveis  
    decl dExecTime, i, j, mc, a, b, ir, n, cont_falhas, m_aux,  
        m_inv, somm, parametro, v_parametro_ini, vX, Y, mZ,  
        m_alpha, m_emv, Sigma_inv, mZ_Estim, m_resid_trans,  
        teste, m_sigma2_trans, mB_verd, m_B_ini, phi_verd,  
        phi, phi_ini, phi_est, vec_mB_verd, vp1, vp2, vp4,  
        vp5, vp6, vp7, dfunc1, dfunc4, dfunc5, dfunc6, dfunc7,  
        // Variveis com relao ao teste sequencial  
        cont_aceitar_seq, aux_seq, posto_estimado_seq, Prop_seq,  
        E_posto_estimado_seq, EQM_posto_estimado_seq,  
        E_posto2_estimado_seq, Vies_posto_estimado_seq,
```

```

// Variveis com relao ao critrio de informao BIC
cont_aceitar_BIC, aux_BIC, posto_estimado_BIC, Prop_BIC,
E_posto_estimado_BIC, EQM_posto_estimado_BIC,
E_posto2_estimado_BIC, Vies_posto_estimado_BIC,
// Variveis com relao ao critrio de informao HQIC
cont_aceitar_HQIC, aux_HQIC, posto_estimado_HQIC, Prop_HQIC,
E_posto_estimado_HQIC, EQM_posto_estimado_HQIC,
E_posto2_estimado_HQIC, Vies_posto_estimado_HQIC;

// Inicio do relógio para o cálculo do tempo de execução
dExecTime = timer();

oxwarning(0);

// Gerador de números uniformes pseudo-aleatórios( George Marsaglia )
ranseed("GM");

// Numero de colunas da matriz X (matriz de regressores)
k = 3; // k = 3; k = 4; k = 5; k = 6;
// Numero de colunas da matriz Y (num. de componentes)
p = 3; // p-1 = 1; p-1 = 2; p-1 = 3;

//Parâmetros verdadeiros do modelo
//Matriz B verdadeira

mB_verd = zeros(k,p-1); //posto=0

//Matriz 3x1
//mB_verd = <1;0;0>; //posto=1

//Matrizes 3x2
//mB_verd = <1,0;0,0;0,0>; //posto=1
//mB_verd = <1,0;0,1;0,0>; //posto=2

//Matrizes 3x3
//mB_verd = <1,0,0;0,0,0;0,0,0>; //posto=1
//mB_verd = <1,0,0;0,1,0;0,0,0>; //posto=2
//mB_verd = <1,0,0;0,1,0;0,0,1>; //posto=3

//Matriz 4x1
//mB_verd = <1;0;0>; //posto=1

//Matrizes 4x2
//mB_verd = <1,0;0,0;0,0>; //posto=1
//mB_verd = <1,0;0,1;0,0>; //posto=2

//Matrizes 4x3
//mB_verd = <1,0,0;0,0,0;0,0,0>; //posto=1
//mB_verd = <1,0,0;0,1,0;0,0,0>; //posto=2
//mB_verd = <1,0,0;0,1,0;0,0,1>; //posto=3

//Matriz 5x1
//mB_verd = <1;0;0>; //posto=1

//Matrizes 5x2
//mB_verd = <1,0;0,0;0,0>; //posto=1
//mB_verd = <1,0;0,1;0,0>; //posto=2

//Matrizes 5x3
//mB_verd = <1,0,0;0,0,0;0,0,0>; //posto=1
//mB_verd = <1,0,0;0,1,0;0,0,0>; //posto=2
//mB_verd = <1,0,0;0,1,0;0,0,1>; //posto=3

//Matriz 6x1
//mB_verd = <1;0;0>; //posto=1

//Matrizes 6x2
//mB_verd = <1,0;0,0;0,0>; //posto=1
//mB_verd = <1,0;0,1;0,0>; //posto=2

//Matrizes 6x3
//mB_verd = <1,0,0;0,0,0;0,0,0>; //posto=1
//mB_verd = <1,0,0;0,1,0;0,0,0>; //posto=2
//mB_verd = <1,0,0;0,1,0;0,0,1>; //posto=3

parametro = vec(mB_verd); // vetorização da matriz B verdadeira

```

```

// Impressao de resultados da simulacao
print("\n\t\t DATA: ", date());
print("\n\t\t HORA: ", time());
print("\n\t\t PROGRAMA OX: ", oxfilename(0));
print("\n\t\t VERSAO OX: ", oxversion());
print("\n\t\t NUM. MIN. DE OBS.: ", MIN);
print("\n\t\t NUM. MAX. DE OBS.: ", MAX);
print("\n\t\t NUM. REP. DE MONTE CARLO: ", REP);
print("\n\t\t NUM. REGRESSORES DA MATRIZ X: k = ", k);
print("\n\t\t NUM. COMPONENTES DE Y: p = ", p);
print("\n\t\tMatriz B verdadeira ");
println("%10.1f",mB_verd);
println("\n\t\tPosto Verdadeiro = ", rank(mB_verd),"n");

// Parametro phi verdadeiro
phi_verd = 30;

parametro = vec(mB_verd) phi_verd;

// Loop para os diferentes tamanhos de amostra
for (ta = MIN; ta <= MAX; ta+= 50)
{
    ranseed({1965, 2001});

    //Inicializacao das matrizes
    m_mi = zeros(ta,p);
    m_mi_est = zeros(ta,p);
    m_mi_trans = zeros(ta,p-1);
    eta_trans = zeros(ta,p-1);
    m_resid_trans = zeros(ta,p-1);
    m_sigma2_trans = zeros(ta,p-1);
    teste = zeros(ta,p-1);
    m_emv = zeros(k*(p-1)+1,REP);
    mZ = zeros(ta,p-1);
    m_aux = zeros(ta,p);
    mY = zeros(ta,p);
    m_alpha = zeros(ta,p);
    somm = zeros(ta,1);
    Y = zeros(ta,p-1);
    cont_aceitar_seq = zeros(p,1);
    aux_seq = zeros(p,1);
    cont_aceitar_BIC = zeros(p,1);
    aux_HQIC = zeros(p,1);
    cont_aceitar_HQIC = zeros(p,1);
    aux_BIC = zeros(p,1);

    // Matriz de dados explicativos gerador dos regressores X
    vX = ranu(ta,k-1);

    // Matriz das variaveis regressoras
    mX = ones(ta,1)~vX;

    // Matriz eta
    m_eta = mX*mB_verd;

    // Loop para preencher a matriz das medias
    for (i = 0; i < ta; i++)
    { soma = 0.0;
      m_mi[i][0] = exp(m_eta[i][0])/(1.0 + exp(m_eta[i][0]));
      for (j = 1; j < p-1; j++)
      { soma = soma + m_mi[i][j-1];
        m_mi[i][j] = (exp(m_eta[i][j])/(1.0 + exp(m_eta[i][j])))*(1.0 - soma);
      }
      m_mi[i][p-1] = 1.0 - sumr(m_mi[i][0:(p-2)]);
    } // Fim do loop para preencher a matriz das medias

    // Inversao auxiliar: ((X'X)^-1)X'
    m_inv = (invertsym(mX'*mX))*mX';

    // Inicializacao do contador de falhas de convergencia
    cont_falhas = 0;

    ranseed( -1 );

    // Matriz alpha

```

```

m_alpha = phi_verd*m_mi;

// Loop de Monte Carlo
for (mc = 0; mc < REP; ++mc)
{
    // Loop para gerar a variavel resposta Y
    for (i = 0; i < ta; i++)
    {
        Y[i][] = randirichlet(1, m_alpha[i][]);
    } // Fim do loop para gerar a variavel resposta Y

    // Matriz Y
    mY = Y^(1.0 - sumr(Y[][0:p-2]));

    // Loop para preencher a matriz resposta transformada Z
    for (i = 0; i < ta; i++)
    {
        soma3 = 0.0;
        mZ[][0] = log(mY[][0]) - log(1.0 - mY[][0]);
        for (j = 1; j < p-1; j++)
        {
            soma3 = soma3 + mY[i][j-1];
            m_aux[i][j] = mY[i][j] / (1 - soma3);
            mZ[i][j] = log(m_aux[i][j]) - log(1.0 - m_aux[i][j]);
        }
    } // Fim do loop para preencher a matriz resposta transformada Z

    // Estimacao inicial da matriz B
    m_B_ini = m_inv*mZ;

    // Estimacao inicial de phi
    // Loop para preencher a matriz mi transformada
    for (i = 0; i < ta; i++)
    {
        soma1 = 0.0;
        eta_trans[i][0] = mX[i][]*m_inv*mZ[][0];
        m_mi_trans[i][0] = (exp(eta_trans[i][0])/(1.0 + exp(eta_trans[i][0])));
        m_resid_trans[][0] = mZ[][0] - (mX*m_inv*mZ[][0]);
        teste[i][0] = (m_resid_trans[][0]*m_resid_trans[][0])*
            (m_mi_trans[i][0] .* (1.0 - m_mi_trans[i][0]));
        for (j = 1; j < p-1; j++)
        {
            soma1 = soma1 + m_mi_trans[i][j-1];
            eta_trans[i][j] = mX[i][j]*m_inv*mZ[][j];
            m_mi_trans[i][j] = (exp(eta_trans[i][j])/(1.0 + exp(eta_trans[i][j])))*
                (1 - soma1);
            m_resid_trans[][j] = mZ[][j] - (mX*m_inv*mZ[][j]);
            teste[i][j] = (m_resid_trans[][j]*m_resid_trans[][j])*
                (m_mi_trans[i][j] .* (1.0 - m_mi_trans[i][j]));
        }
    } // Fim do loop para preencher a matriz mi transformada

    a = (ta - k)/(ta*(p-1));
    b = sumc(sumr(1.0 ./teste));

    // Valor inicial para o phi
    phi_ini = (a*b) - 1.0;

    // Vetor dos parametros iniciais (valores iniciais)
    v_parametro_ini = vec(m_B_ini) phi_ini;

    // Valores iniciais dos parametros
    vp1 = v_parametro_ini;

    // Maximizacao da funcao de log-verossimilhanca pelo Metodo BFGS
    // Parametros da maximizacao por BFGS
    MaxControl(50, -1);

    // Estimacao por Maxima verossimilhanca
    ir = MaxBFGS(floglik1, &vp1, &dfunc1, 0, TRUE);

    // Checar a convergencia
    if(ir == MAX_CONV | ir == MAX_WEAK_CONV)
    {
        // Vetor de parametros estimado por maxima verossimilhanca
        m_env[mc] = vp1;
    } // Fim do if da convergencia
    else
    {
        ++cont_falhas;
    }

    // Matriz B estimada e phi estimado
    B_est1 = shape(vp1[0:(k*(p-1))-1], k, p-1);
}

```

```

phi_est = vp1[k*(p-1)];

// Matriz eta estimada
m_eta_est = mX*B_est1;

// Loop para preencher a matriz das medias estimadas
for (i = 0; i < ta; i++)
{ soma_est = 0.0;
  m_mi_est[i][0] = exp(m_eta_est[i][0])/(1.0 + exp(m_eta_est[i][0]));
  for (j = 1; j < p-1; j++)
  { soma_est = soma_est + m_mi_est[i][j-1];
    m_mi_est[i][j] = (exp(m_eta_est[i][j])/(1.0 + exp(m_eta_est[i][j])))*
      (1.0 - soma_est);
  }
  m_mi_est[i][p-1] = 1.0 - (sumr(m_mi_est[i][0:(p-2)]))';
} // Fim do loop para preencher a matriz das medias estimadas

// funcao de ligacao
gg = (exp(m_eta_est)./(1.0 + exp(m_eta_est)));

// primeira derivada da funcao de ligacao
gg1 = exp(m_eta_est)./((1.0 + exp(m_eta_est)).^2);

// segunda derivada da funcao de ligacao
gg2 = (exp(m_eta_est) - exp(2 .* m_eta_est))./((1.0 + exp(m_eta_est)).^3);

// matriz mi estimada
mi1 = m_mi_est;

// Preenche a matriz inversa de informacao de Fisher
vp5 = phi_est;
finfFishermodelo1(vp5, &dfunc5);
Sigma_inv = dfunc5;

// Inicio do teste sequencial
vp4 = Sigma_inv;
ftestemodelo1(vp4, &dfunc4);
posto_estimado_seq = dfunc4;
for (i = 0; i <= p-1; i++)
{
  if( posto_estimado_seq == i)
  {cont_aceitar_seq[i][0]++;}
  aux_seq[i][0] = i;
} // Fim do teste sequencial

// Inicio do criterio de informacao BIC
vp6 = Sigma_inv;
ftesteBIC1(vp6, &dfunc6);
posto_estimado_BIC = dfunc6;
for (i = 0; i <= p-1; i++)
{
  if( posto_estimado_BIC == i)
  {cont_aceitar_BIC[i][0]++;}
  aux_BIC[i][0] = i;
} // Fim do criterio de informacao BIC

// Inicio do criterio de informacao HQIC
vp7 = Sigma_inv;
ftesteHQIC1(vp7, &dfunc7);
posto_estimado_HQIC = dfunc7;
for (i = 0; i <= p-1; i++)
{
  if( posto_estimado_HQIC == i)
  {cont_aceitar_HQIC[i][0]++;}
  aux_HQIC[i][0] = i;
} // Fim do criterio de informacao HQIC

} // Fim do loop de Monte Carlo

Prop_seq = (cont_aceitar_seq/REP);
Prop_BIC = (cont_aceitar_BIC/REP);
Prop_HQIC = (cont_aceitar_HQIC/REP);

//***** Vieses e EQM para o posto estimado usando o teste sequencial*****

E_posto_estimado_seq = 0.0;
for (i = 0; i < p; i++)

```

```

{E_posto_estimado_seq = E_posto_estimado_seq + i*Prop_seq[i][0];}

Vies_posto_estimado_seq = E_posto_estimado_seq - rank(mB_verd);

E_posto2_estimado_seq = 0.0;
for (i = 0; i < p; i++)
{E_posto2_estimado_seq = E_posto2_estimado_seq + ((i^2)*Prop_seq[i][0]);} // E(k^2)

EQM_posto_estimado_seq = E_posto2_estimado_seq - 2*rank(mB_verd)*E_posto_estimado_seq +
(rank(mB_verd)^2);

//***** Vieses e EQM para o posto estimado usando o crit. BIC *****

E_posto_estimado_BIC = 0.0;
for (i = 0; i < p; i++)
{E_posto_estimado_BIC = E_posto_estimado_BIC + i*Prop_BIC[i][0];}

Vies_posto_estimado_BIC = E_posto_estimado_BIC - rank(mB_verd);

E_posto2_estimado_BIC = 0.0;
for (i = 0; i < p; i++)
{E_posto2_estimado_BIC = E_posto2_estimado_BIC + ((i^2)*Prop_BIC[i][0]);} // E(k^2)

EQM_posto_estimado_BIC = E_posto2_estimado_BIC - 2*rank(mB_verd)*E_posto_estimado_BIC +
(rank(mB_verd)^2);

//***** Vieses e EQM para o posto estimado usando o crit. HQIC *****

E_posto_estimado_HQIC = 0.0;
for (i = 0; i < p; i++)
{E_posto_estimado_HQIC = E_posto_estimado_HQIC + i*Prop_HQIC[i][0];}

Vies_posto_estimado_HQIC = E_posto_estimado_HQIC - rank(mB_verd);

E_posto2_estimado_HQIC = 0.0;
for (i = 0; i < p; i++)
{E_posto2_estimado_HQIC = E_posto2_estimado_HQIC + ((i^2)*Prop_HQIC[i][0]);} // E(k^2)

EQM_posto_estimado_HQIC = E_posto2_estimado_HQIC - 2*rank(mB_verd)*E_posto_estimado_HQIC +
(rank(mB_verd)^2);

//***** IMPRESSAO DOS RESULTADOS *****

println("\n\n\n***** MODELO 1 *****");
println("\n\t\t\t\t\tTamanho amostral: ", ta);
println("\n-----Teste Sequencial-----");
println("\nPosto \t\t", "Num. de replicas \t\t", "Probabilidade ");
for (i = 0; i <= p-1; i++)
{
    println("    ",aux_seq[i][0], " \t\t\t ",
        cont_aceitar_seq[i][0], " \t\t ", "%15.4f",Prop_seq[i][0]);
}
println("\n Vies_posto_estimado_seq: ", "%10.4f",Vies_posto_estimado_seq);
println("\n EQM_posto_estimado_seq: ", "%10.4f",EQM_posto_estimado_seq);

println("\n----- Criterio: BIC-----");
println("\nPosto \t\t", "Num. de replicas \t\t", "Probabilidade ");
for (i = 0; i <= p-1; i++)
{
    println("    ",aux_BIC[i][0], " \t\t\t ",
        cont_aceitar_BIC[i][0], " \t\t ", "%15.4f",Prop_BIC[i][0]);
}
println("\n Vies_posto_estimado_BIC: ", "%10.4f",Vies_posto_estimado_BIC);
println("\n EQM_posto_estimado_BIC: ", "%10.4f",EQM_posto_estimado_BIC);

println("\n----- Criterio: HQIC-----");
println("\nPosto \t\t", "Num. de replicas \t\t", "Probabilidade ");
for (i = 0; i <= p-1; i++)
{
    println("    ",aux_HQIC[i][0], " \t\t\t ",
        cont_aceitar_HQIC[i][0], " \t\t ", "%15.4f",Prop_HQIC[i][0]);
}
println("\n Vies_posto_estimado_HQIC: ", "%10.4f",Vies_posto_estimado_HQIC);
println("\n EQM_posto_estimado_HQIC: ", "%10.4f",EQM_posto_estimado_HQIC);

} // Fim do loop do tamanho de amostras

print("\n\n\t\t Data: ", date());

```

```

        print("\n\n\t\t Hora: ", time());
        print("\n\n\t\t Tempo total de execucao: ", timespan(dExecTime), "\tminutos");
        print("\n");
    } // Fim do corpo principal

```

● Modelo 2

```

/*****
* PROGRAMA: modelo2.ox
* AUTORA: Tatiane Ferreira do Nascimento Melo
* DATA: 27 de fevereiro de 2004
* ULTIMA MODIFICACAO: 30 de agosto de 2004   Hora: 19:00
*****/

/* Arquivos principais */
#include <oxstd.h>
#include <oxfloat.h>
#include <oxprob.h>
#include <maximize>
#include <solvenle>
#include <funcoes2.h>
#include <funcoes2.ox>

/* Parametros da simulacao */
const decl REP = 10000; // numero de replicas Monte Carlo
const decl MIN = 50; // tamanho minimo da amostra
const decl MAX = 200; // tamanho maximo da amostra

/* Corpo Principal */
main()
{
    // Declaracao de variaveis
    decl dExecTime, phi_verd, vX, cont_falhas, mc, m_alpha, i, j, m_g_eta, mY1,
    Matriz_inf_Fisher, ir, m_emv, somm, Y, mB_verd, mB_ini, phi, posto,
    dfunc, dfunc1, dfunc2, dfunc3, dfunc4, vp1, vp3, vp4, vp5, vp7,
    B_est_rep, B_est_barra,
    // Variveis com relao ao seq
    posto_estimado_seq, Prop_seq, aux_seq, cont_aceitar_seq,
    E_posto_estimado_seq, Vies_posto_estimado_seq, E_posto2_estimado_seq,
    EQM_posto_estimado_seq,
    // Variveis com relao ao BIC
    posto_estimado_BIC, Prop_BIC, aux_BIC, cont_aceitar_BIC,
    E_posto_estimado_BIC, Vies_posto_estimado_BIC, E_posto2_estimado_BIC,
    EQM_posto_estimado_BIC,
    // Variveis com relao ao HQIC
    posto_estimado_HQIC, Prop_HQIC, aux_HQIC, cont_aceitar_HQIC,
    E_posto_estimado_HQIC, Vies_posto_estimado_HQIC, E_posto2_estimado_HQIC,
    EQM_posto_estimado_HQIC;

    // Inicio do relógio para o calculo do tempo de execucao
    dExecTime = timer();

    // turn off warnings
    oxwarning(0);

    // Gerador de numeros uniformes pseudo-aleatorios (George Marsaglia)
    ranseed("GM");

    // Numero de colunas da matriz X (matriz de regressores)
    k = 3; // k = 3; k = 4; k = 5; k = 6;
    // Numero de colunas da matriz Y (num. de componentes)
    p = 3; // p = 2; p = 3; p = 4;

    // Parametros verdadeiros do modelo
    // Matriz B verdadeira

    mB_verd = zeros(k,p); //posto=0

    // Matrices 3x2
    // mB_verd = <1,0;0,0;0,0>; //posto=1

```

```

// mB_verd = <1,0;0,1;0,0>; //posto=2

// Matrices 3x3
// mB_verd = <1,0,0;0,0,0;0,0,0>; //posto=1
// mB_verd = <1,0,0;0,1,0;0,0,0>; //posto=2
// mB_verd = <1,0,0;0,1,0;0,0,1>; //posto=3

// Matrices 4x2
// mB_verd = <1,0;0,0;0,0;0,0>; //posto=1
// mB_verd = <1,0;0,1;0,0;0,0>; //posto=2

// Matrices 4x3
// mB_verd = <1,0,0;0,0,0;0,0,0;0,0,0>; //posto=1
// mB_verd = <1,0,0;0,1,0;0,0,0;0,0,0>; //posto=2
// mB_verd = <1,0,0;0,1,0;0,0,1;0,0,0>; //posto=3

// Matrices 4x4
// mB_verd = <1,0,0,0;0,0,0,0;0,0,0,0;0,0,0,0>; //posto=1
// mB_verd = <1,0,0,0;0,1,0,0;0,0,0,0;0,0,0,0>; //posto=2
// mB_verd = <1,0,0,0;0,1,0,0;0,0,1,0;0,0,0,0>; //posto=3
// mB_verd = <1,0,0,0;0,1,0,0;0,0,1,0;0,0,0,1>; //posto=4

// Matrices 5x2
// mB_verd = <1,0;0,0;0,0;0,0;0,0>; //posto=1
// mB_verd = <1,0;0,1;0,0;0,0;0,0>; //posto=2

// Matrices 5x3
// mB_verd = <1,0,0;0,0,0;0,0,0;0,0,0;0,0,0>; //posto=1
// mB_verd = <1,0,0;0,1,0;0,0,0;0,0,0;0,0,0>; //posto=2
// mB_verd = <1,0,0;0,1,0;0,0,1;0,0,0;0,0,0>; //posto=3

// Matrices 5x4
// mB_verd = <1,0,0,0;0,0,0,0;0,0,0,0;0,0,0,0;0,0,0,0>; //posto=1
// mB_verd = <1,0,0,0;0,1,0,0;0,0,0,0;0,0,0,0;0,0,0,0>; //posto=2
// mB_verd = <1,0,0,0;0,1,0,0;0,0,1,0;0,0,0,0;0,0,0,0>; //posto=3
// mB_verd = <1,0,0,0;0,1,0,0;0,0,1,0;0,0,0,1;0,0,0,0>; //posto=4

// Matrices 6x2
// mB_verd = <1,0;0,0;0,0;0,0;0,0;0,0>; //posto=1
// mB_verd = <1,0;0,1;0,0;0,0;0,0;0,0>; //posto=2

// Matrices 6x3
// mB_verd = <1,0,0;0,0,0;0,0,0;0,0,0;0,0,0;0,0,0>; //posto=1
// mB_verd = <1,0,0;0,1,0;0,0,0;0,0,0;0,0,0;0,0,0>; //posto=2
// mB_verd = <1,0,0;0,1,0;0,0,1;0,0,0;0,0,0;0,0,0>; //posto=3

// Matrices 6x4
// mB_verd = <1,0,0,0;0,0,0,0;0,0,0,0;0,0,0,0;0,0,0,0;0,0,0,0>; //posto=1
// mB_verd = <1,0,0,0;0,1,0,0;0,0,0,0;0,0,0,0;0,0,0,0;0,0,0,0>; //posto=2
// mB_verd = <1,0,0,0;0,1,0,0;0,0,1,0;0,0,0,0;0,0,0,0;0,0,0,0>; //posto=3
// mB_verd = <1,0,0,0;0,1,0,0;0,0,1,0;0,0,0,1;0,0,0,0;0,0,0,0>; //posto=4

// Impressao de resultados da simulacao
print("\n\t\t DATA: ", date());
print("\n\t\t HORA: ", time());
print("\n\n\t\t PROGRAMA OX: ", oxfilename(0));
print("\n\t\t VERSAO OX: ", oxversion());
print("\n\t\t NUM. MIN. DE OBS.: ", MIN);
print("\n\t\t NUM. MAX. DE OBS.: ", MAX);
print("\n\t\t NUM. REP. DE MONTE CARLO: ", REP);
print("\n\t\t NUM. REGRESSORES DA MATRIZ X: k = ", k);
print("\n\t\t NUM. COMPONENTES DE Y: p = ", p);
print("\n\n\t\tMatriz B verdadeira ");
println("%10.1f", mB_verd);
println("\n\t\tPosto Verdadeiro = ", rank(mB_verd), "\n");

// Loop para os diferentes tamanhos de amostra
for (ta = MIN; ta <= MAX; ta+= 50)
{
    ranseed({1965, 2001});

    //Inicializacao das matrizes
    m_mi = zeros(ta,p);
    m_emv = zeros(k*p,REP);
    mY = zeros(ta,p);
    somm = zeros(ta,1);
    Y = zeros(ta,p-1);
    B_est_rep = zeros(k*p,REP);

```

```

cont_aceitar_seq = zeros(p+1,1);
aux_seq = zeros(p+1,1);
cont_aceitar_BIC = zeros(p+1,1);
aux_BIC = zeros(p+1,1);
cont_aceitar_HQIC = zeros(p+1,1);
aux_HQIC = zeros(p+1,1);

// Matriz de dados explicativos gerador dos regressores X
vX = ranu(ta,k-1);

// Matriz das variaveis regressoras
mX = ones(ta,1)~vX;

// Inicializacao do contador de falhas de convergencia
cont_falhas = 0;

ranseed(-1);

// Matrizes eta, g(eta), mi e alpha
m_eta = mX*mB_verd;
m_g_eta = exp(m_eta);
m_mi = m_g_eta ./ sumr(m_g_eta);
m_alpha = m_g_eta;

// Loop de Monte Carlo
for (mc = 0; mc < REP; ++mc)
{
    // Loop para gerar a variavel resposta Y
    for (i = 0; i < ta; i++)
    {
        Y[i][] = randirichlet(1, m_alpha[i][]);
        somm[i][] = sumr(Y[i][]);
    } // Fim do loop para gerar a variavel resposta Y

    // Matriz Y
    mY = Y^(1.0 - somm);

    // Estimacao inicial da matriz B
    m_B_ini = zeros(k,p);

    // Vetor dos parametros iniciais (chute inicial)
    vp1 = vec(m_B_ini);

    // Maximizacao da funcao de log-verossimilhanca pelo Metodo BFGS
    // Parametros da maximizacao por BFGS
    MaxControl(50, -1);

    // Estimacao por Maxima verossimilhanca
    ir = MaxBFGS(floglik2, &vp1, &dfunc1, 0, TRUE);

    // Checar a convergencia
    if(ir == MAX_CONV | ir == MAX_WEAK_CONV)
    {
        // Vetor de parametros estimado por maxima verossimilhanca
        m_env[mc] = vp1;
    } // Fim do if da convergencia

    else
    {++cont_falhas;}

    // Matriz B estimada
    B_est = shape(vp1[0:((k*p)-1)], k, p);

    m_eta_est = mX*B_est;
    m_g_eta_est = exp(m_eta_est);
    m_mi_est = m_g_eta_est ./ sumr(m_g_eta_est);
    g_est = m_g_eta_est;
    g1_est = m_g_eta_est;

    // Preenche a matriz inversa de informacao de Fisher
    phi = sumr(m_g_eta_est); //phi
    finfFishermodelo2(phi, &dfunc);
    Matriz_inf_Fisher = dfunc;

    // Inicio do teste sequencial
    vp4 = Matriz_inf_Fisher;
    ftestmodelo2(vp4, &dfunc2);
    posto_estimado_seq = dfunc2;
    for (i = 0; i <= p; i++)
    {

```

```

        if(posto_estimado_seq == i)
        {cont_aceitar_seq[i][0]++;}
        aux_seq[i][0] = i;
    } // Fim do teste sequencial

    // Inicio do criterio BIC
    vp5 = Matriz_inf_Fisher;
    ftesteBIC2(vp5, &dfunc3);
    posto_estimado_BIC = dfunc3;
    for (i = 0; i <= p; i++)
    {
        if( posto_estimado_BIC == i)
        {cont_aceitar_BIC[i][0]++;}
        aux_BIC[i][0] = i;
    } // Fim do criterio BIC

    // Inicio do criterio HQIC
    vp7 = Matriz_inf_Fisher;
    ftesteHQIC2(vp7, &dfunc4);
    posto_estimado_HQIC = dfunc4;
    for (i = 0; i <= p; i++)
    {
        if( posto_estimado_HQIC == i)
        {cont_aceitar_HQIC[i][0]++;}
        aux_HQIC[i][0] = i;
    } // Fim do criterio HQIC

    B_est_rep[mc] = vec(B_est);
} // Fim do loop de Monte Carlo

Prop_BIC = cont_aceitar_BIC/REP;
Prop_seq = cont_aceitar_seq/REP;
Prop_HQIC = cont_aceitar_HQIC/REP;

//***** Vieses e EQM para o posto estimado usando o teste sequencial*****

E_posto_estimado_seq = 0.0;
for (i = 0; i < p+1; i++)
{E_posto_estimado_seq = E_posto_estimado_seq + i*Prop_seq[i][0];}

Vies_posto_estimado_seq = E_posto_estimado_seq - rank(mB_verd);

E_posto2_estimado_seq = 0.0;
for (i = 0; i < p+1; i++)
{E_posto2_estimado_seq = E_posto2_estimado_seq + ((i^2)*Prop_seq[i][0]);} // E(k^2)

EQM_posto_estimado_seq = E_posto2_estimado_seq - 2*rank(mB_verd)
                        *E_posto_estimado_seq + (rank(mB_verd)^2);

//***** Vieses e EQM para o posto estimado usando o teste BIC *****

E_posto_estimado_BIC = 0.0;
for (i = 0; i < p+1; i++)
{E_posto_estimado_BIC = E_posto_estimado_BIC + i*Prop_BIC[i][0];}

Vies_posto_estimado_BIC = E_posto_estimado_BIC - rank(mB_verd);

E_posto2_estimado_BIC = 0.0;
for (i = 0; i < p+1; i++)
{E_posto2_estimado_BIC = E_posto2_estimado_BIC + ((i^2)*Prop_BIC[i][0]);} // E(k^2)

EQM_posto_estimado_BIC = E_posto2_estimado_BIC - 2*rank(mB_verd)
                        *E_posto_estimado_BIC + (rank(mB_verd)^2);

//***** Vieses e EQM para o posto estimado usando o teste HQIC *****

E_posto_estimado_HQIC = 0.0;
for (i = 0; i < p+1; i++)
{E_posto_estimado_HQIC = E_posto_estimado_HQIC + i*Prop_HQIC[i][0];}

Vies_posto_estimado_HQIC = E_posto_estimado_HQIC - rank(mB_verd);

E_posto2_estimado_HQIC = 0.0;
for (i = 0; i < p+1; i++)
{E_posto2_estimado_HQIC = E_posto2_estimado_HQIC + ((i^2)*Prop_HQIC[i][0]);} // E(k^2)

```

```

EQM_posto_estimado_HQIC = E_posto2_estimado_HQIC - 2*rank(mB_verd)
                        *E_posto_estimado_HQIC + (rank(mB_verd)^2);

//##### IMPRESSAO DOS RESULTADOS #####

println("\n\n***** Modelo 2 *****");
println("\n\t\t\t\t\tTamanho amostral: ", ta);
println("\n-----Teste Sequencial-----");
println("\nPosto \t\t", "Num. de replicas \t\t", "Probabilidade ");
for (i = 0; i <= p; i++)
{
    println("    ",aux_seq[i][0], " \t\t\t\t ",
            cont_aceitar_seq[i][0], " \t\t\t\t ", "%15.4f",Prop_seq[i][0]);
}
println("\n Vies_posto_estimado_seq: ", "%10.4f",Vies_posto_estimado_seq);
println("\n EQM_posto_estimado_seq: ", "%10.4f",EQM_posto_estimado_seq);
println("\n----- Criterio: BIC-----");
println("\nPosto \t\t", "Num. de replicas \t\t", "Probabilidade ");
for (i = 0; i <= p; i++)
{
    println("    ",aux_BIC[i][0], " \t\t\t\t ",
            cont_aceitar_BIC[i][0], " \t\t\t\t ", "%15.4f",Prop_BIC[i][0]);
}
println("\n Vies_posto_estimado_BIC: ", "%10.4f",Vies_posto_estimado_BIC);
println("\n EQM_posto_estimado_BIC: ", "%10.4f",EQM_posto_estimado_BIC);
println("\n----- Criterio: HQIC-----");
println("\nPosto \t\t", "Num. de replicas \t\t", "Probabilidade ");
for (i = 0; i <= p; i++)
{
    println("    ",aux_HQIC[i][0], " \t\t\t\t ",
            cont_aceitar_HQIC[i][0], " \t\t\t\t ", "%15.4f",Prop_HQIC[i][0]);
}
println("\n Vies_posto_estimado_HQIC: ", "%10.4f",Vies_posto_estimado_HQIC);
println("\n EQM_posto_estimado_HQIC: ", "%10.4f",EQM_posto_estimado_HQIC);

} // Fim do loop do tamanho de amostras

print("\n\n\t\t\t\t\tData: ", date());
print("\n\n\t\t\t\t\tHora: ", time());
print("\n\n\t\t\t\t\tTempo total de execucao: ", timespan(dExecTime), "\tminutos");
print("\n\n");

} // Fim do corpo principal

```

A5.2. Biblioteca de funções

• Modelo 1

```
/******
* PROGRAMA: funcoes1.ox
* AUTORA: Tatiane Ferreira do Nascimento Melo
* DATA: 27 de fevereiro de 2004
* ULTIMA MODIFICACAO: 30 de agosto de 2004   Hora: 19:00
*****/

/* Arquivos principais */
#include <oxstd.h>
#include <oxfloat.h>
#include <oxprob.h>
#import <maximize>
#import <solvenle>

// Variaveis Globais
static decl mX, mY, p, k, ta, gg, gg1, gg2;
static decl m_mi, m_eta, m_mi_trans, eta_trans, mi1;
static decl soma, soma1, soma2, soma3, M, M4;
static decl B_est, B_est1, m_eta_est, m_g_eta_est, m_mi_est, g_est, g1_est;

//=====

// Funcao log-verossimilhanca do modelo1 da regressao Dirichlet
floglik1(const vP, const adFunc, const avScore, const amHess)
{
    // Variaveis locais
    decl i, j, som;
    decl B = shape(vP[0:((k*(p-1))-1)], k, p-1);
    decl eta = mX*B;
    decl phi = vP[k*(p-1)];
    decl mi = zeros(ta,p);

    // Loop para preencher a matriz mi
    for (i = 0; i < ta; i++)
    {
        som = 0.0;
        mi[i][0] = exp(eta[i][0])/(1.0 + exp(eta[i][0]));
        for (j = 1; j < p-1; j++)
        {
            som = som + mi[i][j-1];
            mi[i][j] = (exp(eta[i][j])/(1.0 + exp(eta[i][j])))*(1.0 - som);
        }
        mi[i][p-1] = 1.0 - sumr(mi[i][0:(p-2)]);
    } // Fim do loop para preencher a matriz mi

    // Matriz alpha
    decl alpha = phi .* mi;

    adFunc[0] = double(sumc(loggamma(phi) - sumr(loggamma(alpha))
        + sumr((alpha - 1.0) .* log(mY))));

    if (isnan(adFunc[0]) || isdotinf(adFunc[0]))
        return 0;
    else
        return 1;    // 1 indica sucesso
} // Fim da funcao log-verossimilhanca do modelo1 de regressao Dirichlet

//=====

// Funcao para preencher a matriz inversa de informacao de Fisher do modelo 1
finfFishermodelo1(const vp5, const adFunc)
{
    decl r, m, t, P, K, Q, V1, V2, V3, V4, M2, M3,
```

```

m_inv_Inf_Fisch_B1, i, j, U, mIdent_Kronec_X1,
mInf_Fisch_B1, I_bb_inv, Sigma_inv, I_bb, I_phi_phi,
I_b_phi, I_phi_b, meval3, ab3, I_bb_aux, n, v,
I_b_phi_aux, I_phi_phi_aux, S1, S2, S3, S4, S5, S6,
S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18,
S19, S20, S21, S22, S23, P1, P2, P3, P4, P6, P7, P8,
P9, P10, P11, P12, SS4, SS44, SS5, SS13, SS15, SS17,
SS19, PS, PS1, PS3, PS4, PS5, PS6, PS8, PS9, PS10,
PS12, PS16, PS17, PS18, PS20, PS21, PS22, PS23, SP1,
SP2, SP3, SP4, PP1, PP2, PP7, PP10, S1_aux, S2_aux,
Sigma_inv_aux;

V1 = zeros(ta,1);
V2 = zeros(ta,1);
V3 = zeros(ta,1);
V4 = zeros(ta,1);
M3 = zeros(ta,p-1);
mInf_Fisch_B1 = zeros((k*(p-1))+1,(k*(p-1))+1);
I_bb_aux = zeros(k,k);
I_b_phi_aux = zeros(k,p-1);

for (n = 0; n < k; n++)
{
    for (v = 0; v < k; v++)
    {
        S1_aux = 0.0;
        for (i = 0; i < ta; i++)
        {
            S1_aux = S1_aux + (mX[i][v] * mX[i][n]*vp5^2*(polygamma(vp5*mi1[i][0],1) +
                polygamma(vp5*(1 - mi1[i][0]),1))*(gg1[i][0]^2));
        }
        I_bb_aux [n][v] = S1_aux;
    }
}

for (v = 0; v < k; v++)
{
    S2_aux = 0.0;
    for (i = 0; i < ta; i++)
    {
        S2_aux = S2_aux + (mX[i][v]*vp5*gg1[i][0]*(polygamma(vp5*mi1[i][0],1)
            * mi1[i][0] - polygamma(vp5*(1 - mi1[i][0]),1)*(1 - mi1[i][0])));
    }
    I_b_phi_aux[v][0] = S2_aux;
}

I_phi_phi_aux = 0.0;
for (i = 0; i < ta; i++)
{
    I_phi_phi_aux = I_phi_phi_aux + (polygamma(vp5 .* mi1[i][0],1)* mi1[i][0]^2 +
        polygamma(vp5*(1 - mi1[i][0]),1)*(1 - mi1[i][0])^2 -
        polygamma(vp5,1));
}

Sigma_inv_aux = I_bb_aux - (I_b_phi_aux * invert(I_phi_phi_aux) * I_b_phi_aux');

//***** Matriz M4 usada para calcular I_bb *****

// Preenchimento do U00,i
for (i = 0; i < ta; i++)
{
    S1 = 0.0;
    S2 = 0.0;
    if(p-2 >= 1)
    {
        for (j = 1; j <= p-2; j++)
        {
            if(j > 1)
            {
                PS1 = 1;
                for (t = 1; t <= j-1; t++)
                {
                    PS1 = PS1*(1.0 - gg[i][t]);
                }
            }
            else if(j == 1)
            {
                PS1 = 1;
            }
            S2 = S2 + (gg[i][j]*PS1);
            S1 = S1 + (polygamma(vp5*mi1[i][j],1) * gg[i][j]^2*PS1^2);
        }
    }
    else if(p-2 < 1)
    {
        S1 = 0.0;
        S2 = 0.0;
    }
}

```

```

        if(p > 2)
        {V1[i][0] = vp5^2*gg1[i][0]^2*(polygamma(vp5*mi1[i][0],1) + S1 +
            polygamma(vp5*(1.0 - sumr(mi1[i][0:p-2])),1)*(1.0 - S2)^2);}

        else if(p == 2)
        {V1[i][0] = vp5^2*gg1[i][0]^2*(polygamma(vp5*mi1[i][0],1) +
            polygamma(vp5*(1.0 - sumr(mi1[i][0:p-2])),1));}

    }//Fim do loop i e fim do preenchimento do U00,i

    M4 = diag(V1);

    // Preenchimento do U0r,i
    for (r = 1; r <= p-2; r++)
    { V2 = zeros(ta,1);
        for (i = 0; i < ta; i++)
        { S3 = 0.0;
            for (t = 0; t <= r-1; t++)
            {S3 = S3 + mi1[i][t];}
            P1 = 1;
            if(r-1 >= 1)
            {
                for (t = 1; t <= r-1; t++)
                {P1 = P1* (1.0 - gg[i][t]);}
            }
            else if(r-1 < 1)
            {P1 = 1;}

            //CALCULAR S4 E S6
            if(r < p-2)
            { S4 = 0.0;
                S6 = 0.0;
                if(r+1 <= p-2)
                {
                    for (j = r+1; j <= p-2; j++)
                    {
                        if(j == r+1)
                        {PS6 = 1;}
                        else if(j > r+1)
                        { PS6 = 1;
                            for (t = r+1; t <= j-1; t++)
                            {PS6 = PS6*(1.0 - gg[i][t]);}
                        }

                        PS4 = 1;
                        for (t = 1; t <= j-1; t++)
                        {PS4 = PS4*(1.0 - gg[i][t]);}
                        S6 = S6 + gg[i][j]*PS6;
                        S4 = S4 + polygamma(vp5*mi1[i][j],1) * gg[i][j]^2*PS4*PS6;
                    }
                }
                else if(r+1 > p-2)
                {S4 = 0.0; S6 = 0.0;}
            }
        }
    }

    else if(r == p-2)
    {S4 = 0.0; S6 = 0.0;}// Fim do else if(r == p-2)

    //CALCULAR S5
    if(1 <= p-2)
    { S5 = 0.0;
        for (j = 1; j <= p-2; j++)
        {
            if(j == 1)
            {PS5 = 1;}
            else if(j > 1)
            { PS5 = 1;
                for (t = 1; t <= j-1; t++)
                {PS5 = PS5*(1.0 - gg[i][t]);}
            }
            S5 = S5 + gg[i][j]*PS5;
        }
    }
    else if(1 > p-2)
    {S5 = 0.0;}

    V2[i][0] = (-vp5^2)*gg1[i][r]*gg1[i][0]*(1.0 - S3)*
        (polygamma(vp5*mi1[i][r],1)*gg[i][r]*P1 - S4 -

```



```

P = diag(V3);

for (r = 1; r <= p-2; r++)
{
    V4 = zeros(ta,1);
    // Preenchimento do Umr,i
    for (i = 0; i < ta; i++)
    {
        if(r == m)
        {
            //Calculo de S11
            S11 = 0.0;
            for (t = 0; t <= r-1; t++)
            {S11 = S11 + mi1[i][t];}

            //Calculo de S12 e S13
            if(r < p-2)
            {
                S12 = 0.0;
                S13 = 0.0;
                if(r+1 <= p-2)
                {
                    for (j = r+1; j <= p-2; j++)
                    {
                        if(j == r+1)
                        {PS12 = 1;}
                        else if(j > r+1)
                        { PS12 = 1;
                            for (t = r+1; t <= j-1; t++)
                            {PS12 = PS12*(1.0 - gg[i][t]);}
                        }

                        S13 = S13 + gg[i][j]*PS12;
                        S12 = S12 + polygamma(vp5*mi1[i][j],1) * gg[i][j]^2*PS12^2;
                    }

                    }
                else if(r+1 > p-2)
                {S13 = 0.0; S12 = 0.0;}
            }
            // Fim do if(r < p-2)
            else if(r == p-2)
            {S12 = 0.0; S13 = 0.0;}

            V4[i][0] = vp5^2*gg1[i][r]^2*(1.0 - S11)^2*(polygamma(vp5*mi1[i][r],1) +
                S12 + polygamma(vp5*(1.0 - sumr(mi1[i][0:p-2])),1)*(1.0 - S13)^2);

        }
        //Fim do if(r == m)
        else if(r > m)
        {
            //Calculo de S14
            S14 = 0.0;
            for (t = 0; t <= m-1; t++)
            {S14 = S14 + mi1[i][t];}

            //Calculo de S15
            S15 = 0.0;
            for (t = 0; t <= r-1; t++)
            {S15 = S15 + mi1[i][t];}

            //Calculo de S16 e S17
            if(r < p-2)
            {
                S16 = 0.0;
                S17 = 0.0;
                if(r+1 <= p-2)
                {
                    for (j = r+1; j <= p-2; j++)
                    {
                        if(j == r+1)
                        {PS17 = 1;}
                        else if(j > r+1)
                        { PS17 = 1;
                            for (t = r+1; t <= j-1; t++)
                            {PS17 = PS17*(1.0 - gg[i][t]);}
                        }

                        if(m+1 > j-1)
                        {PS16 = 1;}
                    }
                }
            }
        }
    }
}

```

```

        else if(m+1 <= j-1)
        { PS16 = 1;
          for (t = m+1; t <= j-1; t++)
            {PS16 = PS16*(1.0 - gg[i][t]);}
        }

        S17 = S17 + gg[i][j]*PS17;
        S16 = S16 + polygamma(vp5*mi1[i][j],1) * gg[i][j]^2*PS16*PS17;
    }

    }
    else if(r+1 > p-2)
    {S16 = 0.0; S17 = 0.0;}
} // Fim do if r < p-2
else if(r == p-2)
{S16 = 0.0; S17 = 0.0;}

//Calculo de S18
if(m+1 <= p-2)
{ S18 = 0.0;
  for (j = m+1; j <= p-2; j++)
  {
    if(j == m+1)
    {PS18 = 1;}
    else if(j > m+1)
    { PS18 = 1;
      for (t = m+1; t <= j-1; t++)
        {PS18 = PS18*(1.0 - gg[i][t]);}
    }
    S18 = S18 + gg[i][j]*PS18;
  }
}
else if(m+1 > p-2)
{S18 = 0.0;}

if(r-1 >= m+1)
{ P3 = 1;
  for (t = m+1; t <= r-1; t++)
    {P3 = P3*(1.0 - gg[i][t]);}
}
else if(r-1 < m+1)
{P3 = 1;}

V4[i][0] = (-vp5^2)*gg1[i][r]*gg1[i][m]*(1.0 - S14)*(1.0 - S15)*
(polygamma(vp5*mi1[i][r],1)*gg[i][r]*P3 - S16 -
polygamma(vp5*(1.0 - sumr(mi1[i][0:p-2])),1)*(1.0 - S17)*(1.0 - S18));

} //Fim do else if(r > m)

else if(r < m)
{
  //Calculo de S19
  S19 = 0.0;
  for (t = 0; t <= m-1; t++)
    {S19 = S19 + mi1[i][t];}

  //Calculo de S20
  S20 = 0.0;
  for (t = 0; t <= r-1; t++)
    {S20 = S20 + mi1[i][t];}

  //Calculo de S21 e S22
  if(m < p-2)
  {
    S21 = 0.0;
    S22 = 0.0;
    if(m+1 <= p-2)
    {
      for (j = m+1; j <= p-2; j++)
      {
        if(j == m+1)
        {PS22 = 1;}
        else if(j > m+1)
        { PS22 = 1;
          for (t = m+1; t <= j-1; t++)
            {PS22 = PS22*(1.0 - gg[i][t]);}
        }
      }
    }
  }
}

```

```

        if(r+1 > j-1)
        {PS21 = 1;}
        else if(r+1 <= j-1)
        { PS21 = 1;
          for (t = r+1; t <= j-1; t++)
            {PS21 = PS21*(1.0 - gg[i][t]);}
        }

        S22 = S22 + gg[i][j]*PS22;
        S21 = S21 + polygamma(vp5*mi1[i][j],1) * gg[i][j]^2*PS21*PS22;
    }

}
else if(m+1 > p-2)
{S21 = 0.0; S22 = 0.0;}
} // Fim do if m < p-2

else if(m == p-2)
{S21 = 0.0; S22 = 0.0;}

//Calculo de S23
if(r+1 <= p-2)
{ S23 = 0.0;
  for (j = r+1; j <= p-2; j++)
  {
    if(j == r+1)
    {PS23 = 1;}
    else if(j > r+1)
    { PS23 = 1;
      for (t = r+1; t <= j-1; t++)
        {PS23 = PS23*(1.0 - gg[i][t]);}
    }
    S23 = S23 + gg[i][j]*PS23;
  }
}
else if(r+1 > p-2)
{S23 = 0.0;}

if(m-1 >= r+1)
{ P4 = 1;
  for (t = r+1; t <= m-1; t++)
    {P4 = P3* (1.0 - gg[i][t]);}
}
else if(m-1 < r+1)
{P4 = 1;}

V4[i][0] = (-vp5^2)*gg1[i][r]*gg1[i][m]*(1.0 - S19)*(1.0 - S20)*
(polygamma(vp5*mi1[i][m],1)*gg[i][m]*P4 - S21 -
polygamma(vp5*(1.0 - sumr(mi1[i][0:p-2])),1)*(1.0 - S22)*(1.0 - S23));

} //Fim do else if(r < m)
} //Fim do loop i
Q = diag(V4);
P = P^Q;

} //Fim do loop r e fim do preenchimento do Umr,i
M4 = M4|P;
} //Fim do loop m

//***** Matriz M3 usada para calcular I_b_phi *****
for (i = 0; i < ta; i++)
{
  for (r = 0; r <= p-2; r++)
  {
    // Calcular SP1 e SP2
    if(r-1 >= 0)
    { SP1 = 0.0;
      for (t = 0; t <= r-1; t++)
        {SP1 = SP1 + mi1[i][t];}
    }
    else if(r-1 < 0)
    {SP1 = 0.0;}

    if(r < p-2)
    {
      SP3 = 0.0;
      if(p-2 >= r+1)

```

```

        {
            for (j = r+1; j <= p-2; j++)
            {
                if(j == r+1)
                {PS = 1;}
                else if(j > r+1)
                { PS = 1;
                 for (t = r+1; t <= j-1; t++)
                 {PS = PS*(1.0 - gg[i][t]);}
                }
                SP3 = SP3 + ((polygamma(vp5*mi1[i][j],1)* mi1[i][j] -
                    polygamma(vp5*(1.0 - sumr(mi1[i][0:p-2])),1)*
                    (1.0 - sumr(mi1[i][0:p-2])))* gg[i][j]*PS);
            }
            else if(p-2 < r+1)
            {SP3 = 0.0;}
        } // Fim do if(r < p-2)
        else if(r == p-2)
        {SP3 = 0.0;}

        M3[i][r] = gg1[i][r] * (1.0 - SP1) * vp5 *(polygamma(vp5*mi1[i][r],1) * mi1[i][r]
            - polygamma(vp5*(1 - sumr(mi1[i][0:p-2])),1) * (1 - sumr(mi1[i][0:p-2]))
            - SP3);

        } // Fim do loop do r
    } // Fim do loop do i

    M2 = mX'*M3;

    mIdent_Kronec_X1 = unit(p-1) ** mX;
    U = (mIdent_Kronec_X1')*M4*(mIdent_Kronec_X1);

    I_bb = U;
    I_phi_phi = - sumc(polygamma(vp5,1) - sumr((mi1.^2) .* polygamma(vp5 .* mi1,1)));
    I_b_phi = vec(M2);
    I_phi_b = vec(M2);
    Sigma_inv = I_bb - (I_b_phi * invert(I_phi_phi) * I_b_phi');

    adFunc[0] = Sigma_inv;

} // Fim da funcao para preencher a matriz inversa de informacao de Fisher do modelo 1

//=====

// Funcao para o teste sequencial
ftestemodolo1(const vp4, const adFunc)
{
    decl posto, V, D, U, S, w, svd, i, j, p_B, mu, mv, mw,
        v, min, U1, U2, V1, V2, D1, D2, Q_inv, l, n, L,
        quantil, nivel, posto_estimado;

    // decomposicao em valores singulares da matriz B estimada
    svd = decsvd(B_est1, &mu, &mw, &mv);
    V = mv;
    D = diag(mw);
    U = zeros(k,k);
    S = zeros(k,k);
    w = zeros(k,k);
    v = mu^(unit(k,k-(p-1)));

    // Loop para completar a matriz U
    // atraves do metodo de Gram-Schmidt
    for (j = 0; j < (p-1); j++)
    {U[j][j] = v[j][j];}
    for (i = p-1; i < k; i++)
    {
        for (j = 0; j < i; j++)
        {S[i][j] = S[i][j] + sumc(v[i][i] .* U[j][j])* U[j][j];}

        w[i] = v[i] - S[i];
        U[i][i] = w[i] ./ ((sumc(w[i].^2)).^0.5);
    } // Fim do loop para completar a matriz U

    // Loop para testar H0
    for (posto = 0; posto <= p-2; posto +=1)
    {

```

```

        if(posto == 0)
        { U2 = U; V2 = V; }

        else if(posto > 0)
        { U2 = U[posto:k-1];
          V2 = V[posto:p-2]; }

        Q_inv = (V2' ** U2') * vp4 * (V2 ** U2);
        l = vec(U2'*B_est1*V2);
        // a estatística de teste L(b)
        L = l'*Q_inv*l;
        // graus de liberdade para a qui-quadrado
        n = (k-posto)*((p-1)-posto);
        // quantil da qui-quadrado vezes a funcao C_N = sqrt(log(N))
        quantil = quanchi(0.95,n)*((log(ta))^0.5);

        // Aceita H0 se
        if(L <= quantil)
        {
            posto_estimado = posto;
            break;
        }

    } // Fim do loop para testar H0

    // Condicao para rejeitar H0
    if(L > quantil)
    {posto_estimado = p-1;} // Fim da condicao para rejeitar H0

    adFunc[0] = posto_estimado;
} // Fim da funcao para o teste sequencial

//=====

// Funcao para o criterio de informacao BIC
ftesteBIC1(const vp6, const adFunc)
{
    decl posto, V, D, U, S, w, svd, i, j, p_B, mu, mv, mw,
        v, U1, U2, V1, V2, D1, D2, Q_inv, l, n, L,
        BIC, posto_estimado;

    // decomposicao em valores singulares da matriz B estimada
    svd = decsvd(B_est1, &mu, &mw, &mv);
    V = mv;
    D = diag(mw);
    U = zeros(k,k);
    S = zeros(k,k);
    w = zeros(k,k);
    v = mu*(unit(k,k-(p-1)));
    BIC = zeros(1,p);
    L = zeros(1,p-1);

    // Loop para completar a matriz U
    // atraves do metodo de Gram-Schmidt
    for (j = 0; j < p-1; j++)
    {U[j][j] = v[j][j];}
    for (i = p-1; i < k; i++)
    {
        for (j = 0; j < i; j++)
        {S[i][i] = S[i][i] + sumc(v[i][i] * U[j][j]) * U[j][j];}

        w[i][i] = v[i][i] - S[i][i];
        U[i][i] = w[i][i] ./ ((sumc(w[i][i]^2))^0.5);
    } // Fim do loop para completar a matriz U

    // Loop para o calculo do vetor BIC
    for (posto = 0; posto <= p-2; posto+=1)
    {
        if(posto == 0)
        {U2 = U; V2 = V;}

        else if(posto > 0)
        {U2 = U[posto:k-1]; V2 = V[posto:(p-1)-1]; }
    }
}

```

```

        Q_inv = (V2' ** U2') * vp6 * (V2 ** U2);
        l = vec(U2'*B_est1*V2);
        // a estatística de teste L(b)
        L[posto] = l'*Q_inv*l;
        // vetor IC(b, C_N)
        BIC[posto] = (L[posto] + (posto)*log(ta));

    } // Fim do loop para o calculo do vetor BIC

    // se b = p-1 entao L(b) = 0
    BIC[p-1] = (p-1)*log(ta);

    posto_estimado = (vecindex(BIC, min(BIC))); // o minimo do vetor BIC

    adFunc[0] = posto_estimado;

} // Fim da funcao para o criterio de informacao BIC

//=====

// Funcao para o teste HQIC
ftesteHQIC1(const vp7, const adFunc)
{
    decl posto, V, D, U, S, w, svd, i, j, p_B, mu, mv, mw,
        v, U1, U2, V1, V2, D1, D2, Q_inv, l, n, L,
        HQIC, posto_estimado, const_c;

    // decomposicao em valores singulares da matriz B estimada
    svd = decsvd(B_est1, &mu, &mw, &mv);
    V = mv;
    D = diag(mw);
    U = zeros(k,k);
    S = zeros(k,k);
    w = zeros(k,k);
    v = mu*(unit(k,k-(p-1)));
    HQIC = zeros(1,p);
    L = zeros(1,p-1);

    const_c = 2.01;

    // Loop para completar a matriz U
    // atraves do metodo de Gram-Schmidt
    for (j = 0; j < p-1; j++)
    {U[j] = v[j];}
    for (i = p-1; i < k; i++)
    {
        for (j = 0; j < i; j++)
        {S[i] = S[i] + sumc(v[i].* U[j])* U[j];}

        w[i] = v[i] - S[i];
        U[i] = w[i] ./ ((sumc(w[i].^2)).^0.5);
    } // Fim do loop para completar a matriz U

    // Loop para o calculo do vetor HQIC
    for (posto = 0; posto <= p-2; posto+=1)
    {
        if(posto == 0)
        {U2 = U; V2 = V;}

        else if(posto > 0)
        {U2 = U[posto:k-1]; V2 = V[posto:(p-1)-1];}

        Q_inv = (V2' ** U2') * vp7 * (V2 ** U2);
        l = vec(U2'*B_est1*V2);
        // a estatística de teste L(b)
        L[posto] = l'*Q_inv*l;
        // vetor IC(b, C_N)
        HQIC[posto] = (L[posto] + (posto)*(const_c*(log(log(ta)))));
    } // Fim do loop para o calculo do vetor HQIC

    // se b = p-1 entao L(b) = 0
    HQIC[p-1] = (p-1)*const_c*(log(log(ta)));

    posto_estimado = (vecindex(HQIC, min(HQIC))); // o minimo do vetor HQIC

```

```

        adFunc[0] = posto_estimado;
} // Fim da funcao para o criterio de informacao HIC

```

● Modelo 2

```

/*****
* PROGRAMA: funcoes2.ox
* AUTORA: Tatiane Ferreira do Nascimento Melo
* DATA: 27 de fevereiro de 2004
* ULTIMA MODIFICACAO: 30 de agosto de 2004   Hora: 19:00
*****/

/* Arquivos principais */
#include <oxstd.h>
#include <oxfloat.h>
#include <oxprob.h>
#import <maximize>
#import <solvenle>

// Variaveis Globais
static decl mX, mY, m_mi, m_eta, k, p, ta, M;
static decl B_est, m_eta_est, m_g_eta_est, m_mi_est, g_est, g1_est;

//=====

// Funcao log-verossimilhanca do modelo2 de regressao Dirichlet
floglik2(const vP, const adFunc, const avScore, const amHess)
{
    // Variveis locais
    decl i, j, B, eta, phi2, alpha, mi, g_eta;

    B = shape(vP[0:((k*p)-1)], k, p);
    eta = mX*B;
    g_eta = exp(eta);
    phi2 = sumr(g_eta);
    //matriz mi
    mi = g_eta ./ phi2;
    //matriz alpha
    alpha = g_eta;

    adFunc[0] = double(sumc(loggamma(phi2) - sumr(loggamma(alpha)) +
        sumr((alpha - 1.0) .* log(mY))));

    if (isnan(adFunc[0]) || isdotinf(adFunc[0]))
        return 0;
    else
        return 1;    // 1 indica sucesso
} // Fim da funcao log-verossimilhanca do modelo2 de regressao Dirichlet

//=====

// Funcao para preencher a matriz inversa de informacao de Fisher
finfFishermodelo2(const phi, const adFunc)
{
    // Variaveis locais
    decl i, j, r, m, V1, V2, V3, V4, P, Q, mIdent_Kronec_X, mInf_Fisch_B_est;

    V1 = zeros(ta,1);
    for (i = 0; i < ta; i++)
        {V1[i][0] = - g1_est[i][0]^2 * (polygamma(phi[i][0],1)
            - polygamma(phi[i][0]*m_mi_est[i][0], 1));} // U00

```

```

M = diag(V1);

for (r = 1; r <= p-1; r++)
{
    V2 = zeros(ta,1);
    for (i = 0; i < ta; i++)
        {V2[i][0] = -g1_est[i][r]*g1_est[i][0]*polygamma(phi[i][0],1);} // U0r

    M = M ~ diag(V2);
} // Fim do loop do r

for (m = 1; m <= p-1; m++)
{
    V3 = zeros(ta,1);
    for (i = 0; i < ta; i++)
        {V3[i][0] = -g1_est[i][m]*g1_est[i][0]*polygamma(phi[i][0],1);} // Um0

    P = diag(V3);

    for (r = 1; r <= p-1; r++)
    {
        V4 = zeros(ta,1);
        for (i = 0; i < ta; i++)
        {
            if(r == m)
                {V4[i][0] = -g1_est[i][m]^2*(polygamma(phi[i][0],1)
                    - polygamma(phi[i][0]*m_mi_est[i][m], 1));} // Umm

            else
                {V4[i][0] = -g1_est[i][r]*g1_est[i][m]*polygamma(phi[i][0],1);} // Umr

        }
    } // Fim do loop do i

    Q = diag(V4);
    P = P ~ Q;

    } // Fim do loop do r

    M = M | P;

} // Fim do loop do m

mIdent_Kronec_X = unit(p) ** mX;
mInf_Fisch_B_est = ((mIdent_Kronec_X)')*M*(mIdent_Kronec_X);

adFunc[0] = mInf_Fisch_B_est;

return 1; // 1 indica sucesso

} // Fim da funcao para preencher a matriz inversa de informacao de Fisher

//=====

// Funcao para o teste sequencial
ftestemodelo2(const vp4, const adFunc)
{
    // Variaveis locais
    decl posto, V, D, U, S, w, svd, i, j, mu, mv, mw,
        v, U1, U2, V1, V2, Q_inv, l, n, L, quantil,
        posto_estimado;

    // decomposicao em valores singulares a matriz B estimada
    svd = decsvd(B_est, &mu, &mw, &mv);
    V = mv;
    D = diag(mw);
    U = zeros(k,k);
    S = zeros(k,k);
    w = zeros(k,k);
    v = mu~(unit(k,k-p));

    // Loop para completar a matriz U
    // atraves do metodo de Gram-Schmidt
    for (j = 0; j < p; j++)
        {U[j][j] = v[j][j];}

```

```

for (i = p; i < k; i++)
{
    for (j = 0; j < i; j++)
    {S[i] = S[i] + sumc(v[i] .* U[j]) * U[j];}

    w[i] = v[i] - S[i];
    U[i] = w[i] ./ ((sumc(w[i].^2)).^0.5);

} // Fim do loop para completar a matriz U

// Loop para testar H0
for (posto = 0; posto <= p-1; posto++)
{
    if(posto == 0)
    { U2 = U; V2 = V;}
    else
    {
        U1 = U[0:posto-1];
        U2 = U[posto:k-1];
        V1 = V[0:posto-1];
        V2 = V[posto:p-1];
    }

    Q_inv = (V2' ** U2') * vp4 * (V2 ** U2);
    l = vec(U2'*B_est*V2);
    // a estatística de teste L(b)
    L = l'*Q_inv*l;
    // graus de liberdade para a qui-quadrado
    n = (k-posto)*(p-posto);
    // quantil da qui-quadrado vezes a função C_N = sqrt(log(N))
    quantil = quanchi(0.95,n)*((log(ta))^0.5);

    // Aceita H0 se
    if(L <= quantil)
    {
        posto_estimado = posto;
        break;
    }
} // Fim do loop para testar H0

// Condio para rejeitar H0
if(L > quantil)
{posto_estimado = p;} // Fim da condio para rejeitar H0

adFunc[0] = posto_estimado;

} // Fim da função para o teste sequencial

//=====

// Função para o critério de informação BIC
ftesteBIC2(const vp5, const adFunc)
{
    // Variáveis locais
    decl posto, V, D, U, S, w, svd, i, j, mu, mv, mw, v, U1,
        U2, V1, V2, Q_inv, l, L, BIC, posto_estimado;

    // decomposição em valores singulares da matriz B estimada
    svd = decsvd(B_est, &mu, &mw, &mv);
    V = mv;
    D = diag(mw);
    U = zeros(k,k);
    S = zeros(k,k);
    w = zeros(k,k);
    v = mu~(unit(k,k-p));
    BIC = zeros(1,p+1);
    L = zeros(1,p);

    // Loop para completar a matriz U
    // através do método de Gram-Schmidt
    for (j = 0; j < p; j++)
    {U[j] = v[j];}

    for (i = p; i < k; i++)
    {

```

```

    for (j = 0; j < i; j++)
    {S[i] = S[i] + sumc(v[i].* U[j])* U[j];}

    w[i] = v[i] - S[i];
    U[i] = w[i] ./ ((sumc(w[i].^2)).^0.5);

} // Fim do loop para completar a matriz U

// Loop para testar H0
for (posto = 0; posto <= p-1; posto++)
{
    if(posto == 0)
    {U2 = U; V2 = V;}

    else if(posto >= 1)
    {
        U1 = U[0:posto-1];
        U2 = U[posto:k-1];
        V1 = V[0:posto-1];
        V2 = V[posto:p-1];
    }

    Q_inv = (V2' ** U2') * vp5 * (V2 ** U2);
    l = vec(U2'*B_est*V2);
    // a estatística de teste L(b)
    L[posto] = l'*Q_inv*l;
    // vetor IC(b, C_N)
    BIC[posto] = L[posto] + posto*log(ta);

} // Fim do loop para testar H0

// se b = p então L(b) = 0
BIC[p] = p*log(ta);

posto_estimado = (vecindex(BIC, min(BIC))); // o mínimo do vetor BIC

adFunc[0] = posto_estimado;

} // Fim da função para o critério de informação BIC

//=====

// Função para o critério de informação HQIC
ftesteHQIC2(const vp7, const adFunc)
{
    // Variáveis locais
    decl posto, V, D, U, S, w, svd, i, j, mu, mv, mw,
    v, U1, U2, V1, V2, Q_inv, l, L, HQIC,
    posto_estimado, const_c;

    // decomposição em valores singulares da matriz B estimada
    svd = decsvd(B_est, &mu, &mw, &mv);
    V = mv;
    D = diag(mw);
    U = zeros(k,k);
    S = zeros(k,k);
    w = zeros(k,k);
    v = mu~(unit(k,k-p));
    HQIC = zeros(1,p+1);
    L = zeros(1,p);

    const_c = 2.01;

    // Loop para completar a matriz U
    // através do método de Gram-Schmidt
    for (j = 0; j < p; j++)
    {U[j] = v[j];}

    for (i = p; i < k; i++)
    {
        for (j = 0; j < i; j++)
        {S[i] = S[i] + sumc(v[i].* U[j])* U[j];}

        w[i] = v[i] - S[i];
        U[i] = w[i] ./ ((sumc(w[i].^2)).^0.5);
    }
}

```

```

} // Fim do loop para completar a matriz U

// Loop para testar H0
for (posto = 0; posto <= p-1; posto++)
{
    if(posto == 0)
    {U2 = U; V2 = V;}

    else if(posto >= 1)
    {
        U1 = U[] [0:posto-1];
        U2 = U[] [posto:k-1];
        V1 = V[] [0:posto-1];
        V2 = V[] [posto:p-1];
    }

    Q_inv = (V2' ** U2') * vp7 * (V2 ** U2);
    l = vec(U2'*B_est*V2);
    // a estatística de teste L(b)
    L[posto] = l'*Q_inv*l;
    // vetor IC(b, C_N)
    HQIC[posto] = L[posto] + posto*const_c*(log(log(ta)));

} // Fim do loop para testar H0

// se b = p então L(b) = 0
HQIC[p] = p*const_c*(log(log(ta)));

posto_estimado = (vecindex(HQIC, min(HQIC))); // o mínimo do vetor HQIC

adFunc[0] = posto_estimado;

} // Fim da função para o critério de informação HQIC

```

A5.3. Programas de ligação

• Modelo 1

```
/******  
* PROGRAMA: funcoes1.h  
* AUTORA: Tatiane Ferreira do Nascimento Melo  
* DATA: 27 de fevereiro de 2004  
* ULTIMA MODIFICACAO: 30 de agosto de 2004   Hora: 19:00  
*****/  
  
// Funcao log-verossimilhanca no modelo 1 de regressao Dirichlet  
floglik1(const vP, const adFunc, const avScore, const amHess);  
  
// Funcao para preencher a matriz inversa de informacao de Fisher no modelo 1  
finfFishermodelo1(const vp5, const adFunc);  
  
// Funcao para o teste sequencial  
ftestemodelo1(const vp4, const adFunc);  
  
// Funcao para o criterio de informacao BIC  
ftesteBIC1(const vp6, const adFunc);  
  
// Funcao para o criterio de informacao HQIC  
ftesteHQIC1(const vp7, const adFunc);
```

• Modelo 2

```
/******  
* PROGRAMA: funcoes2.h  
* AUTORA: Tatiane Ferreira do Nascimento Melo  
* DATA: 27 de fevereiro de 2004  
* ULTIMA MODIFICACAO: 30 de agosto de 2004   Hora: 19:00  
*****/  
  
// Funcao log-verossimilhanca no modelo 2 de regressao Dirichlet  
floglik2(const vP, const adFunc, const avScore, const amHess);  
  
// Funcao para preencher a matriz de informacao de Fisher no modelo 2  
finfFishermodelo2(const phi, const adFunc);  
  
// Funcao para o teste sequencial  
ftestemodelo2(const vp4, const adFunc);  
  
// Funcao para o criterio de informacao BIC  
ftesteBIC2(const vp5, const adFunc);  
  
// Funcao para o o criterio de informacao HQIC  
ftesteHQIC2(const vp7, const adFunc);
```

Referências

- [1] Aitchison, J. (2003). *The Statistical Analysis of Compositional Data*. New Jersey: The Blackburn Press.
- [2] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions in Automatic Control*, **AC-19**, 716–723.
- [3] Anderson, T. W. & Kunitomo N. (1994). Asymptotic robustness of tests of overidentification and predeterminedness. *Journal of Econometrics*, **62**, 383–414.
- [4] Bury, K. (1999). *Statistical Distributions in Engineering*. New York: Cambridge University Press.
- [5] Cordeiro G. M. (1992). *Introdução à Teoria de Verossimilhança*. Livro Texto do 10^o Simpósio Nacional de Probabilidade e Estatística, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- [6] Cordeiro G. M. (1999). *Introdução à Teoria Assintótica*. Livro Texto do 22^o Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro.
- [7] Cragg J. G. & Donald S. G. (1996). On the asymptotic properties of LDU based tests for the rank of a matrix. *Journal of the American Statistical Association*, **91**, 1301–1309.
- [8] Cragg J. G. & Donald S. G. (1997). Inferring the rank of matrix. *Journal of Econometrics*, **76**, 223–250.
- [9] Cribari-Neto, F. & Zarkos, S. G. (1999). R: yet another econometric programming environment. *Journal of Applied Econometrics*, **14**, 319–329.
- [10] Cribari-Neto, F. & Zarkos, S. G. (2003). Econometric and statistical computing using Ox. *Computational Economics*, **21**, 277–295.
- [11] Doornik, J. A. (2001). *Ox: an Object-oriented Matrix Programming Language*, 4th ed. London: Timberlake Consultants & Oxford: <http://www.nuff.ox.ac.uk/Users/Doornik>.
- [12] Ferrari, S.L.P. & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, **31(7)**, 799–815.
- [13] Gill L. & Lewbel A. (1992). Testing the rank and definiteness of estimated matrices with applications to factor, state space and ARMA models. *Journal of the American Statistical Association*, **87**, 766–776.
- [14] Golub, G. H. & Loan, C. F. V. (1996). *Matrix Computations*. 3^a ed. Johns Hopkins University Press. Baltimore.
- [15] Good, I. J. (1965). Estimation of probabilities. *Cambridge*. Mass: The M. I. T. Press.

- [16] Hannan, E. J. & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of Royal Statistical Society Series B*, **41**, 190–195.
- [17] Hsiao C. (1986). Identification . *Handbook of Econometrics*, Volume 1, eds Z. Griliches and M.D Intrilligator. Amsterdam: North-Holland.
- [18] Janardan, K. G. & Padmanabhan, G. (1986). Double bounded beta distribution for hydrologic variables. *Proc. 17th Annual Pittsburgh Conference* (parte 3), **17**, 1107–1111.
- [19] Kato T. (1982). *A Short Introduction to Perturbation Theory for Linear Operators*. Springer Verlag, New York.
- [20] Lewbel A. (1991). The rank of demand systems: theory and nonparametric estimation. *Econometrica*, **59**, 711–730.
- [21] Lewbel A. & Perraudin (1995). A Theorem on portfolio separation with general preferences. *Journal of Economic Theory*, **65**, 624–626.
- [22] Mauldon, J. G. (1951). Random division of an interval. *Proceedings of the Cambridge Philosophical Society*, **47**, 331–336.
- [23] Rao C. R. (1973). *Linear Statistical Inference and Its Applications*. 2^a ed. New York: John Wiley and Sons.
- [24] Ratsimalahelo Z. (2003). Strongly consistent determination of the rank of matrix, artigo ewp-em/0307007, disponível em <http://econwpa.wustl.edu/eprints/em/papers/0307/0307007.abs>.
- [25] Robin J-M & Smith R. J. (2000). Tests of rank. *Econometric Theory*, **16**, 151–175.
- [26] Schwarz G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.
- [27] Stewart G. W.. (1984). Rank degeneracy. *SIAM Journal on Scientific and Statistical Computing*, **5**, 403–413.
- [28] Stewart G. W. & Sun J. G. (1990). *Matrix Perturbation Theory*, Boston, Academic Press.