



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
CURSO DE GRADUAÇÃO EM ESTATÍSTICA
DISCENTE: DOUGLAS CHAVES MOURA
DOCENTE: PROF. DR. JUVÊNCIO SANTOS NOBRE

EEGs em modelos de regressão beta para dados de
medidas repetidas - ANÁLISE DE DADOS
LONGITUDINAIS (CC0300)

CAPITAL DO CEARÁ

2025

Sumário

1	Introdução	4
2	Distribuição Beta	4
3	Regressão Beta	6
4	Equações de Estimação Generalizadas (EEGs)	9
4.1	Formulação das Equações de Estimação	9
4.2	Estrutura de Covariância de Trabalho	9
4.3	Estimação dos Parâmetros	10
4.4	Abordagens EEG1 vs. EEG2	10
5	Aplicação das EEGs em Modelos de Regressão Beta para Medidas Repetidas	10
5.1	Modelagem do Parâmetro de Posição	11
5.1.1	Estimação de β , ϕ e α	13
5.2	Modelagem Conjunta dos Parâmetros de Posição e Precisão	15
5.2.1	Estimação de β , γ e α	17
6	Análise de Diagnóstico Para Dados de Medidas Repetidas	18
6.1	Processo Iterativo Reponderado	19
6.2	Ponto Alavanca Baseado na Matriz de Projecção.	19
6.3	Ponto Aberrante Baseado no Resíduo Padronizado	20
6.4	Ponto Influente Baseado na Distância de Cook	20
7	Seleção de Modelos e de Matriz de Correlação de Trabalho	21
8	Influência Local para Equações de Estimação	22
8.1	Influência Local Generalizada	23
8.2	Influência Local para Equações de Estimação	24
8.3	Esquemas de Perturbação sob Homogeneidade da Dispersão	25
8.3.1	Ponderação de Casos	25
8.3.2	Perturbação da Variável Resposta	26
8.3.3	Perturbação Individual das Covariáveis	26
8.3.4	Perturbação do Parâmetro de Precisão	28
8.3.5	Perturbação na Matriz de Correlação de Trabalho $\mathbf{R}(\alpha)$	29
8.4	Esquemas de Perturbação sob Heterogeneidade da Dispersão	30
8.4.1	Ponderação de Casos	30
8.4.2	Perturbação da Variável Resposta	31
8.4.3	Perturbação Individual das Covariáveis	31
8.4.4	Matriz \mathbf{Q} Totalmente Diferente da Matriz \mathbf{X}	32
8.4.5	Matriz \mathbf{Q} Totalmente Igual à Matriz \mathbf{X}	33
8.4.6	A k' -ésima coluna da matriz \mathbf{Q} igual à k -ésima coluna da matriz \mathbf{X}	34
8.4.7	A k' -ésima coluna da matriz \mathbf{Q} função da k -ésima coluna da matriz \mathbf{X}	35

Lista de Figuras

1	Funções densidade de probabilidade e distribuição acumulada da distribuição Beta para diferentes parametrizações de p e q	5
2	Efeito do aumento da precisão (ϕ) na forma da distribuição Beta reparametrizada para diferentes valores de μ	7

1 Introdução

Estudos longitudinais – ou estudos com medidas repetidas – são caracterizados pela coleta sequencial de dados em múltiplos momentos ao longo do tempo para os mesmos indivíduos ou unidades amostrais. Essa abordagem possibilita o acompanhamento da evolução dos sujeitos, fornecendo uma visão dinâmica dos fenômenos, mas também impõe desafios analíticos importantes, pois as observações realizadas em momentos distintos para o mesmo indivíduo geralmente exibem correlação intrínseca.

A não aleatorização na ordem das observações, inerente ao desenho longitudinal, impede a aplicação direta de modelos de regressão clássicos que assumem independência entre as observações. Ademais, essa modalidade de estudo pode apresentar problemas práticos, como espaçamentos irregulares entre as medições, ausência de dados em determinados momentos (dados censurados ou faltantes) e outros “inconvenientes” relacionados à coleta ao longo do tempo.

Para lidar com a dependência entre as observações e possibilitar a análise adequada dos dados, as Equações de Estimação Generalizadas (EEGs) de Liang e Zeger (1986) surgiram como uma ferramenta robusta. Essa abordagem baseia-se na utilização de modelos de quasi-verossimilhança, conforme discutido por Wedderburn (1974), e fundamenta-se na especificação das distribuições marginais da variável resposta, seguindo o paradigma dos Modelos Lineares Generalizados (NELDER; WEDDERBURN, 1972). Dessa forma, não é necessário conhecer a distribuição conjunta dos dados, focando-se apenas nas características marginais, o que é particularmente útil quando se trabalha com dados correlacionados.

Quando a variável resposta assume valores contínuos restringidos ao intervalo $(0, 1)$ – como proporções, taxas e índices percentuais –, a literatura tem adotado frequentemente o modelo de regressão beta proposto por Ferrari e Cribari-Neto (2004). Nesse contexto, os autores introduziram uma nova parametrização da distribuição beta, na qual a média é modelada por meio de um componente sistemático e um parâmetro adicional, denominado parâmetro de precisão. Este, diferentemente do parâmetro de dispersão encontrado em outros modelos, permite capturar com maior acurácia a variabilidade esperada para respostas que naturalmente se restringem ao intervalo $(0, 1)$.

Portanto, o presente relatório tem como objetivo explorar a aplicação das Equações de Estimação Generalizadas em modelos de regressão beta voltados à análise de dados longitudinais. Para tanto, o trabalho está estruturado de modo a, inicialmente, apresentar a distribuição beta (Seção 2), elemento fundamental para compreender a regressão beta, cuja abordagem detalhada ocorrerá na Seção 3. Em seguida, serão discutidos os princípios teóricos das EEGs, enfatizando suas limitações, desafios e as vantagens que essa metodologia oferece para a modelagem de variáveis contínuas restritas ao intervalo $(0, 1)$.

2 Distribuição Beta

A distribuição Beta é parte fundamental do pressuposto do modelo desenvolvido por Ferrari e Cribari-Neto (2004). Ela é bastante flexível para a modelagem de proporções, pois sua densidade pode ser ajustada de acordo com os parâmetros associados à distribuição. A sua função densidade é dada por:

$$f(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1} (1-y)^{q-1}, \quad 0 < y < 1$$

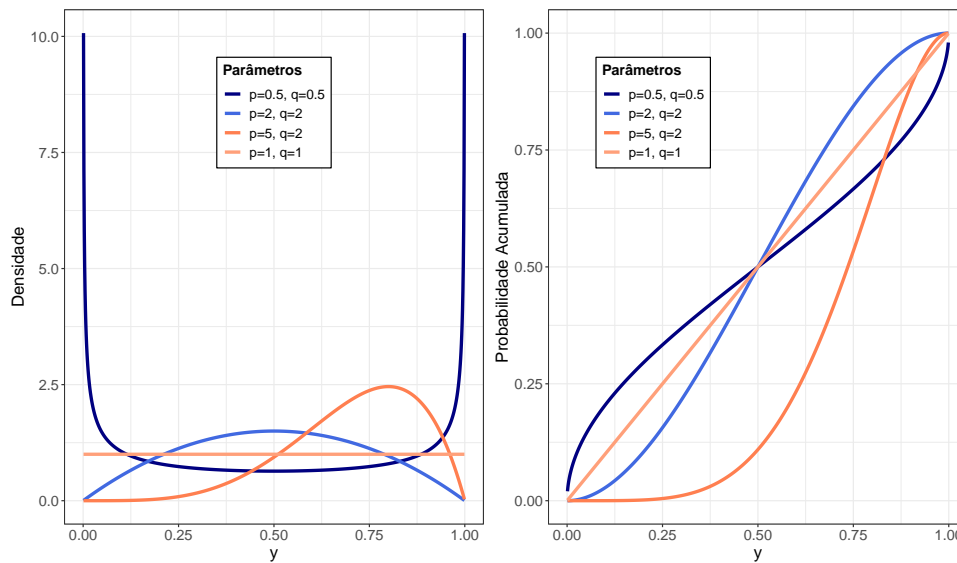
onde $p > 0$, $q > 0$, e $\Gamma(\cdot)$ representa a função gama:

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

sendo z um número complexo com parte real positiva. Para números naturais, $\Gamma(z) = (z - 1)!$, estendendo o conceito de fatorial.

Valores maiores de p resultam em maior concentração dos valores de y próximos a 1. Se p for pequeno a distribuição terá maior concentração perto de 0. O mesmo ocorre de forma inversa para o parâmetro q . Quando $p = q$, a distribuição é simétrica.

Figura 1: Funções densidade de probabilidade e distribuição acumulada da distribuição Beta para diferentes parametrizações de p e q .



Casos especiais

- **Distribuição Uniforme:** Quando $p = q = 1$, a função densidade de probabilidade da distribuição Beta torna-se constante em todo o intervalo $(0, 1)$.
- **Distribuição U-Shaped:** Quando $p, q < 1$, a densidade apresenta maior concentração nas extremidades, resultando em uma forma de “U”.
- **Distribuição Concentrada no Centro:** Quando $p, q > 1$, a densidade apresenta um pico na região central, com valores mais frequentes próximos de $y = \frac{p-1}{p+q-2}$.

A média e a variância de y são, respectivamente,

$$E(y) = \frac{p}{p + q} \tag{2.1}$$

e

$$\text{Var}(y) = \frac{pq}{(p + q)^2(p + q + 1)}. \tag{2.2}$$

As distribuições Beta são altamente versáteis, permitindo a modelagem de uma ampla variedade de incertezas de forma eficaz. Essa flexibilidade incentiva sua utilização empírica

em diversas aplicações práticas. Diversas aplicações da distribuição Beta são discutidas por Bury (1999) e por Johnson, Kotz e Balakrishnan (1995). No entanto, essas aplicações geralmente não exigem que o profissional imponha uma estrutura de regressão para a variável de interesse. A proposta de Ferrari e Cribari-Neto (2004) é investigar o comportamento de uma variável que segue uma distribuição Beta em relação a outras variáveis explicativas, permitindo assim a construção de modelos preditivos mais 'simples' e adequados para dados restritos ao intervalo unitário.

3 Regressão Beta

A distribuição Beta tradicional é parametrizada por dois parâmetros de forma, p e q . Ferrari e Cribari-Neto (2004) introduziram uma nova parametrização para essa distribuição, incorporando um parâmetro de posição (μ) e um parâmetro de precisão (ϕ). Esses parâmetros estão relacionados a p e q pelas seguintes equações:

$$\mu = \frac{p}{(p+q)}, \quad \phi = p+q.$$

Com isso, podemos expressar p e q em função de μ e ϕ :

$$p = \mu\phi \quad \text{e} \quad q = (1-\mu)\phi.$$

A nova parametrização permite expressar a média e a variância de \mathbf{y} de forma mais intuitiva:

$$E(\mathbf{y}) = \mu,$$
$$\text{Var}(\mathbf{y}) = \frac{\mu(1-\mu)}{1+\phi}.$$

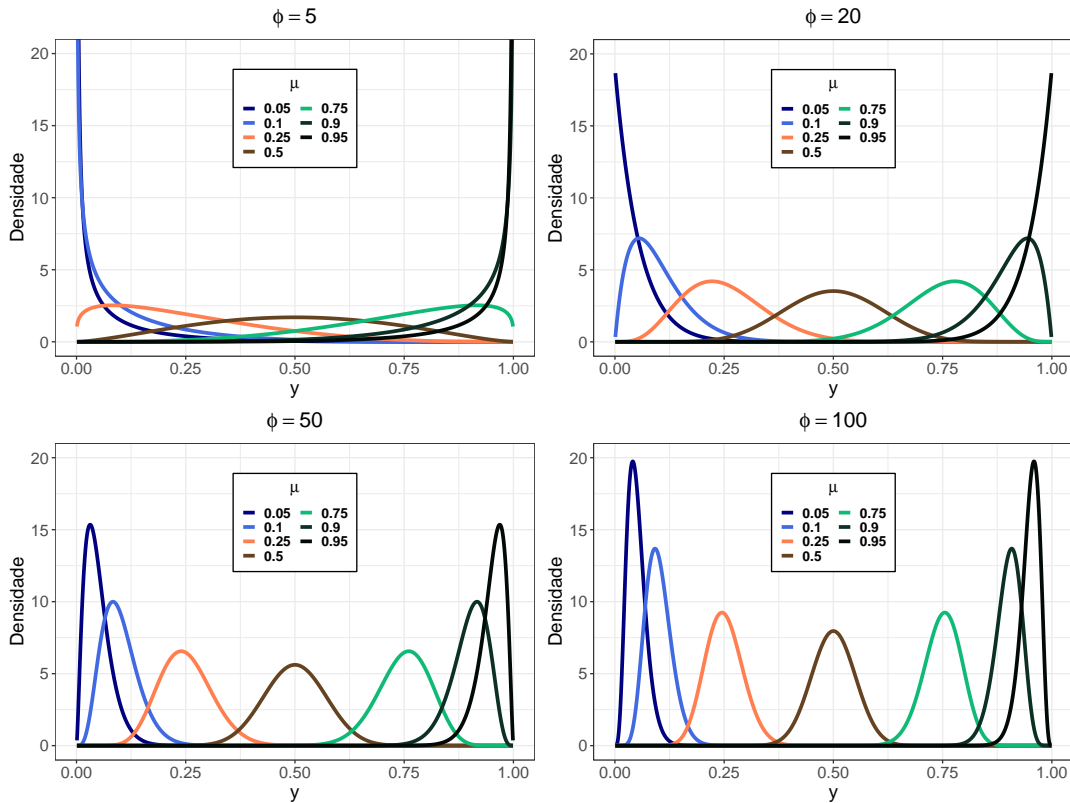
Aqui, ϕ controla a dispersão dos dados: valores maiores de ϕ reduzem a variância, concentrando os valores em torno de μ .

A densidade de probabilidade da distribuição Beta, na nova parametrização, é dada por:

$$f(\mathbf{y}; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1. \quad (3.1)$$

Para evitar problemas com valores extremos (próximos de 0 ou 1), Cribari-Neto e Zeileis (2010) sugerem a transformação $\mathbf{y}^* = (y \cdot (n-1) + 0.5)/n$, onde n é o tamanho da amostra (SMITHSON; VERKUILEN, 2006).

Figura 2: Efeito do aumento da precisão (ϕ) na forma da distribuição Beta reparametrizada para diferentes valores de μ .



NOTA: Imagem adaptada de Ferrari e Cribari-Neto (2004).

A regressão beta modela a média μ_t da variável resposta y_t como uma combinação linear de covariáveis, utilizando uma função de ligação $g(\cdot)$:

$$g(\mu_t) = \sum_{i=1}^k x_{ti}\beta_i = \eta_t, \quad (3.2)$$

onde η_t é o preditor linear. A função $g(\cdot)$ deve ser monotônica e diferenciável. As funções mais comuns incluem:

Ligação Logito

$$g(\mu_t) = \log\left(\frac{\mu_t}{1 - \mu_t}\right), \quad \mu_t = \frac{\exp(\mathbf{x}_t^\top \beta)}{1 + \exp(\mathbf{x}_t^\top \beta)}.$$

Ideal para proporções, com interpretação direta em termos de razão de chances (COLLETT, 2003).

Ligação Probit

$$g(\mu_t) = \Phi^{-1}(\mu_t), \quad \mu_t = \Phi(\mathbf{x}_t^\top \beta),$$

onde $\Phi(\cdot)$ é a função de distribuição normal padrão. Útil para dados com comportamento semelhante ao modelo probito clássico.

Outras ligações são: *cauchit*, *log*, *log-log* e *complementar log-log*, cada uma com interpretações específicas para diferentes contextos (COX; SNELL, 1989).

O logaritmo da função de verossimilhança para n observações independentes é dado por:

$$\ell(\beta, \phi) = \sum_{t=1}^n [\log \Gamma(\phi) - \log \Gamma(\mu_t \phi) - \log \Gamma((1 - \mu_t)\phi) + (\mu_t \phi - 1) \log y_t + \{(1 - \mu_t)\phi - 1\} \log(1 - y_t)] \quad (3.3)$$

Essa função é maximizada para obter os estimadores de máxima verossimilhança (EMV) de β e ϕ .

Os EMV são obtidos resolvendo as equações de verossimilhança:

$$U_\beta(\beta, \phi) = \phi \mathbf{X}^\top \mathbf{T}(\mathbf{y}^* - \mu^*) = 0,$$

$$U_\phi(\beta, \phi) = \sum_{t=1}^n [\mu_t(y_t^* - \mu_t^*) + \log(1 - y_t) - \psi((1 - \mu_t)\phi) + \psi(\phi)] = 0,$$

onde $y_t^* = \log(y_t/(1 - y_t))$, $\mu_t^* = \psi(\mu_t \phi) - \psi((1 - \mu_t)\phi)$, e $\psi(\cdot)$ é a função digama. A solução requer métodos numéricos como Newton-Raphson (FERRARI; CRIBARI-NETO, 2004).

A matriz de informação de Fisher para o modelo é:

$$\mathbf{K}(\beta, \phi) = \begin{pmatrix} \phi \mathbf{X}^\top \mathbf{W} \mathbf{X} & \mathbf{X}^\top \mathbf{T} \mathbf{c} \\ \mathbf{c}^\top \mathbf{T}^\top \mathbf{X} & \text{tr}(\mathbf{D}) \end{pmatrix}, \quad (3.4)$$

Com as matrizes e vetores utilizados na expressão (3.4) sendo definidos da seguinte forma:

- $\mathbf{W} = \text{diag}\{w_1, \dots, w_n\}$, com:

$$w_t = \phi \{\psi'(\mu_t \phi) + \psi'((1 - \mu_t)\phi)\} \frac{1}{\{g'(\mu_t)\}^2}$$

onde $\psi'(\cdot)$ é a função trigama, ou seja, a segunda derivada da função digama;

- $\mathbf{c} = (c_1, \dots, c_n)^\top$, com:

$$c_t = \phi \{\psi'(\mu_t \phi) \mu_t - \psi'((1 - \mu_t)\phi) (1 - \mu_t)\};$$

- $\mathbf{D} = \text{diag}\{d_1, \dots, d_n\}$, com:

$$d_t = \psi'(\mu_t \phi) \mu_t^2 + \psi'((1 - \mu_t)\phi) (1 - \mu_t)^2 - \psi'(\phi).$$

Importante notar que os parâmetros β e ϕ não são ortogonais. Essa não-ortogonalidade implica que a estimativa de um parâmetro influencia a variabilidade do outro, tornando a análise mais complexa.

A matriz inversa $\mathbf{K}^{-1}(\beta, \phi)$, descrita em Ferrari e Cribari-Neto (2004), permite calcular os erros padrão assintóticos dos estimadores de máxima verossimilhança. Sob as condições usuais de regularidade para estimativa de máxima verossimilhança, quando o tamanho da amostra é grande,

$$\begin{pmatrix} \widehat{\beta} \\ \widehat{\phi} \end{pmatrix} \sim \mathcal{N}_{k+1} \left(\begin{pmatrix} \beta \\ \phi \end{pmatrix}, \mathbf{K}^{-1} \right).$$

4 Equações de Estimação Generalizadas (EEGs)

Enquanto os Modelos Lineares Generalizados (MLGs) e a Regressão Beta (Seção 3) pressupõem independência entre observações, estudos longitudinais exigem métodos que incorporem a correlação intraunidade. As EEGs, introduzidas por Liang e Zeger (1986), focam nas distribuições marginais da variável resposta, evitando a complexidade da modelagem conjunta. Essa abordagem, baseada em quasi-verossimilhança, produz estimadores consistentes e assintoticamente normais sob condições regulares.

4.1 Formulação das Equações de Estimação

As Equações de Estimação Generalizadas (EEGs) estabelecem um sistema de equações que garantem o equilíbrio entre valores observados e preditos através de três componentes fundamentais:

- **Função de Ligação:** Conecta a média marginal μ_{ij} ao preditor linear η_{ij} via

$$g(\mu_{ij}) = \mathbf{x}_{ij}^\top \boldsymbol{\beta},$$

onde $g(\cdot)$ é uma função monotônica e diferenciável (e.g., logito, probito).

- **Função de Variância:** Especifica a heteroscedasticidade dos dados através de

$$\text{Var}(y_{ij}) = \phi^{-1} \nu(\mu_{ij}),$$

sendo $\nu(\mu_{ij})$ a função de variância, e ϕ o parâmetro de escala.

- **Matriz de Covariância de Trabalho:** Modela a dependência intraunidade por

$$\mathbf{V}_i = \mathbf{A}_i^{1/2} \mathbf{R}_i(\boldsymbol{\alpha}) \mathbf{A}_i^{1/2},$$

sendo $\mathbf{A}_i = \text{diag}(\nu(\mu_{i1}), \dots, \nu(\mu_{iT_i}))$ a matriz de variâncias marginais.

O sistema de equações é formalizado por:

$$\sum_{i=1}^n \mathbf{D}_i^\top \mathbf{V}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) = \mathbf{0},$$

onde:

- $\mathbf{D}_i = \mathbf{X}_i \mathbf{G}_i$ é a matriz de sensibilidade, com $\mathbf{G}_i = \text{diag}(g'(\mu_{i1}), \dots, g'(\mu_{iT_i}))$
- \mathbf{V}_i^{-1} pondera os resíduos considerando a estrutura de dependência.

4.2 Estrutura de Covariância de Trabalho

A matriz de covariância de trabalho $\boldsymbol{\Omega}_{W_i}$ é construída como:

$$\boldsymbol{\Omega}_{W_i}(\boldsymbol{\alpha}, \phi) = \phi \mathbf{A}_i^{1/2} \mathbf{R}_W(\boldsymbol{\alpha}) \mathbf{A}_i^{1/2},$$

onde $\mathbf{R}_W(\boldsymbol{\alpha})$ é uma matriz de correlação parametrizada. Quando $\mathbf{R}_W(\boldsymbol{\alpha})$ coincide com a matriz de correlação verdadeira, temos:

$$\boldsymbol{\Omega}_{W_i} = \text{Var}(\mathbf{y}_i).$$

A escolha de $\mathbf{R}_W(\boldsymbol{\alpha})$ afeta a eficiência, mas não a consistência dos estimadores.

4.3 Estimação dos Parâmetros

Os estimadores de β são obtidos via algoritmo iterativo:

$$\begin{aligned} \hat{\beta}^{(h+1)} = \hat{\beta}^{(h)} + & \left[\sum_{i=1}^n \mathbf{X}_i^\top \hat{\Delta}_i^{(h)} [\Omega_{W_i}(\hat{\theta}^{(h)})]^{-1} \hat{\Delta}_i^{(h)} \mathbf{X}_i \right]^{-1} \\ & \times \left[\sum_{i=1}^n \mathbf{X}_i^\top \hat{\Delta}_i^{(h)} [\Omega_{W_i}(\hat{\theta}^{(h)})]^{-1} [\mathbf{y}_i - \boldsymbol{\mu}_i(\mathbf{X}_i^\top \hat{\beta}^{(h)})] \right], \end{aligned} \quad (4.1)$$

onde $\Delta_i = \text{diag}(\partial\mu_{ij}/\partial\eta_{ij})$ ajusta as ligações não lineares.

Sob condições de regularidade, a distribuição assintótica é dada por:

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{D} \mathcal{N}_p(\mathbf{0}, \mathbf{A}^{-1}),$$

em que

$$\mathbf{A}^{-1} = \lim_{n \rightarrow \infty} n \mathbf{S}^{-1} \left[\sum_{i=1}^n \mathbf{X}_i^\top \Delta_i [\Omega_{W_i}(\theta)]^{-1} \mathbf{V}(\mathbf{y}_i) [\Omega_{W_i}(\theta)]^{-1} \Delta_i \mathbf{X}_i \right] \mathbf{S}^{-1},$$

e sua matriz de covariância é robusta (estimador sanduíche):

$$\mathbf{A}^{-1} = \mathbf{S}^{-1} \mathbf{B} \mathbf{S}^{-1},$$

com:

$$\mathbf{S} = \sum_{i=1}^n \mathbf{X}_i^\top \Delta_i \Omega_{W_i}^{-1} \Delta_i \mathbf{X}_i$$

e

$$\mathbf{B} = \sum_{i=1}^n \mathbf{X}_i^\top \Delta_i \Omega_{W_i}^{-1} \mathbf{V}(\mathbf{y}_i) \Omega_{W_i}^{-1} \Delta_i \mathbf{X}_i$$

onde \mathbf{S} representa a informação de Fisher observada e \mathbf{B} captura a variabilidade empírica. Notavelmente, quando $\mathbf{R}_W(\alpha)$ está corretamente especificada, $\mathbf{B} \approx \mathbf{S}$, simplificando \mathbf{A}^{-1} para \mathbf{S}^{-1} .

4.4 Abordagens EEG1 vs. EEG2

- **EEG1 (LIANG; ZEGGER, 1986)**: Trata α como parâmetro de perturbação. Mantém consistência mesmo com $\mathbf{R}_W(\alpha)$ mal especificada, porém com perda de eficiência.
- **EEG2 (PRENTICE; ZHAO, 1991)**: Estima conjuntamente β e α . Exige correta especificação de $\mathbf{R}_W(\alpha)$ para eficiência ótima, sendo preferível quando a estrutura de dependência é de interesse científico.

5 Aplicação das EEGs em Modelos de Regressão Beta para Medidas Repetidas

Nesta seção, exploramos a implementação das Equações de Estimação Generalizadas (EEGs) em modelos de regressão beta, abordando o tratamento de dados longitudinais

— ou seja, dados provenientes de medidas repetidas — onde a resposta assume valores no intervalo $(0, 1)$. Inicialmente, focamos na modelagem do parâmetro de posição, que relaciona a média da resposta com as variáveis explicativas. Em seguida, consideramos a modelagem conjunta dos parâmetros de posição e precisão, permitindo que a variabilidade dos dados seja capturada por meio de um componente adicional.

5.1 Modelagem do Parâmetro de Posição

Para variáveis resposta que estão restritas ao intervalo $(0, 1)$, o modelo de regressão beta reparametrizado, conforme proposto por Ferrari e Cribari-Neto (2004), apresenta-se como uma abordagem natural. Suponha que, para a i -ésima unidade experimental ($i = 1, \dots, n$) e para a j -ésima observação ($j = 1, \dots, t$), a resposta y_{ij} siga uma distribuição beta:

$$y_{ij} \sim \mathcal{B}(\mu_{ij}, \phi),$$

onde μ_{ij} representa a média e ϕ é o parâmetro de precisão, assumido inicialmente conhecido e constante para todas as observações. A densidade dessa distribuição é dada por:

$$p(y_{ij}; \mu_{ij}, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu_{ij}\phi)\Gamma((1-\mu_{ij})\phi)} y_{ij}^{\mu_{ij}\phi-1} (1-y_{ij})^{(1-\mu_{ij})\phi-1}. \quad (5.1)$$

A modelagem da média μ_{ij} relaciona-a às covariáveis por meio de uma função de ligação $g(\cdot)$ (monotônica e duas vezes diferenciável). Assim, define-se o componente sistemático como:

$$g(\mu_{ij}) = \eta_{ij} = \mathbf{x}_{ij}^\top \boldsymbol{\beta}, \quad (5.2)$$

onde \mathbf{x}_{ij} é o vetor de covariáveis correspondente à j -ésima observação da i -ésima unidade e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros desconhecidos, com $p < n$.

Na construção das equações de estimação para modelos de regressão beta com medidas repetidas, é essencial definir funções de estimação que possuam propriedades regulares, isto é, que os vetores de erro tenham média zero, sejam mutuamente independentes e satisfaçam condições de regularidade. Inicialmente, pode-se considerar a definição simples

$$\mathbf{b}_i = \mathbf{y}_i - \boldsymbol{\mu}_i,$$

para

$$\mathbf{y}_i = (y_{i1}, \dots, y_{it})^\top \quad \text{e} \quad \boldsymbol{\mu}_i = (\mu_{i1}, \dots, \mu_{it})^\top.$$

Contudo, quando as observações de uma mesma unidade experimental são presumidas independentes, esses vetores \mathbf{b}_i não englobam, na classe $\mathcal{L}(\mathbf{b})$.

Para contornar essa limitação e formular equações de estimação capazes de generalizar tanto o caso de correlação intraunidade quanto o de independência, propõe-se definir:

$$\mathbf{b}_i = \mathbf{y}_i^* - \boldsymbol{\mu}_i^*,$$

sendo

$$\mathbf{y}_i^* = (y_{i1}^*, \dots, y_{it}^*)^\top \quad \text{e} \quad \boldsymbol{\mu}_i^* = (\mu_{i1}^*, \dots, \mu_{it}^*)^\top$$

são transformações das observações e da média, respectivamente, escolhidas de modo a garantir que tais vetores possuam média zero, sejam mutuamente independentes e incluam a função escore do modelo beta clássico.

Dessa forma, as equações de estimação para o parâmetro β são formuladas a partir da densidade marginal de y_{ij} (Equação 5.1) e do componente sistemático (Equação 5.2), adotando a definição:

$$\Psi_1^o(\beta) = \sum_{i=1}^n \mathbf{X}_i^\top \Lambda_i \text{Cov}(\mathbf{b}_i)^{-1} (\mathbf{y}_i^* - \boldsymbol{\mu}_i^*). \quad (5.3)$$

Nesta expressão:

- $\mathbf{X}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{it})^\top$ é a matriz das covariáveis;
- $\Lambda_i = \phi \mathbf{G}_i \mathbf{A}_i$, onde

$$\mathbf{G}_i = \text{diag} \left(\frac{\partial g^{-1}(\eta_{i1})}{\partial \eta_{i1}}, \dots, \frac{\partial g^{-1}(\eta_{it})}{\partial \eta_{it}} \right)$$

e

$$\mathbf{A}_i = \text{diag} (\mathbf{a}_{i1}, \dots, \mathbf{a}_{it}),$$

com cada

$$\mathbf{a}_{ij} = \psi'(\mu_{ij}\phi) + \psi'((1 - \mu_{ij})\phi);$$

- $\text{Cov}(\mathbf{b}_i)$ é a matriz de covariância dos vetores b_i , expressa em termos da transformação \mathbf{y}_i^* como

$$\text{Cov}(\mathbf{b}_i) = \text{Var}(\mathbf{y}_i^*)^{1/2} \mathbf{R}(\mathbf{y}_i^*) \text{Var}(\mathbf{y}_i^*)^{1/2} = \mathbf{A}_i^{1/2} \mathbf{R}(\mathbf{y}_i^*) \mathbf{A}_i^{1/2},$$

onde $\mathbf{R}(\mathbf{y}_i^*)$ é a verdadeira matriz de correlação das transformações \mathbf{y}_i^* .

Sob as condições do Teorema 1, de Venezuela (2008), o estimador $\hat{\beta}$, definido como a solução da equação de estimação $\Psi_1^o(\hat{\beta}) = \mathbf{0}$, é consistente e possui a distribuição assintótica

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathcal{D}} N_p \left(\mathbf{0}, \lim_{n \rightarrow \infty} n \left(\sum_{i=1}^n \mathbf{X}_i^\top \Lambda_i \text{Cov}(\mathbf{b}_i)^{-1} \Lambda_i \mathbf{X}_i \right)^{-1} \right).$$

Na prática, como a verdadeira matriz de correlação é geralmente desconhecida, segue-se a proposta de Liang e Zeger (1986) e define-se uma matriz de correlação de trabalho, $\mathbf{R}(\alpha)$, que possui dimensão $t \times t$ e satisfaz as condições necessárias para ser uma matriz de correlação. Essa matriz é caracterizada por um vetor de parâmetros α de dimensão $s \times 1$ e não precisa coincidir com a verdadeira matriz de correlação dos \mathbf{y}_i^* .

Dessa forma, a função de estimação generalizada adotada torna-se

$$\Psi_1(\beta) = \sum_{i=1}^n \mathbf{X}_i^\top \Lambda_i \Omega_i^{-1} (\mathbf{y}_i^* - \boldsymbol{\mu}_i^*) = \sum_{i=1}^n \mathbf{X}_i^\top \mathbf{W}_i \Lambda_i^{-1} \mathbf{b}_i, \quad (5.4)$$

onde a matriz de covariância de trabalho é definida por

$$\Omega_i = \mathbf{A}_i^{1/2} \mathbf{R}(\alpha) \mathbf{A}_i^{1/2},$$

e os pesos são dados por

$$\mathbf{W}_i = \Lambda_i \Omega_i^{-1} \Lambda_i.$$

Nota-se que, uma vez que a especificação de $\mathbf{R}(\boldsymbol{\alpha})$ pode impactar a eficiência dos estimadores sem afetar sua consistência, é necessário obter, por meio de métodos de estimação (por exemplo, o método dos momentos), estimativas consistentes para $\boldsymbol{\alpha}$ e para ϕ . Quando tais parâmetros adicionais são incorporados ao procedimento, o estimador de $\boldsymbol{\beta}$ mantém propriedades desejáveis (como consistência e normalidade assintótica), desde que as condições do teorema sejam satisfeitas, (VENEZUELA, 2008).

Por fim, a matriz de covariância de $\hat{\boldsymbol{\beta}}$ pode ser consistentemente estimada por meio do estimador robusto (ou “*sandwich*”), expresso por:

$$\hat{\mathbf{J}}_n^{-1} = \left\{ \sum_{i=1}^n \hat{\mathbf{S}}_i \right\}^{-1} \left\{ \sum_{i=1}^n \mathbf{X}_i^\top \hat{\boldsymbol{\Lambda}}_i \hat{\boldsymbol{\Omega}}_i^{-1} \hat{\mathbf{b}}_i \hat{\mathbf{b}}_i^\top \hat{\boldsymbol{\Omega}}_i^{-1} \hat{\boldsymbol{\Lambda}}_i \mathbf{X}_i \right\} \left\{ \sum_{i=1}^n \hat{\mathbf{S}}_i \right\}^{-1}.$$

Quando a matriz de correlação de trabalho $\mathbf{R}(\boldsymbol{\alpha})$ coincide com a verdadeira matriz de correlação dos \mathbf{y}_i^* , o estimador robusto se reduz à forma “*naive*” ou “*model-based*”:

$$\hat{\mathbf{J}}_n^{-1} = - \left\{ \sum_{i=1}^n \hat{\mathbf{S}}_i \right\}^{-1}.$$

Portanto, enquanto o estimador “*naive*” é consistente se a matriz de correlação de trabalho estiver corretamente especificada, o estimador robusto conserva consistência independentemente dessa especificação, embora possa apresentar viés elevado em amostras pequenas (por exemplo, quando $n < 20$). Uma boa indicação acerca da qualidade da estrutura escolhida para $\mathbf{R}(\boldsymbol{\alpha})$ é a proximidade entre as estimativas “*naive*” e robusta (JOHNSTON; DINARDO, 1996).

5.1.1 Estimação de $\boldsymbol{\beta}$, ϕ e $\boldsymbol{\alpha}$

O processo iterativo proposto para a estimação dos parâmetros $\boldsymbol{\beta}$, ϕ e $\boldsymbol{\alpha}$ combina o método de pontuação de Fisher para $\boldsymbol{\beta}$ com o método dos momentos para $\boldsymbol{\alpha}$ e ϕ . Expandindo a função de estimação $\boldsymbol{\Psi}_1(\boldsymbol{\beta})$ em torno de um valor inicial $\hat{\boldsymbol{\beta}}^{(0)}$, obtém-se a atualização iterativa dada por:

$$\begin{aligned} \hat{\boldsymbol{\beta}}^{(m+1)} = & \hat{\boldsymbol{\beta}}^{(m)} - \left\{ \mathbb{E} \left[\frac{\partial \boldsymbol{\Psi}_1(\hat{\boldsymbol{\beta}}^{(m)})}{\partial \boldsymbol{\beta}^\top} \right] \right\}^{-1} \boldsymbol{\Psi}_1(\hat{\boldsymbol{\beta}}^{(m)}) = \\ & \hat{\boldsymbol{\beta}}^{(m)} + \left\{ \left[\sum_{i=1}^n \mathbf{X}_i^\top \widehat{\mathbf{W}}_i \mathbf{X}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{X}_i^\top \widehat{\mathbf{W}}_i \hat{\boldsymbol{\Lambda}}_i^{-1} \hat{\mathbf{b}}_i \right] \right\}, \end{aligned} \quad (5.5)$$

para $m = 0, 1, 2, \dots$. Nessa expressão, todos os componentes (matrizes e vetores) são atualizados a cada iteração com base nas estimativas correntes de $\boldsymbol{\beta}$, ϕ e $\boldsymbol{\alpha}$.

Uma formulação alternativa do processo iterativo, utilizando o procedimento de mínimos quadrados ponderados, expressa a atualização de $\boldsymbol{\beta}$ por:

$$\hat{\boldsymbol{\beta}}^{(m+1)} = \left\{ \left[\sum_{i=1}^n \mathbf{X}_i^\top \widehat{\mathbf{W}}_i \mathbf{X}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{X}_i^\top \widehat{\mathbf{W}}_i \mathbf{z}_i \right] \right\}^{(m)}, \quad (5.6)$$

em que a variável dependente modificada é definida por

$$\mathbf{z}_i = \hat{\boldsymbol{\eta}}_i + \hat{\boldsymbol{\Lambda}}_i^{-1} \hat{\mathbf{b}}_i.$$

Esse formato iterativo é particularmente útil para o desenvolvimento de medidas de diagnóstico.

Adicionalmente, utilizando o método dos momentos e considerando que

$$\text{Var}(y_{ij}) = \frac{v(\mu_{ij})}{1 + \phi},$$

a estimação de ϕ no m -ésimo passo é dada por:

$$\hat{\phi}^{(m)} = \left(\left\{ \sum_{i=1}^n \sum_{j=1}^t \left(\frac{y_{ij} - \hat{\mu}_{ij}^{(m)}}{\sqrt{\hat{\mu}_{ij}^{(m)}(1 - \hat{\mu}_{ij}^{(m)})}} \right)^2 / (nt - p) \right\}^{-1} - 1 \right). \quad (5.7)$$

O procedimento descrito permite ainda a especificação de diferentes estruturas de correlação para as observações repetidas. Para esse fim, são desenvolvidos estimadores para a matriz de correlação de trabalho $\mathbf{R}(\boldsymbol{\alpha})$ sob diversas configurações, tais como:

- **Matriz de Correlação Padrão Uniforme:** Nesta estrutura, assume-se que

$$\text{Corr}(b_{ij}, b_{il}) = \alpha, \quad \forall j \neq l.$$

A estimativa de α no passo m é dada por:

$$\hat{\alpha}^{(m)} = \left\{ \frac{\sum_{i=1}^n \sum_{j>l}^t \hat{b}_{ij} \hat{b}_{il}}{\sum_{i=1}^n \sum_{j=1}^t \hat{b}_{ij}^2} \cdot \frac{2}{t-1} \right\}^{(m)}.$$

- **Matriz de Correlação AR-1:** Assume-se que

$$\text{Corr}(b_{ij}, b_{il}) = \alpha^{|j-l|}, \quad 1 \leq j, l \leq t.$$

Um estimador simples para α é:

$$\hat{\alpha}^{(m)} = \left\{ \frac{\sum_{i=1}^n \sum_{j=1}^{t-1} \hat{b}_{ij} \hat{b}_{ij+1}}{\sqrt{\left(\sum_{i=1}^n \sum_{j=1}^{t-1} \hat{b}_{ij}^2 \right) \left(\sum_{i=1}^n \sum_{j=2}^t \hat{b}_{ij}^2 \right)}} \right\}^{(m)}.$$

Note que esse estimador explora apenas a correlação entre observações consecutivas; estimativas alternativas podem considerar dependências mais longas.

- **Matriz de Correlação Não Estruturada:** Aqui, o vetor $\boldsymbol{\alpha} = (\alpha_{12}, \alpha_{13}, \dots, \alpha_{t-1,t})^\top$ possui $t(t-1)/2$ componentes, onde cada α_{jl} representa a correlação entre b_{ij} e b_{il} para $j < l$. Assim, estima-se:

$$\hat{\alpha}_{jl}^{(m)} = \left\{ \frac{\sum_{i=1}^n \hat{b}_{ij} \hat{b}_{il}}{\sqrt{\left(\sum_{i=1}^n \hat{b}_{ij}^2 \right) \left(\sum_{i=1}^n \hat{b}_{il}^2 \right)}} \right\}^{(m)}.$$

O procedimento prático para a estimação dos parâmetros β , ϕ e α pode ser resumido nas seguintes etapas:

- Inicialmente, supondo independência entre as observações dentro de cada unidade experimental, ajusta-se um modelo de regressão linear de $g(\mathbf{y})$ sobre \mathbf{X} mediante mínimos quadrados ordinários. Nesta etapa, os parâmetros α e ϕ não são considerados.
- Com as estimativas iniciais de β obtidas na etapa anterior, procede-se à estimação iterativa de β e ϕ utilizando as equações (5.6) e a fórmula para $\hat{\phi}$.
- Por fim, especifica-se uma estrutura de correlação de trabalho, $\mathbf{R}(\alpha)$. Se essa estrutura for, por exemplo, a de independência, os valores finais de β e ϕ correspondem aos obtidos na etapa anterior. Caso contrário, estima-se os parâmetros α e, em seguida, reestima-se β e ϕ de forma iterativa até que haja convergência.

5.2 Modelagem Conjunta dos Parâmetros de Posição e Precisão

Em cenários onde a precisão varia entre as observações, estende-se o modelo de regressão beta para incorporar essa heterogeneidade. Seja

$$\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{it_i})^\top$$

o vetor ($t_i \times 1$) de respostas da i -ésima unidade experimental, com $i = 1, \dots, n$. Assumimos que a densidade marginal da variável resposta siga uma distribuição beta, ou seja,

$$y_{ij} \sim \mathcal{B}(\mu_{ij}, \phi_{ij}),$$

com densidade dada por

$$p(y_{ij}; \mu_{ij}, \phi_{ij}) = \frac{\Gamma(\phi_{ij})}{\Gamma(\mu_{ij}\phi_{ij})\Gamma((1-\mu_{ij})\phi_{ij})} y_{ij}^{\mu_{ij}\phi_{ij}-1} (1-y_{ij})^{(1-\mu_{ij})\phi_{ij}-1}. \quad (5.8)$$

Nesse modelo, cada observação y_{ij} possui sua própria precisão, ϕ_{ij} , para $i = 1, \dots, n$ e $j = 1, \dots, t_i$.

Consideramos que o componente sistemático que modela a média permaneça o mesmo, conforme a Equação (5.2):

$$g(\mu_{ij}) = \eta_{ij} = \mathbf{x}_{ij}^\top \beta,$$

e que o parâmetro de precisão também seja modelado em função de covariáveis. Para isso, define-se um componente sistemático para ϕ_{ij} como

$$f(\phi_{ij}) = \delta_{ij} = \mathbf{q}_{ij}^\top \gamma, \quad (5.9)$$

onde $f(\cdot)$ é uma função monótona e duas vezes diferenciável cuja inversa é positiva, \mathbf{q}_{ij} representa o vetor de covariáveis (que pode ser um subconjunto de \mathbf{x}_{ij}) e $\gamma = (\gamma_1, \dots, \gamma_q)^\top$ é um vetor de parâmetros desconhecidos, com $q < n$. Para simplificar a notação, assumimos sem perda de generalidade que $t_i = t$ para todo $i = 1, \dots, n$.

A construção das equações de estimação para o modelo conjunto leva em consideração a densidade marginal em (5.8), os componentes sistemáticos (Equações (5.2) e (5.9)) e utiliza os vetores

$$\mathbf{b}_i = \mathbf{y}_i^* - \boldsymbol{\mu}_i^*,$$

onde \mathbf{y}_i^* e $\boldsymbol{\mu}_i^*$ são as versões transformadas que garantem as propriedades necessárias para a construção das funções de estimação.

Assumindo a correlação entre as observações da mesma unidade, a função de estimação ótima para os parâmetros conjuntos, definidos como $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$, é dada por

$$\Psi_2^o(\boldsymbol{\theta}) = \sum_{i=1}^n \begin{pmatrix} \mathbf{X}_i^\top \mathbf{G}_i \boldsymbol{\Phi}_i \mathbf{A}_i \\ \mathbf{Q}_i^\top \mathbf{F}_i \mathbf{C}_i \end{pmatrix} \left[\mathbf{A}_i^{1/2} \mathbf{R}(\mathbf{y}_i^*) \mathbf{A}_i^{1/2} \right]^{-1} (\mathbf{y}_i^* - \boldsymbol{\mu}_i^*), \quad (5.10)$$

onde:

- $\mathbf{X}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{it})^\top$;
- $\boldsymbol{\Phi}_i = \text{diag}(\phi_{i1}, \dots, \phi_{it})$;
- $\mathbf{A}_i = \text{diag}(a_{i1}, \dots, a_{it})$, com $a_{ij} = \psi'(\mu_{ij}\phi_{ij}) + \psi'((1 - \mu_{ij})\phi_{ij})$;
- $\mathbf{G}_i = \text{diag}\left(\frac{\partial g^{-1}(\eta_{i1})}{\partial \eta_{i1}}, \dots, \frac{\partial g^{-1}(\eta_{it})}{\partial \eta_{it}}\right)$;
- $\mathbf{Q}_i = (\mathbf{q}_{i1}, \dots, \mathbf{q}_{it})^\top$;
- $\mathbf{C}_i = \text{diag}(c_{i1}, \dots, c_{it})$, com $c_{ij} = \mu_{ij}\psi'(\mu_{ij}\phi_{ij}) - (1 - \mu_{ij})\psi'((1 - \mu_{ij})\phi_{ij})$;
- $\mathbf{F}_i = \text{diag}\left(\frac{\partial f^{-1}(\delta_{i1})}{\partial \delta_{i1}}, \dots, \frac{\partial f^{-1}(\delta_{it})}{\partial \delta_{it}}\right)$;
- $\mathbf{R}(\mathbf{y}_i^*)$ é a verdadeira matriz de correlação associada a \mathbf{y}_i^* .

Sob as condições do Teorema 1 (VENEZUELA, 2008), a solução $\widehat{\boldsymbol{\theta}}$ de $\Psi_2^o(\widehat{\boldsymbol{\theta}}) = \mathbf{0}$ é um estimador consistente para $\boldsymbol{\theta}$, e temos que

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathcal{D}} N_{p+q} \left(\mathbf{0}, \lim_{n \rightarrow \infty} n \left\{ \sum_{i=1}^n \begin{pmatrix} \mathbf{X}_i^\top \mathbf{G}_i \boldsymbol{\Phi}_i \mathbf{A}_i \\ \mathbf{Q}_i^\top \mathbf{F}_i \mathbf{C}_i \end{pmatrix} \text{Cov}(\mathbf{b}_i)^{-1} \begin{pmatrix} \mathbf{X}_i^\top \mathbf{G}_i \boldsymbol{\Phi}_i \mathbf{A}_i \\ \mathbf{Q}_i^\top \mathbf{F}_i \mathbf{C}_i \end{pmatrix}^\top \right\}^{-1} \right).$$

Como na modelagem somente do parâmetro de posição, substitui-se a verdadeira matriz de correlação por uma matriz de correlação de trabalho, $\mathbf{R}(\boldsymbol{\alpha})$, que satisfaça as condições para ser uma matriz de correlação, onde o vetor $\boldsymbol{\alpha}$, de dimensão $(s \times 1)$, a caracteriza completamente. Note que essa matriz de trabalho não precisa coincidir com a verdadeira matriz de correlação dos \mathbf{y}_i^* .

Logo, $\boldsymbol{\theta}$ tem sua função de estimação generalizada dada por

$$\begin{aligned} \Psi_2(\boldsymbol{\theta}) &= \begin{pmatrix} \Psi_2(\boldsymbol{\beta}) \\ \Psi_2(\boldsymbol{\gamma}) \end{pmatrix} = \sum_{i=1}^n \begin{pmatrix} \mathbf{X}_i^\top \mathbf{G}_i \boldsymbol{\Phi}_i \mathbf{A}_i \\ \mathbf{Q}_i^\top \mathbf{F}_i \mathbf{C}_i \end{pmatrix} \boldsymbol{\Omega}_i^{-1} (\mathbf{y}_i^* - \boldsymbol{\mu}_i^*) = \\ &= \sum_{i=1}^n \mathbf{M}_i^\top \boldsymbol{\Lambda}_i \boldsymbol{\Omega}_i^{-1} \mathbf{b}_i = \sum_{i=1}^n \mathbf{M}_i^\top \mathbf{W}_i \boldsymbol{\Lambda}_i^\top \mathbf{b}_i, \end{aligned} \quad (5.11)$$

sendo $\mathbf{A}_i^{-\top}$ a inversa generalizada de \mathbf{A}_i^\top , $\mathbf{M}_i = \begin{pmatrix} \mathbf{X}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_i \end{pmatrix}$, $\boldsymbol{\Lambda}_i = \begin{pmatrix} \mathbf{G}_i \boldsymbol{\Phi}_i \mathbf{A}_i \\ \mathbf{F}_i \mathbf{C}_i \end{pmatrix}$, $\boldsymbol{\Omega}_i = \mathbf{A}_i^{1/2} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{A}_i^{1/2}$ e $\mathbf{W}_i = \mathbf{A}_i \boldsymbol{\Omega}_i^{-1} \mathbf{A}_i^\top$. Como \mathbf{A}_i^\top é uma matriz com dimensão $t \times 2t$ com posto t , temos que $\mathbf{A}_i^\top \mathbf{A}_i^{-\top} = \mathbf{I}_t$.

No caso do ajuste conjunto dos parâmetros de posição e precisão, o estimador robusto da matriz de covariância de $\hat{\boldsymbol{\theta}}$ é dado por

$$\hat{\mathbf{J}}_n^{-1} = \left\{ \sum_{i=1}^n \hat{\mathbf{S}}_i \right\}^{-1} \left\{ \sum_{i=1}^n \mathbf{M}_i^\top \hat{\mathbf{\Lambda}}_i \hat{\mathbf{\Omega}}_i^{-1} \hat{\mathbf{b}}_i \hat{\mathbf{b}}_i^\top \hat{\mathbf{\Lambda}}_i^{-1} \mathbf{M}_i \right\} \left\{ \sum_{i=1}^n \hat{\mathbf{S}}_i \right\}^{-1},$$

obtida substituindo $\boldsymbol{\theta}$ e $\boldsymbol{\alpha}$ por suas estimativas consistentes. Quando a matriz de correlação de trabalho, $\mathbf{R}(\boldsymbol{\alpha})$, representar a verdadeira matriz de correlação dos \mathbf{y}_i^* , o estimador robusto se reduz à forma “naive”:

$$\hat{\mathbf{J}}_n^{-1} = - \left\{ \sum_{i=1}^n \hat{\mathbf{S}}_i \right\}^{-1}.$$

5.2.1 Estimação de $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ e $\boldsymbol{\alpha}$

O processo iterativo para calcular $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\boldsymbol{\gamma}}^\top)^\top$ e $\hat{\boldsymbol{\alpha}}$, combina o método de pontuação de Fisher para os parâmetros de regressão $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ com o método dos momentos para a estimação dos parâmetros de correlação $\boldsymbol{\alpha}$. Expandindo a EEGs dada na Equação (5.11) em torno de um valor inicial $\hat{\boldsymbol{\theta}}^{(0)} = (\hat{\boldsymbol{\beta}}^{(0)\top}, \hat{\boldsymbol{\gamma}}^{(0)\top})^\top$, o processo iterativo para atualizar $\boldsymbol{\theta}$ é dado por

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{(m+1)} &= \hat{\boldsymbol{\theta}}^{(m)} - \left\{ \mathbb{E} \left[\frac{\partial \boldsymbol{\Psi}_2(\hat{\boldsymbol{\theta}}^{(m)})}{\partial \boldsymbol{\theta}^\top} \right] \right\}^{-1} \boldsymbol{\Psi}_2(\hat{\boldsymbol{\theta}}^{(m)}) = \\ &= \hat{\boldsymbol{\theta}}^{(m)} + \left\{ \left[\sum_{i=1}^n \mathbf{M}_i^\top \widehat{\mathbf{W}}_i \mathbf{M}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{M}_i^\top \widehat{\mathbf{W}}_i \hat{\mathbf{A}}_i^{-\top} \hat{\mathbf{b}}_i \right] \right\}^{(m)}, \end{aligned} \quad (5.12)$$

para $m = 0, 1, 2, \dots$, onde as matrizes \mathbf{M}_i apresentam as derivadas dos preditores lineares em relação a $\boldsymbol{\theta}$, e $\widehat{\mathbf{W}}_i$ é a matriz de pesos baseada na estrutura de correlação de trabalho.

Essa atualização também pode ser reescrita como um procedimento de mínimos quadrados ponderados:

$$\hat{\boldsymbol{\theta}}^{(m+1)} = \left\{ \left[\sum_{i=1}^n \mathbf{M}_i^\top \widehat{\mathbf{W}}_i \mathbf{M}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{M}_i^\top \widehat{\mathbf{W}}_i \mathbf{z}_i \right] \right\}^{(m)}, \quad (5.13)$$

onde a variável modificada \mathbf{z}_i é definida por

$$\mathbf{z}_i = \hat{\boldsymbol{\nu}}_i + \hat{\mathbf{A}}_i^{-\top} \hat{\mathbf{b}}_i, \quad \text{com } \boldsymbol{\nu}_i = \begin{pmatrix} \boldsymbol{\eta}_i \\ \boldsymbol{\delta}_i \end{pmatrix}.$$

Para a estimação dos parâmetros de correlação, $\boldsymbol{\alpha}$, que determinam as correlações entre \mathbf{y}_{ij}^* e \mathbf{y}_{il}^* (para $j \neq l$), utiliza-se o procedimento descrito na Seção 5.1.1.

As etapas práticas para a estimação conjunta dos parâmetros $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ e $\boldsymbol{\alpha}$ podem ser resumidas da seguinte forma:

- Supondo inicialmente independência entre as observações de cada unidade experimental, ajusta-se um modelo de regressão linear de $g(\mathbf{y})$ sobre \mathbf{X} e outro de $f(\check{\boldsymbol{\phi}})$ sobre \mathbf{Q} por meio de mínimos quadrados ordinários, onde $\boldsymbol{\phi} = (\boldsymbol{\phi}_1^\top, \dots, \boldsymbol{\phi}_n^\top)^\top$ e

$\mathbf{Q} = (\mathbf{Q}_1^\top, \dots, \mathbf{Q}_n^\top)^\top$. Na prática, para o parâmetro de precisão, utiliza-se a sugestão de Ferrari e Cribari-Neto (2004):

$$\check{\phi}_{ij} = \frac{\check{\mu}_{ij}(1 - \check{\mu}_{ij})}{\check{\sigma}_{ij}^2} - 1,$$

sendo

$$\check{\mu}_{ij} = g^{-1} \left(\mathbf{x}_{ij}^\top \widehat{\boldsymbol{\beta}}_{\text{MQO}} \right) \quad \text{e} \quad \check{\sigma}_{ij}^2 = \frac{\check{\mathbf{e}}^\top \check{\mathbf{e}} \check{\mathbf{G}}_{ij}^2}{n - p}.$$

- Com as estimativas iniciais $\widehat{\boldsymbol{\beta}}^{(0)}$ e $\widehat{\boldsymbol{\gamma}}^{(0)}$, define-se uma estrutura de correlação de trabalho, $\mathbf{R}(\boldsymbol{\alpha})$, e estima-se uma primeira aproximação de $\boldsymbol{\alpha}$, denotada por $\widehat{\boldsymbol{\alpha}}^{(0)}$, utilizando os componentes do vetor \mathbf{b}_i .
- Em seguida, procede-se iterativamente à atualização dos parâmetros. Inicialmente, estima-se $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ separadamente usando:

$$\widehat{\boldsymbol{\beta}}^{(m+1)} = \left\{ \left[\sum_{i=1}^n \mathbf{X}_i^\top \widehat{\mathbf{W}}_{\beta_i} \mathbf{X}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{X}_i^\top \widehat{\mathbf{W}}_{\beta_i} \mathbf{z}_{\beta_i} \right] \right\}^{(m)}, \quad (5.14)$$

com

$$\widehat{\mathbf{W}}_{\beta_i} = \widehat{\mathbf{G}}_i \widehat{\Phi}_i \widehat{\mathbf{A}}_i \widehat{\Omega}_i^{-1} \widehat{\mathbf{A}}_i \widehat{\Phi}_i \widehat{\mathbf{G}}_i \quad \text{e} \quad \mathbf{z}_{\beta_i} = \widehat{\boldsymbol{\eta}}_i + \left(\widehat{\mathbf{A}}_i \widehat{\Phi}_i \widehat{\mathbf{G}}_i \right)^{-1} \widehat{\mathbf{b}}_i,$$

e por

$$\widehat{\boldsymbol{\gamma}}^{(m+1)} = \left\{ \left[\sum_{i=1}^n \mathbf{Q}_i^\top \widehat{\mathbf{W}}_{\gamma_i} \mathbf{Q}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{Q}_i^\top \widehat{\mathbf{W}}_{\gamma_i} \mathbf{z}_{\gamma_i} \right] \right\}^{(m)}, \quad (5.15)$$

com

$$\widehat{\mathbf{W}}_{\gamma_i} = \widehat{\mathbf{F}}_i \widehat{\mathbf{C}}_i \widehat{\Omega}_i^{-1} \widehat{\mathbf{C}}_i \widehat{\mathbf{F}}_i \quad \text{e} \quad \mathbf{z}_{\gamma_i} = \widehat{\boldsymbol{\delta}}_i + \left(\widehat{\mathbf{C}}_i \widehat{\mathbf{F}}_i \right)^{-1} \widehat{\mathbf{b}}_i.$$

As atualizações de $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ e $\boldsymbol{\alpha}$ prosseguem iterativamente até a convergência dos estimadores.

- Finalmente, utiliza-se o processo iterativo conjunto (Equação 5.13) com as estimativas iniciais $\widehat{\boldsymbol{\beta}}^{(0)}$, $\widehat{\boldsymbol{\gamma}}^{(0)}$ e $\widehat{\boldsymbol{\alpha}}^{(0)}$, repetindo cada iteração até que se atinja a convergência de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$.

6 Análise de Diagnóstico Para Dados de Medidas Repetidas

A análise diagnóstica em modelos de regressão é crucial para a verificação das hipóteses assumidas no ajuste, a identificação de possíveis violações e a detecção de observações que possam exercer influência desproporcional sobre os estimadores. Em modelos com medidas repetidas, essa análise torna-se ainda mais complexa, pois é necessário lidar com a estrutura de correlação intragrupo – frequentemente especificada através de uma matriz de correlação de trabalho $\mathbf{R}(\boldsymbol{\alpha})$.

Considerando as abordagens apresentadas na primeira parte do trabalho de Venezuela (2008), os parâmetros $\boldsymbol{\theta}$ podem ser interpretados de duas formas distintas. Se o interesse

for construir medidas de diagnóstico para avaliar apenas a modelagem da média, então $\boldsymbol{\theta} = \boldsymbol{\beta}$. Por outro lado, se for de desejo avaliar o comportamento conjunto da média e da dispersão dos dados, então $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$.

As medidas de diagnóstico resultantes serão vetores de dimensão $N \times 1$ quando $\boldsymbol{\theta} = \boldsymbol{\beta}$ (com $N = nt$) ou vetores $2N \times 1$ quando $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$.

6.1 Processo Iterativo Reponderado

Seja o processo iterativo reponderado para a estimação de $\boldsymbol{\theta}$ dado por

$$\hat{\boldsymbol{\theta}}^{(m+1)} = \left\{ \left[\sum_{i=1}^n \mathbf{M}_i^\top \widehat{\mathbf{W}}_i \mathbf{M}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{M}_i^\top \widehat{\mathbf{W}}_i \mathbf{z}_i \right] \right\}^{(m)}, \quad (6.1)$$

sendo:

- $\mathbf{M}_i = \mathbf{X}_i$ se $\boldsymbol{\theta} = \boldsymbol{\beta}$, ou $\mathbf{M}_i = \begin{bmatrix} \mathbf{X}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_i \end{bmatrix}$ se $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$;
- $\widehat{\mathbf{W}}_i = \widehat{\boldsymbol{\Lambda}}_i \widehat{\mathbf{Q}}_i^{-1} \widehat{\boldsymbol{\Lambda}}_i$;
- $\mathbf{z}_i = \widehat{\boldsymbol{\nu}}_i + \widehat{\boldsymbol{\Lambda}}_i^{-1} \widehat{\mathbf{b}}_i$, com $\widehat{\boldsymbol{\nu}}_i = \mathbf{M}_i \widehat{\boldsymbol{\theta}}$, para $i = 1, \dots, n$.

A convergência do processo iterativo leva à seguinte expressão para a estimativa de $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}} = (\mathbf{M}^\top \widehat{\mathbf{W}} \mathbf{M})^{-1} \mathbf{M}^\top \widehat{\mathbf{W}} \mathbf{z}, \quad (6.2)$$

em que $\mathbf{M} = (\mathbf{M}_1^\top, \dots, \mathbf{M}_n^\top)^\top$, $\widehat{\mathbf{W}} = \text{diag}(\widehat{\mathbf{W}}_1, \dots, \widehat{\mathbf{W}}_n)$ e $\mathbf{z} = (\mathbf{z}_1^\top, \dots, \mathbf{z}_n^\top)^\top$.

6.2 Ponto Alavanca Baseado na Matriz de Projção.

A estimativa $\hat{\boldsymbol{\theta}}$ pode ser interpretada como a solução de mínimos quadrados da regressão normal linear de $\widehat{\mathbf{W}}^{1/2} \mathbf{z}$, considerando $\widehat{\mathbf{W}}^{1/2} \mathbf{M}$ como matriz de planejamento. Assim, o resíduo ordinário é dado por

$$\mathbf{r}_O = \widehat{\mathbf{W}}^{1/2} (\mathbf{z} - \widehat{\boldsymbol{\nu}}) = (\mathbf{I} - \mathbf{H}) \widehat{\mathbf{W}}^{1/2} \mathbf{z}, \quad (6.3)$$

onde $\widehat{\boldsymbol{\nu}} = (\widehat{\boldsymbol{\nu}}_1^\top, \dots, \widehat{\boldsymbol{\nu}}_n^\top)^\top$, \mathbf{I} é a matriz identidade e \mathbf{H} é uma matriz bloco diagonal, com

$$\mathbf{H}_i = \widehat{\mathbf{W}}_i^{1/2} \mathbf{M}_i (\mathbf{M}^\top \widehat{\mathbf{W}} \mathbf{M})^{-1} \mathbf{M}_i^\top \widehat{\mathbf{W}}_i^{1/2}. \quad (6.4)$$

tendo dimensão $t' \times t'$, para todo $i = 1, \dots, n$. A matriz \mathbf{H} é simétrica ($\mathbf{H}^\top = \mathbf{H}$) e idempotente ($\mathbf{H}\mathbf{H} = \mathbf{H}$).

Os elementos da diagonal principal de \mathbf{H} nos auxiliam a detectar pontos alavanca. Se \mathbf{h}_{ij} for alto em relação aos outros valores, então a j -ésima observação da i -ésima unidade experimental pode ser um ponto alavanca.

6.3 Ponto Aberrante Baseado no Resíduo Padronizado

Assumindo que $\text{Cov}(\mathbf{z}) \approx \widehat{\mathbf{W}}^{-1}$, a matriz de covariâncias de \mathbf{r}_O pode ser aproximada por:

$$\begin{aligned}\text{Cov}(\mathbf{r}_O) &= \text{Cov}\left(\widehat{\mathbf{W}}^{1/2}\mathbf{z} - \widehat{\mathbf{W}}^{1/2}\mathbf{M}\widehat{\boldsymbol{\theta}}\right) \\ &= \text{Cov}\left(\widehat{\mathbf{W}}^{1/2}\mathbf{z} - \widehat{\mathbf{W}}^{1/2}\mathbf{M}(\mathbf{M}^\top\widehat{\mathbf{W}}\mathbf{M})^{-1}\mathbf{M}^\top\widehat{\mathbf{W}}\mathbf{z}\right) \\ &= (\mathbf{I} - \mathbf{H})\widehat{\mathbf{W}}^{1/2}\text{Cov}(\mathbf{z})\widehat{\mathbf{W}}^{1/2}(\mathbf{I} - \mathbf{H}) \cong (\mathbf{I} - \mathbf{H}).\end{aligned}$$

Como os elementos de \mathbf{r}_O possuem variâncias distintas, o resíduo padronizado para a observação y_{ij} é definido por:

$$(\mathbf{r}_{PD})_{ij} = \frac{\mathbf{e}_{(ij)}^\top \widehat{\mathbf{W}}^{1/2}(\mathbf{z} - \widehat{\boldsymbol{\nu}})}{\sqrt{1 - h_{ij}}} = \frac{\mathbf{e}_{(ij)}^\top \widehat{\mathbf{W}}^{1/2}\widehat{\boldsymbol{\Lambda}}^{-1}\widehat{\mathbf{b}}}{\sqrt{1 - h_{ij}}}, \quad (6.5)$$

em que:

- $\widehat{\mathbf{A}} = \text{diag}(\widehat{\mathbf{A}}_1, \dots, \widehat{\mathbf{A}}_n)$,
- $\widehat{\mathbf{b}} = (\widehat{\mathbf{b}}_1^\top, \dots, \widehat{\mathbf{b}}_n^\top)^\top$,
- $\mathbf{e}_{(ij)}$ é um vetor de tamanho conveniente com valor 1 na posição correspondente à observação y_{ij} e 0 nas demais posições,
- h_{ij} é o j -ésimo elemento da diagonal principal de \mathbf{H}_i , com $i = 1, \dots, n$ e $j = 1, \dots, t'$.

Observações com resíduos padronizados excessivamente grandes podem indicar a presença de pontos aberrantes.

6.4 Ponto Influyente Baseado na Distância de Cook

Um ponto influente é caracterizado por apresentar um perfil distinto dos demais em relação aos valores da variável resposta, ao mesmo tempo em que exibe um valor elevado na matriz de projeção \mathbf{H} . Essa influência pode ser identificada por meio da distância de Cook (COOK, 1977), uma métrica que quantifica o afastamento entre as estimativas do vetor paramétrico obtidas com todas as observações ($\widehat{\boldsymbol{\theta}}$) e sem a observação \mathbf{y}_{ij} ($\widehat{\boldsymbol{\theta}}_{(ij)}$), onde $i = 1, \dots, n$ e $j = 1, \dots, t'$.

Como, em geral, não é viável obter uma forma fechada para $\widehat{\boldsymbol{\theta}}_{(ij)}$, utiliza-se uma aproximação de um passo. Essa abordagem consiste em empregar a primeira iteração do processo iterativo do método de pontuação de Fisher, iniciando em $\widehat{\boldsymbol{\theta}}$. Introduzida por Pregibon (1981), essa aproximação é expressa para modelos de regressão com medidas repetidas da seguinte forma:

$$\widehat{\boldsymbol{\theta}}_{(ij)}^{(1)} = \widehat{\boldsymbol{\theta}} - \frac{\left[\mathbf{M}^\top\widehat{\mathbf{W}}\mathbf{M}\right]^{-1} \left[\mathbf{M}^\top\widehat{\mathbf{W}}^{1/2}\mathbf{e}_{(ij)}\right] \left[\mathbf{e}_{(ij)}^\top\widehat{\mathbf{W}}^{1/2}\boldsymbol{\Lambda}^{-1}\mathbf{b}\right]}{1 - h_{ij}}.$$

Conseqüentemente, a distância de Cook, ao excluir a observação \mathbf{y}_{ij} , é definida por:

$$(DC)_{ij} = \frac{1}{d}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(ij)})^\top \mathbf{M}^\top\widehat{\mathbf{W}}\mathbf{M}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(ij)}) = (r_{PD})_{ij}^2 \frac{h_{ij}}{d(1 - h_{ij})}. \quad (6.6)$$

Gráficos de $(DC)_{ij}$ em função do índice i são úteis para identificar pontos altamente influentes. Dessa forma, a distância de Cook permite detectar pontos influentes ao avaliar o impacto da exclusão de cada observação.

7 Seleção de Modelos e de Matriz de Correlação de Trabalho

O critério de informação de Akaike (AIC) é amplamente utilizado para selecionar o modelo mais parcimonioso dentre os ajustados, ou seja, aquele que oferece um bom ajuste com o menor número possível de parâmetros. Desenvolvido como um estimador assintoticamente não viesado para a informação de Kullback-Leibler (KULLBACK; LEIBLER, 1951), o AIC mede a discrepância entre um modelo candidato e o verdadeiro modelo. O modelo selecionado é aquele que minimiza a medida AIC, definida por:

$$AIC = -2\ell(\hat{\beta}) + 2p,$$

onde $\ell(\cdot)$ representa o logaritmo da função de verossimilhança associada aos dados, $\hat{\beta}$ é o estimador de máxima verossimilhança do modelo candidato, e p é a dimensão do vetor de parâmetros β .

No contexto das equações de estimação generalizadas, a construção não se baseia diretamente em verossimilhanças. Para contornar essa limitação, Pan (2001) propôs uma modificação no critério AIC, denominada QIC (Quasi-likelihood under the Independence model Criterion). Essa medida é útil tanto para a seleção de modelos quanto para a escolha de uma matriz de correlação de trabalho. A motivação para o QIC surge do fato de que, ao assumir independência entre todas as observações, as equações de estimação de Liang e Zeger (1986) são equivalentes à função quasi-escore (MCCULLAGH; NELDER, 1989) ou, no caso da família exponencial, à função escore correspondente.

Quando se assume independência entre as observações e homogeneidade da dispersão, as equações de estimação propostas por Venezuela (2008) (Parte I) são equivalentes às suas respectivas funções escores. Nesse contexto, a medida QIC para selecionar uma matriz de correlação de trabalho \mathbf{R} pode ser expressa como:

$$QIC(\mathbf{R}) \equiv -2\ell(\hat{\beta}(\mathbf{R})) + 2\text{tr}(\hat{\mathbf{S}}_I \hat{\mathbf{J}}_{nR}^{-1}), \quad (7.1)$$

onde:

- $\ell(\cdot)$ é o logaritmo da função de verossimilhança que gera a função escore equivalente às equações de estimação sob independência,
- $\hat{\mathbf{S}}_I$ é a matriz de sensibilidade sob a estrutura de independência,
- $\hat{\mathbf{J}}_{nR}^{-1}$ é o estimador robusto da matriz de covariância sob a estrutura \mathbf{R} ,
- $\hat{\beta}(\mathbf{R})$ é a estimativa de β obtida com a matriz de correlação de trabalho \mathbf{R} .

Quando todas as especificações do modelo via EEGs estão corretas, $\hat{\mathbf{S}}_I$ e $\hat{\mathbf{J}}_{nR}^{-1}$ são assintoticamente equivalentes, e $\text{tr}(\hat{\mathbf{S}}_I \hat{\mathbf{J}}_{nR}^{-1}) \approx 2$ (PAN, 2001). Nesse caso, a medida QIC pode ser simplificada para:

$$QIC_s(\mathbf{R}) \equiv -2\ell(\hat{\beta}(\mathbf{R})) + 2p, \quad (7.2)$$

que é mais adequada para a seleção de covariáveis.

As medidas QIC e QIC_s são válidas quando o parâmetro de dispersão ϕ^{-1} é conhecido e único. No caso em que ϕ^{-1} é desconhecido, a medida QIC deve ser calculada utilizando

o maior valor estimado para o parâmetro de dispersão dentre os modelos candidatos. Já a medida QIC_s deve ser calculada utilizando o valor estimado de ϕ^{-1} obtido do modelo ajustado com todas as covariáveis.

8 Influência Local para Equações de Estimação

A influência local tem como objetivo principal avaliar, por meio de uma medida apropriada, o impacto de pequenas perturbações no modelo ou nos dados. Quando essas perturbações resultam em efeitos desproporcionais em determinados resultados do modelo, isso pode indicar um ajuste inadequado ou desvios significativos das suposições feitas. A identificação das observações responsáveis por essas discrepâncias pode auxiliar na escolha de um modelo mais adequado aos dados.

Cook (1986) propôs um procedimento para avaliar a influência de pontos atípicos ou influentes, examinando o efeito de pequenas perturbações em componentes do modelo sobre a superfície de afastamento pela verossimilhança, definida por:

$$LD_\omega = 2 \left[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_\omega) \right],$$

onde:

- $\omega \in \mathcal{O} \subseteq \mathbb{R}^N$ é um vetor de perturbações, geralmente de dimensão $N \times 1$, com $N = nt$,
- $\ell(\cdot)$ é o logaritmo da função de verossimilhança,
- $\hat{\boldsymbol{\theta}}$ é o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ para o modelo postulado,
- $\hat{\boldsymbol{\theta}}_\omega$ é o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ para o modelo perturbado.

Cadigan e Farrell (2002) generalizaram essa medida de influência, considerando o afastamento de uma função de ajuste $\mathcal{F}(\boldsymbol{\theta})$, que é duplamente diferenciável em $\boldsymbol{\theta}$. O estimador de $\boldsymbol{\theta}$, denotado por $\hat{\boldsymbol{\theta}}$, é a solução de:

$$\boldsymbol{\Psi}(\hat{\boldsymbol{\theta}}) = \left. \frac{\partial \mathcal{F}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \mathbf{0}. \quad (8.1)$$

Assim, a medida de afastamento pela função de ajuste é dada por:

$$FD_\omega = 2 \left[\mathcal{F}(\hat{\boldsymbol{\theta}}) - \mathcal{F}(\hat{\boldsymbol{\theta}}_\omega) \right],$$

onde $\hat{\boldsymbol{\theta}}_\omega$ é a estimativa que maximiza a função de ajuste perturbada $\mathcal{F}(\boldsymbol{\theta}|\boldsymbol{\omega})$. Observe que, quando a função de ajuste \mathcal{F} é definida como o logaritmo da função de verossimilhança, temos $FD_\omega = LD_\omega$. No entanto, \mathcal{F} pode ser outra função, como uma quasi-verossimilhança, o que torna a abordagem de Cadigan e Farrell (2002) mais geral do que a proposta original de Cook.

8.1 Influência Local Generalizada

Conforme discutido anteriormente, a medida FD_ω pode ser utilizada para avaliar a influência de perturbações sobre a estimativa de θ ao variar ω em \mathcal{O} . No entanto, avaliar FD_ω para todos os valores possíveis de $\omega \in \mathcal{O}$ é inviável, dada a infinidade de valores que ω pode assumir. Para contornar esse problema, Cook (1986) propôs estudar o comportamento local de FD_ω em uma vizinhança de ω_0 , que representa a ausência de perturbação ($\mathcal{F}(\theta|\omega_0) = \mathcal{F}(\theta)$ ou $FD_{\omega_0} = 0$).

A abordagem de Cook consiste em investigar a curvatura normal (BATES; WATTS, 1980) da linha projetada no gráfico $FD_{\omega_0+ad} \times a$, onde $a \in \mathbb{R}$ e d é uma direção arbitrária com norma unitária ($\|d\| = 1$). Essa curvatura normal pode ser expressa de forma geral como:

$$C_d(\theta) = 2 \left| d^\top \Delta^\top \ddot{\mathcal{F}}^{-1} \Delta d \right|,$$

sendo:

$$\Delta = \frac{\partial^2 \mathcal{F}(\theta|\omega)}{\partial \theta \partial \omega^\top} = \frac{\partial \Psi(\theta|\omega)}{\partial \omega^\top}, \quad (8.2)$$

e

$$\ddot{\mathcal{F}} = \frac{\partial^2 \mathcal{F}(\theta)}{\partial \theta \partial \theta^\top} = \frac{\partial \Psi(\theta)}{\partial \theta^\top}, \quad (8.3)$$

com todas as quantidades avaliadas em $\theta = \hat{\theta}$ e $\omega = \omega_0$.

Para identificar as observações que, sob pequenas perturbações, exercem influência local significativa em FD_{ω_0} , inspecionamos a direção do autovetor d_{max} correspondente à maior curvatura C_{max} , obtida pelo maior autovalor da matriz:

$$-\Delta^\top \ddot{\mathcal{F}}^{-1} \Delta. \quad (8.4)$$

Ospina (2007) apresenta ilustrações que facilitam a compreensão da medida de influência local proposta por Cook, descrita em (8.4).

Além de identificar observações que influenciam conjuntamente a estimação do vetor de parâmetros θ , também é possível avaliar a influência local em apenas uma parte de θ . Suponha que θ seja particionado na forma $\theta = (\theta_1^\top, \theta_2^\top)^\top$. Se o interesse é identificar observações influentes apenas na estimação de θ_1 , a curvatura normal na direção d é dada por:

$$C_d(\theta_1) = 2 \left| d^\top \Delta^\top \left(\ddot{\mathcal{F}}^{-1} - \ddot{\mathcal{F}}^{\theta_2 \theta_2} \right) \Delta d \right|,$$

onde:

$$\ddot{\mathcal{F}}^{\theta_2 \theta_2} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathcal{F}}_{\theta_2 \theta_2}^{-1} \end{bmatrix},$$

com $\ddot{\mathcal{F}}_{\theta_2 \theta_2}$ avaliada em $\hat{\theta}$. O gráfico de índices de d_{max} da matriz $-\Delta^\top \left(\ddot{\mathcal{F}}^{-1} - \ddot{\mathcal{F}}^{\theta_2 \theta_2} \right) \Delta$ pode revelar quais observações são influentes em $\hat{\theta}_1$.

De forma análoga, a curvatura normal para o parâmetro θ_2 na direção d é dada por:

$$C_d(\theta_2) = 2 \left| d^\top \Delta^\top \left(\ddot{\mathcal{F}}^{-1} - \ddot{\mathcal{F}}^{\theta_1 \theta_1} \right) \Delta d \right|,$$

onde:

$$\ddot{\mathcal{F}}^{\theta_1 \theta_1} = \begin{bmatrix} \ddot{\mathcal{F}}_{\theta_1 \theta_1}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

com $\ddot{\mathcal{F}}_{\theta_1 \theta_1}$ avaliada em $\hat{\theta}$. A influência local das observações em $\hat{\theta}_2$ pode ser avaliada por meio do gráfico de índices do maior autovetor d_{max} da matriz $-\Delta^\top \left(\ddot{\mathcal{F}}^{-1} - \ddot{\mathcal{F}}^{\theta_1 \theta_1} \right) \Delta$.

8.2 Influência Local para Equações de Estimação

No contexto das equações de estimação generalizadas, a função de verossimilhança ou a função de ajuste que gera a equação de estimação nem sempre é conhecida. No entanto, podemos garantir sua existência ao assumir que qualquer equação de estimação generalizada, construída utilizando a verdadeira matriz de correlação ou uma matriz de correlação de trabalho $\mathbf{R}(\boldsymbol{\alpha})$ conhecida, satisfaz as propriedades de quasi-verossimilhança descritas por McCullagh e Nelder (1989).

Esses autores destacam que uma função quasi-escore com observações dependentes é o vetor gradiente de uma quasi-verossimilhança, desde que a derivada dessa função quasi-escore em relação a $\boldsymbol{\beta}$ seja uma matriz simétrica. Equivalentemente, isso ocorre quando as derivadas dos componentes de $\text{Cov}(\mathbf{y}_i)^{-1}$ em relação a $\boldsymbol{\mu}_i$ são iguais sob a permutação de três índices. Em outras palavras, a seguinte condição deve ser satisfeita:

$$\frac{\partial \text{Cov}(\mathbf{y}_{ij}, \mathbf{y}_{il})^{-1}}{\partial \mu_{ik}} = \frac{\partial \text{Cov}(\mathbf{y}_{ij}, \mathbf{y}_{ik})^{-1}}{\partial \mu_{il}} = \frac{\partial \text{Cov}(\mathbf{y}_{il}, \mathbf{y}_{ik})^{-1}}{\partial \mu_{ij}},$$

para $i = 1, \dots, n$ e $j, l, k = 1, \dots, t$.

Ao transferir essas propriedades para o caso em estudo, verifica-se que as derivadas dos componentes da matriz $\text{Cov}(\mathbf{b}_i)^{-1}$ em relação a $\boldsymbol{\mu}_i$ são iguais sob a permutação dos índices j, l e k , com $i = 1, \dots, n$ e $j, l, k = 1, \dots, t$. Isso ocorre tanto quando utiliza-se a verdadeira matriz de correlação $\mathbf{R}(\mathbf{b}_i)$ quanto quando emprega-se uma matriz de correlação de trabalho $\mathbf{R}(\boldsymbol{\alpha})$ conhecida, uma vez que ambas não dependem de $\boldsymbol{\mu}$.

Portanto, é garantido que qualquer equação de estimação $\boldsymbol{\Psi}(\boldsymbol{\theta})$, que utiliza a verdadeira matriz de correlação ou uma matriz de correlação de trabalho conhecida, é o vetor gradiente de uma função de ajuste $\mathcal{F}(\boldsymbol{\theta})$. Em outras palavras, podemos assumir que:

$$\exists \mathcal{F}(\boldsymbol{\theta}) \quad \text{tal que} \quad \frac{\partial \mathcal{F}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \boldsymbol{\Psi}(\boldsymbol{\theta}) \quad \text{e} \quad \boldsymbol{\Psi}(\hat{\boldsymbol{\theta}}) = \mathbf{0},$$

onde $\hat{\boldsymbol{\theta}}$ é o ponto de máximo da função de ajuste $\mathcal{F}(\boldsymbol{\theta})$.

Dessa forma, podemos utilizar a matriz $\boldsymbol{\Delta}$ definida na equação (8.2) para obter o autovetor d_{max} a partir da equação (8.4). A matriz hessiana $\ddot{\mathcal{F}}$, definida em (8.3), nem sempre é facilmente obtida para alguns modelos, mas pode ser simplificada utilizando seu valor esperado (CADIGAN, 1995). No nosso caso, o valor esperado de $\ddot{\mathcal{F}}$ é dado pela matriz de sensibilidade.

Assim, considerando que a equação de estimação $\boldsymbol{\Psi}(\boldsymbol{\theta})$ é o vetor gradiente de uma função de ajuste, ainda que desconhecida, e satisfaz (8.1), Venezuela (2008) propõe uma *medida de influência local para equações de estimação*, definida pelo autovetor d_{max} correspondente ao maior autovalor da matriz:

$$-\boldsymbol{\Delta}^\top \mathbf{S}^{-1} \boldsymbol{\Delta}, \tag{8.5}$$

onde:

$$\boldsymbol{\Delta} = \frac{\partial \boldsymbol{\Psi}(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \quad \text{e} \quad \mathbf{S} = \text{E} \left(\ddot{\mathcal{F}} \right) = \text{E} \left(\frac{\partial \boldsymbol{\Psi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top} \right),$$

com todas as quantidades avaliadas em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Influência na Estimação de $\boldsymbol{\beta}$ Para identificar observações influentes somente na estimação de $\boldsymbol{\beta}$, a curvatura normal na direção d é dada por:

$$C_d(\boldsymbol{\beta}) = 2 |d^\top \boldsymbol{\Delta}^\top (\mathbf{S}^{-1} - \mathbf{S}^{\gamma\gamma}) \boldsymbol{\Delta} d|,$$

onde:

$$\mathbf{S}^{\gamma\gamma} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{-1}_{\gamma\gamma} \end{bmatrix},$$

com $\mathbf{S}_{\gamma\gamma}$ avaliada em $\hat{\boldsymbol{\theta}}$. O gráfico de índices do maior autovetor de $-\boldsymbol{\Delta}^\top (\mathbf{S}^{-1} - \mathbf{S}^{\gamma\gamma}) \boldsymbol{\Delta}$ pode revelar quais observações são influentes na estimação de $\boldsymbol{\beta}$.

Influência na Estimação de $\boldsymbol{\gamma}$ De forma análoga, a curvatura normal para os parâmetros de regressão $\boldsymbol{\gamma}$ na direção d é dada por:

$$C_d(\boldsymbol{\gamma}) = 2 |d^\top \boldsymbol{\Delta}^\top (\mathbf{S}^{-1} - \mathbf{S}^{\beta\beta}) \boldsymbol{\Delta} d|,$$

onde:

$$\mathbf{S}^{\beta\beta} = \begin{bmatrix} \mathbf{S}^{-1}_{\beta\beta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

com $\mathbf{S}_{\beta\beta}$ avaliada em $\hat{\boldsymbol{\theta}}$. A influência local das observações em $\hat{\boldsymbol{\gamma}}$ pode ser avaliada por meio do gráfico de índices do maior autovetor d_{max} da matriz $-\boldsymbol{\Delta}^\top (\mathbf{S}^{-1} - \mathbf{S}^{\beta\beta}) \boldsymbol{\Delta}$.

8.3 Esquemas de Perturbação sob Homogeneidade da Dispersão

As equações de estimação generalizadas para a modelagem da média, sob homogeneidade do parâmetro de dispersão, são expressas por:

$$\Psi_1(\boldsymbol{\beta}) = \mathbf{X}^\top \boldsymbol{\Lambda} \boldsymbol{\Omega}^{-1} \mathbf{b} = \mathbf{X}^\top \mathbf{W} \boldsymbol{\Lambda}^{-1} \mathbf{b}, \quad (8.6)$$

onde a matriz de sensibilidade é dada por:

$$\mathbf{S} = -\mathbf{X}^\top \mathbf{W} \mathbf{X}, \quad (8.7)$$

sendo:

- $\mathbf{X} = (\mathbf{X}_1^\top, \dots, \mathbf{X}_n^\top)^\top$ a matriz de delineamento,
- $\mathbf{A} = \text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_n)$ a matriz de variâncias,
- $\boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_n)$ a matriz de covariâncias,
- $\mathbf{W} = \text{diag}(\mathbf{W}_1, \dots, \mathbf{W}_n)$ a matriz de pesos,
- $\mathbf{b} = (\mathbf{b}_1^\top, \dots, \mathbf{b}_n^\top)^\top$ o vetor de resíduos.

8.3.1 Ponderação de Casos

No esquema de perturbação por ponderação de casos, a equação de estimação perturbada é dada por:

$$\Psi_1(\boldsymbol{\beta}|\boldsymbol{\omega}) = \mathbf{X}^\top \mathbf{W} \boldsymbol{\Lambda}^{-1} \text{diag}(\boldsymbol{\omega}) \mathbf{b}, \quad (8.8)$$

onde $\boldsymbol{\omega}$ é o vetor de perturbações. A matriz $\boldsymbol{\Delta}$, que mede a sensibilidade da equação de estimação em relação às perturbações, é expressa por:

$$\boldsymbol{\Delta} = \mathbf{X}^\top \mathbf{W} \boldsymbol{\Lambda}^{-1} \text{diag}(\mathbf{b}).$$

A medida de influência local é obtida a partir da matriz:

$$\text{diag}(\mathbf{b}) \boldsymbol{\Lambda}^{-1} \mathbf{W}^\top \mathbf{X} (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W} \boldsymbol{\Lambda}^{-1} \text{diag}(\mathbf{b}), \quad (8.9)$$

avaliada em $\boldsymbol{\omega}_0$ e $(\hat{\boldsymbol{\beta}}^\top, \hat{\boldsymbol{\phi}})^\top$.

8.3.2 Perturbação da Variável Resposta

No esquema de perturbação aditiva da variável resposta y_{ij} , definimos:

$$y_{\omega ij} = y_{ij} + \omega_{ij} \sqrt{\text{var}(y_{ij})}, \quad (8.10)$$

onde $\omega_{ij} = 0$ indica ausência de perturbação. A equação de estimação perturbada é dada por:

$$\Psi_1(\beta|\omega) = \mathbf{X}^\top \mathbf{W} \Lambda^{-1} \mathbf{b}_\omega,$$

sendo \mathbf{b}_ω o vetor de resíduos perturbado. A matriz Δ , que mede a sensibilidade da equação de estimação em relação às perturbações, é expressa por:

$$\Delta = \mathbf{X}^\top \mathbf{W} \Lambda^{-1} \mathcal{B},$$

onde $\mathcal{B} = \frac{\partial \mathbf{b}_\omega}{\partial \omega^\top}$. A medida de influência local é obtida a partir da matriz:

$$\mathcal{B}^\top \Lambda^{-1} \mathbf{W}^\top \mathbf{X} (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W} \Lambda^{-1} \mathcal{B}, \quad (8.11)$$

avaliada em ω_0 e $(\widehat{\beta}^\top, \widehat{\phi})^\top$.

BETA De acordo com a equação de estimação definida em (5.4), o vetor perturbado $\mathbf{b}_{\omega i}$ referente à i -ésima unidade experimental é dado por:

$$\mathbf{b}_{\omega i} = \mathbf{y}_{\omega i}^* - \boldsymbol{\mu}_i^*, \quad (8.12)$$

onde $\mathbf{y}_{\omega i}^* = (y_{\omega i1}^*, \dots, y_{\omega it}^*)^\top$, com:

$$y_{\omega ij}^* = \log \left(\frac{y_{\omega ij}}{1 - y_{\omega ij}} \right), \quad i = 1, \dots, n \quad \text{e} \quad j = 1, \dots, t.$$

Com isso, temos:

$$\mathcal{B} = \mathcal{S} \Upsilon^{-1}, \quad (8.13)$$

onde:

$$\Upsilon = \text{diag}(\Upsilon_1, \dots, \Upsilon_n) \quad \text{e} \quad \mathcal{S} = \text{diag}(S_1, \dots, S_n).$$

Cada Υ_i e \mathcal{S}_i são definidos como:

$$\Upsilon_i = \text{diag}(v(y_{\omega i1}), \dots, v(y_{\omega it})) \quad \text{e} \quad \mathcal{S}_i = \text{diag}(s_{i1}, \dots, s_{it}),$$

para $i = 1, \dots, n$. Notemos que $v(y_{\omega ij})$, avaliado em ω_0 , é dado por:

$$v(y_{ij}) = y_{ij} (1 - y_{ij}), \quad i = 1, \dots, n \quad \text{e} \quad j = 1, \dots, t.$$

8.3.3 Perturbação Individual das Covariáveis

Considere um esquema aditivo de perturbação na k -ésima coluna da matriz de covariáveis \mathbf{X} , onde $\mathbf{x}_k = (x_{1ik}, x_{12k}, \dots, x_{nik})^\top$, conforme proposto por Thomas e Cook (1990). O vetor perturbado $\mathbf{x}_{\omega k}$ tem cada componente dado por:

$$x_{\omega ijk} = x_{ijk} + \omega_{ij} s_{zk}, \quad (8.14)$$

onde s_{zk} é um fator de escala dado pelo desvio padrão de \mathbf{x}_k , com $i = 1, \dots, n$ e $j = 1, \dots, t$. Para ilustrar, se $k \neq 2$ e $k \neq p$, o modelo (5.2) perturbado segundo o esquema definido em (8.14) é expresso por:

$$g(\mu_{\omega ij}) = \eta_{\omega ij} = \beta_1 + x_{ij2}\beta_2 + \dots + x_{\omega ijk}\beta_k + \dots + x_{ijp}\beta_p.$$

Neste esquema de perturbação, $\boldsymbol{\omega}_0 = \mathbf{0}$ indica ausência de perturbação.

Na equação de estimação definida em (8.6), todos os seus componentes são modificados quando uma covariável é perturbada. Assim, a equação de estimação perturbada é dada por:

$$\Psi_1(\boldsymbol{\beta}|\boldsymbol{\omega}) = \mathbf{X}_\omega^\top \boldsymbol{\Lambda}_\omega \boldsymbol{\Omega}_\omega^{-1} \mathbf{b}_\omega,$$

onde o índice ω indica que as matrizes \mathbf{X} , $\boldsymbol{\Lambda}$, $\boldsymbol{\Omega}$ e o vetor \mathbf{b} dependem da perturbação definida em (8.14).

A derivada de $\Psi_1(\boldsymbol{\beta}|\boldsymbol{\omega})$ em relação a $\boldsymbol{\omega}^\top$ pode ser expressa, conforme Harville (1997), por:

$$\boldsymbol{\Delta} = \mathbf{X}_\omega^\top \boldsymbol{\Lambda}_\omega \left[\boldsymbol{\Omega}_\omega^{-1} \frac{\partial \mathbf{b}_\omega}{\partial \boldsymbol{\omega}^\top} + \frac{\partial \boldsymbol{\Omega}_\omega^{-1}}{\partial \boldsymbol{\omega}^\top} \text{diag}(\mathbf{b}_\omega) \right] + \left[\mathbf{X}_\omega^\top \frac{\partial \boldsymbol{\Lambda}_\omega}{\partial \boldsymbol{\omega}^\top} + \frac{\partial \mathbf{X}_\omega^\top}{\partial \boldsymbol{\omega}^\top} \right] \boldsymbol{\Lambda}_\omega \boldsymbol{\Omega}_\omega^{-1} \text{diag}(\mathbf{b}_\omega),$$

onde a derivada de \mathbf{X}_ω^\top em relação a $\boldsymbol{\omega}^\top$ é uma matriz $p \times N$ de zeros, exceto a k -ésima linha, que é composta pela constante s_{zk} , com $N = nt$. Além disso,

$$\frac{\partial \boldsymbol{\Omega}_\omega^{-1}}{\partial \boldsymbol{\omega}^\top} = -\boldsymbol{\Omega}_\omega^{-1} \frac{\partial \boldsymbol{\Omega}_\omega}{\partial \boldsymbol{\omega}^\top} \boldsymbol{\Omega}_\omega^{-1}. \quad (8.15)$$

As derivadas de $\boldsymbol{\Lambda}_\omega$, $\boldsymbol{\Omega}_\omega$ e \mathbf{b}_ω em relação a $\boldsymbol{\omega}^\top$ são calculadas de acordo com a equação $\Psi_1(\boldsymbol{\beta})$. Essas equações representam as equações de estimação generalizadas para o modelo de regressão beta sob homogeneidade da dispersão.

As derivadas são expressas como:

$$\frac{\partial \boldsymbol{\Lambda}_\omega}{\partial \boldsymbol{\omega}^\top} = \text{diag} \left(\frac{\partial \boldsymbol{\Lambda}_{\omega 1}}{\partial \boldsymbol{\omega}^\top}, \dots, \frac{\partial \boldsymbol{\Lambda}_{\omega n}}{\partial \boldsymbol{\omega}^\top} \right),$$

$$\frac{\partial \boldsymbol{\Omega}_\omega}{\partial \boldsymbol{\omega}^\top} = \text{diag} \left(\frac{\partial \boldsymbol{\Omega}_{\omega 1}}{\partial \boldsymbol{\omega}^\top}, \dots, \frac{\partial \boldsymbol{\Omega}_{\omega n}}{\partial \boldsymbol{\omega}^\top} \right),$$

e

$$\frac{\partial \mathbf{b}_\omega}{\partial \boldsymbol{\omega}^\top} = \text{diag} \left(\frac{\partial \mathbf{b}_{\omega 1}}{\partial \boldsymbol{\omega}^\top}, \dots, \frac{\partial \mathbf{b}_{\omega n}}{\partial \boldsymbol{\omega}^\top} \right).$$

BETA De acordo com a equação de estimação definida em (5.4), temos:

$$\boldsymbol{\Lambda}_{\omega i} = \phi \mathbf{G}_{\omega i} \mathbf{A}_{\omega i}, \quad \boldsymbol{\Omega}_{\omega i} = \mathbf{A}_{\omega i}^{1/2} \mathbf{R}(\alpha) \mathbf{A}_{\omega i}^{1/2} \quad \text{e} \quad \mathbf{b}_{\omega i} = \mathbf{y}_i^* - \boldsymbol{\mu}_{\omega i}^*, \quad (6.24)$$

sendo:

- $\mathbf{G}_{\omega i} = \text{diag} \left(\frac{\partial g^{-1}(\eta_{\omega i 1})}{\partial \eta_{\omega i 1}}, \dots, \frac{\partial g^{-1}(\eta_{\omega i t})}{\partial \eta_{\omega i t}} \right),$
- $\mathbf{A}_{\omega i} = \text{diag} (a_{\omega i 1}, \dots, a_{\omega i t}),$

- $\boldsymbol{\mu}_{\omega_i}^* = (\mu_{\omega_{i1}}^*, \dots, \mu_{\omega_{it}}^*)^\top$,

com $a_{\omega_{ij}} = \psi'(\mu_{\omega_{ij}}\phi) + \psi'((1 - \mu_{\omega_{ij}})\phi)$ e $\mu_{\omega_{ij}}^* = \psi(\mu_{\omega_{ij}}\phi) - \psi((1 - \mu_{\omega_{ij}})\phi)$, para $i = 1, \dots, n$ e $j = 1, \dots, t$.

Da Seção B.1.2 do Apêndice B de Venezuela (2008), obtemos:

$$\begin{aligned}\frac{\partial \boldsymbol{\Lambda}_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} &= \phi \beta_k s_{x_k} \left[\dot{\mathbf{G}}_{\omega_i} \mathbf{A}_{\omega_i} + \mathbf{G}_{\omega_i}^2 \boldsymbol{\mathcal{E}}_{Loi} \right], \\ \frac{\partial \boldsymbol{\Omega}_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} &= \frac{1}{2} \beta_k s_{x_k} \left[\mathbf{A}_{\omega_i}^{1/2} \mathbf{R}(\alpha) \mathcal{A}_{\omega_i} + \mathcal{A}_{\omega_i} \mathbf{R}(\alpha) \mathbf{A}_{\omega_i}^{1/2} \right], \\ \frac{\partial \mathbf{b}_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} &= -\phi \beta_k s_{x_k} \mathbf{G}_{\omega_i} \mathbf{A}_{\omega_i},\end{aligned}$$

sendo:

- $\dot{\mathbf{G}}_{\omega_i} = \text{diag} \left(\frac{\partial^2 g^{-1}(\eta_{\omega_{i1}})}{\partial \eta_{\omega_{i1}}^2}, \dots, \frac{\partial^2 g^{-1}(\eta_{\omega_{it}})}{\partial \eta_{\omega_{it}}^2} \right)$,
- $\mathcal{A}_{\omega_i} = \mathbf{A}_{\omega_i}^{-1/2} \mathbf{G}_{\omega_i} \boldsymbol{\mathcal{E}}_{Loi}$,
- $\boldsymbol{\mathcal{E}}_{Loi} = \text{diag}(\boldsymbol{\mathcal{E}}_{Loi1}, \dots, \boldsymbol{\mathcal{E}}_{Loit})$,

com

$$\boldsymbol{\mathcal{E}}_{Loij} = \phi [\psi''(\mu_{\omega_{ij}}\phi) - \psi''((1 - \mu_{\omega_{ij}})\phi)], \quad (6.25)$$

e a função poligama $\psi''(\cdot)$ avaliada no ponto λ dada por:

$$\psi''(\lambda) = \frac{d\psi'(\lambda)}{d\lambda} = \frac{d^3 \log \Gamma(\lambda)}{d\lambda^3} = \frac{\Gamma''(\lambda)\Gamma(\lambda) - \Gamma'(\lambda)\Gamma'(\lambda)}{\Gamma(\lambda)^2} - \psi(\lambda)^2,$$

para $i = 1, \dots, n$ e $j = 1, \dots, t$.

8.3.4 Perturbação do Parâmetro de Precisão

Para a perturbação do parâmetro de precisão ϕ , definimos:

$$\phi_{\omega_{ij}} = \frac{\phi}{\omega_{ij}}, \quad (8.16)$$

onde $\boldsymbol{\omega}_0 = \mathbf{1}$ indica ausência de perturbação. A equação de estimação perturbada é dada por:

$$\boldsymbol{\Psi}_1(\boldsymbol{\beta}|\boldsymbol{\omega}) = \mathbf{X}^\top \boldsymbol{\Lambda}_\omega \boldsymbol{\Omega}_\omega^{-1} \mathbf{b}_\omega,$$

e a matriz $\boldsymbol{\Delta}$, que mede a sensibilidade da equação de estimação em relação às perturbações, é expressa por:

$$\boldsymbol{\Delta} = \mathbf{X}^\top \boldsymbol{\Lambda}_\omega \boldsymbol{\Omega}_\omega^{-1} \frac{\partial \mathbf{b}_\omega}{\partial \boldsymbol{\omega}^\top} + \mathbf{X}^\top \boldsymbol{\Lambda}_\omega \frac{\partial \boldsymbol{\Omega}_\omega^{-1}}{\partial \boldsymbol{\omega}^\top} \text{diag}(\mathbf{b}_\omega) + \mathbf{X}^\top \frac{\partial \boldsymbol{\Lambda}_\omega}{\partial \boldsymbol{\omega}^\top} \boldsymbol{\Omega}_\omega^{-1} \text{diag}(\mathbf{b}_\omega). \quad (8.17)$$

As derivadas são calculadas de acordo com as equações de estimação específicas.

BETA De acordo com a equação de estimação definida em (5.4), temos:

$$\Lambda_{\omega_i} = \mathbf{G}_i \Phi_{\omega_i} \mathbf{A}_{\omega_i}, \quad \Omega_{\omega_i} = \mathbf{A}_{\omega_i}^{1/2} \mathbf{R}(\alpha) \mathbf{A}_{\omega_i}^{1/2} \quad \text{e} \quad \mathbf{b}_{\omega_i} = \mathbf{y}_i^* - \boldsymbol{\mu}_{\omega_i}^*, \quad (8.18)$$

sendo:

- $\Phi_{\omega_i} = \text{diag}(\phi_{\omega_i 1}, \dots, \phi_{\omega_i t})$,
- $\mathbf{A}_{\omega_i} = \text{diag}(a_{\omega_i 1}, \dots, a_{\omega_i t})$,
- $\boldsymbol{\mu}_{\omega_i}^* = (\mu_{\omega_i 1}^*, \dots, \mu_{\omega_i t}^*)^\top$,

com $a_{\omega_{ij}} = \psi'(\mu_{ij} \phi_{\omega_{ij}}) + \psi'((1 - \mu_{ij}) \phi_{\omega_{ij}})$ e $\mu_{\omega_{ij}}^* = \psi(\mu_{ij} \phi_{\omega_{ij}}) - \psi((1 - \mu_{ij}) \phi_{\omega_{ij}})$, para $i = 1, \dots, n$ e $j = 1, \dots, t$.

Assim, obtemos as seguintes derivadas:

$$\begin{aligned} \frac{\partial \Lambda_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} &= -\phi^{-1} \mathbf{G}_i \Phi_{\omega_i}^2 [\Phi_{\omega_i} \boldsymbol{\mathcal{E}}_{\omega_i} + \mathbf{A}_{\omega_i}], \\ \frac{\partial \Omega_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} &= -\frac{1}{2} \phi^{-1} \left[\mathbf{A}_{\omega_i}^{1/2} \mathbf{R}(\alpha) \mathbf{A}_{\omega_i} + \mathbf{A}_{\omega_i} \mathbf{R}(\alpha) \mathbf{A}_{\omega_i}^{1/2} \right], \\ \frac{\partial \mathbf{b}_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} &= \phi^{-1} \Phi_{\omega_i}^2 \boldsymbol{\mathcal{E}}_{3\omega_i}, \end{aligned}$$

sendo:

- $\mathcal{A}_{\omega_i} = \mathbf{A}_{\omega_i}^{-1/2} \Phi_{\omega_i}^2 \boldsymbol{\mathcal{E}}_{2\omega_i}$,
- $\boldsymbol{\mathcal{E}}_{2\omega_i} = \text{diag}(\mathcal{E}_{2\omega_i 1}, \dots, \mathcal{E}_{2\omega_i t})$,
- $\boldsymbol{\mathcal{E}}_{3\omega_i} = \text{diag}(\mathcal{E}_{3\omega_i 1}, \dots, \mathcal{E}_{3\omega_i t})$,

com

$$\mathcal{E}_{2\omega_{ij}} = [\mu_{ij} \psi''(\mu_{ij} \phi_{\omega_{ij}}) + (1 - \mu_{ij}) \psi''((1 - \mu_{ij}) \phi_{\omega_{ij}})], \quad (8.19)$$

e

$$\mathcal{E}_{3\omega_{ij}} = [\mu_{ij} \psi'(\mu_{ij} \phi_{\omega_{ij}}) - (1 - \mu_{ij}) \psi'((1 - \mu_{ij}) \phi_{\omega_{ij}})], \quad (8.20)$$

para $i = 1, \dots, n$ e $j = 1, \dots, t$.

8.3.5 Perturbação na Matriz de Correlação de Trabalho $\mathbf{R}(\alpha)$

Para a perturbação na matriz de correlação de trabalho, temos:

$$\alpha_{\omega_i(jj')} = \frac{\alpha_{jj'}}{\omega_i(jj')}, \quad (8.21)$$

com $\boldsymbol{\omega}_0 = \mathbf{1}$. A equação de estimação perturbada é:

$$\Psi_1(\boldsymbol{\beta} | \boldsymbol{\omega}) = \mathbf{X}^\top \Lambda \Omega_{\boldsymbol{\omega}}^{-1} \mathbf{b},$$

e a matriz Δ é dada por:

$$\Delta = \mathbf{X}^\top \Lambda \frac{\partial \Omega_{\boldsymbol{\omega}}^{-1}}{\partial \boldsymbol{\omega}^\top} \text{diag}(\mathbf{b}),$$

A medida de influência local com perturbação na matriz de correlação de trabalho é obtida da matriz:

$$\text{diag}(\mathbf{b}) \Omega^{-1} \Lambda^\top \mathbf{X} (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\top \Lambda \Omega^{-1} \text{diag}(\mathbf{b}),$$

avaliada em $\boldsymbol{\omega}_0$ e $(\widehat{\boldsymbol{\beta}}^\top, \widehat{\boldsymbol{\phi}})^\top$.

8.4 Esquemas de Perturbação sob Heterogeneidade da Dispersão

As equações de estimação generalizadas para a modelagem conjunta da média e da dispersão no modelo de regressão beta podem ser escritas de forma geral como:

$$\Psi_2(\boldsymbol{\theta}) = \mathbf{M}^\top \boldsymbol{\Lambda} \boldsymbol{\Omega}^{-1} \mathbf{b} = \mathbf{M}^\top \mathbf{W} \boldsymbol{\Lambda}^{-\top} \mathbf{b}, \quad (8.22)$$

com a matriz de sensibilidade expressa por:

$$\mathbf{S} = -\mathbf{M}^\top \mathbf{W} \mathbf{M}, \quad (8.23)$$

onde:

- $\mathbf{M} = (\mathbf{M}_1^\top, \dots, \mathbf{M}_n^\top)^\top$ é a matriz de delineamento,
- $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{\Lambda}_1, \dots, \boldsymbol{\Lambda}_n)$ é a matriz de variâncias,
- $\boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_n)$ é a matriz de covariâncias,
- $\mathbf{W} = \text{diag}(\mathbf{W}_1, \dots, \mathbf{W}_n)$ é a matriz de pesos,
- $\mathbf{b} = (\mathbf{b}_1^\top, \dots, \mathbf{b}_n^\top)^\top$ é o vetor de resíduos.

Note que, para modelos de regressão beta, $\boldsymbol{\Lambda}$ não é uma matriz quadrada, sendo importante distinguir sua matriz transposta quando necessário.

As medidas de influência local, descritas a seguir para alguns esquemas de perturbação com heterogeneidade da dispersão, são definidas a partir das equações (8.22) e (8.23), avaliadas sob as estimativas do modelo postulado $\hat{\boldsymbol{\theta}}$ e em $\boldsymbol{\omega}_0$.

8.4.1 Ponderação de Casos

No esquema de perturbação por ponderação de casos, a equação de estimação perturbada é dada por:

$$\Psi_2(\boldsymbol{\theta}|\boldsymbol{\omega}) = \mathbf{M}^\top \mathbf{W} \boldsymbol{\Lambda}^{-\top} \text{diag}(\boldsymbol{\omega}) \mathbf{b}, \quad (8.24)$$

onde $\boldsymbol{\omega} = (\boldsymbol{\omega}_1^\top, \dots, \boldsymbol{\omega}_n^\top)^\top$, com $\boldsymbol{\omega}_i = (\omega_{i1}, \dots, \omega_{it})^\top$, para $i = 1, \dots, n$. Aqui, $\boldsymbol{\omega}_0$, que representa ausência de perturbação, é um vetor de uns.

A matriz $\boldsymbol{\Delta}$, que equivale à derivada de $\Psi_2(\boldsymbol{\theta}|\boldsymbol{\omega})$ em relação a $\boldsymbol{\omega}^\top$, é dada por:

$$\boldsymbol{\Delta} = \mathbf{M}^\top \mathbf{W} \boldsymbol{\Lambda}^{-\top} \mathbf{B},$$

onde $\mathbf{B} = \text{diag}(\mathbf{b})$ para o modelo de regressão beta.

A matriz definida em (8.5) é então expressa por:

$$\mathbf{B} \boldsymbol{\Lambda}^{-1} \mathbf{W}^\top \mathbf{M} (\mathbf{M}^\top \mathbf{W} \mathbf{M})^{-1} \mathbf{M}^\top \mathbf{W} \boldsymbol{\Lambda}^{-\top} \mathbf{B},$$

avaliada em $\hat{\boldsymbol{\theta}}$ e em $\boldsymbol{\omega}_0$.

8.4.2 Perturbação da Variável Resposta

Considerando \mathbf{b}_ω como o vetor \mathbf{b} com perturbação na variável resposta, conforme o esquema expresso em (8.10), a equação de estimação perturbada é dada por:

$$\Psi_2(\boldsymbol{\theta}|\boldsymbol{\omega}) = \mathbf{M}^\top \mathbf{W} \boldsymbol{\Lambda}^{-\top} \mathbf{b}_\omega,$$

onde $\mathbf{b}_\omega = (\mathbf{b}_{\omega 1}^\top, \dots, \mathbf{b}_{\omega n}^\top)^\top$, com $\mathbf{b}_{\omega i} = (b_{\omega i 1}, \dots, b_{\omega i t})^\top$, para $i = 1, \dots, n$.

A medida de influência local com perturbação da variável resposta é obtida a partir da matriz:

$$\mathcal{B}^\top \boldsymbol{\Lambda}^{-1} \mathbf{W}^\top \mathbf{M} (\mathbf{M}^\top \mathbf{W} \mathbf{M})^{-1} \mathbf{M}^\top \mathbf{W} \boldsymbol{\Lambda}^{-\top} \mathcal{B},$$

avaliada em $\hat{\boldsymbol{\theta}}$ e em $\boldsymbol{\omega}_0$, onde $\mathcal{B} = \partial \mathbf{b}_\omega / \partial \boldsymbol{\omega}^\top$.

Modelo Beta Para a equação de estimação definida em (5.11), o vetor perturbado na variável resposta referente à i -ésima unidade experimental $\mathbf{b}_{\omega i}$ é dado pela equação (8.12). Assim, no caso de modelagem conjunta da média e da dispersão, \mathcal{B} é expresso conforme (8.13), mas com:

$$s_{ij} = \sqrt{\frac{v(\mu_{ij})}{1 + \phi_{ij}}}, \quad i = 1, \dots, n \quad \text{e} \quad j = 1, \dots, t.$$

8.4.3 Perturbação Individual das Covariáveis

A k -ésima coluna da matriz de covariáveis \mathbf{X} será modificada conforme a proposta apresentada na equação (8.14), para $k = 2, \dots, p$. Perturbando a equação (8.22), obtemos:

$$\Psi_2(\boldsymbol{\theta}|\boldsymbol{\omega}) = \mathbf{M}_\omega^\top \boldsymbol{\Lambda}_\omega \boldsymbol{\Omega}_\omega^{-1} \mathbf{b}_\omega,$$

onde o índice ω indica que as matrizes \mathbf{M} , $\boldsymbol{\Lambda}$, $\boldsymbol{\Omega}$ e o vetor \mathbf{b} dependem da perturbação definida em (8.14). A matriz $\boldsymbol{\Delta}$, definida em (8.2), pode ser expressa por:

$$\boldsymbol{\Delta} = \mathbf{M}_\omega^\top \boldsymbol{\Lambda}_\omega \left[\boldsymbol{\Omega}_\omega^{-1} \frac{\partial \mathbf{b}_\omega}{\partial \boldsymbol{\omega}^\top} + \frac{\partial \boldsymbol{\Omega}_\omega^{-1}}{\partial \boldsymbol{\omega}^\top} \mathbf{B}_\omega \right] + \left[\mathbf{M}_\omega^\top \frac{\partial \boldsymbol{\Lambda}_\omega}{\partial \boldsymbol{\omega}^\top} + \frac{\partial \mathbf{M}_\omega^\top}{\partial \boldsymbol{\omega}^\top} \boldsymbol{\Lambda}_\omega \right] \boldsymbol{\Omega}_\omega^{-1} \mathbf{B}_\omega,$$

onde \mathbf{B}_ω é a matriz definida na subseção "Ponderação de casos", perturbada conforme o esquema (8.14). As derivadas de $\boldsymbol{\Lambda}_\omega$, $\boldsymbol{\Omega}_\omega^{-1}$ e \mathbf{b}_ω em relação a $\boldsymbol{\omega}^\top$ são calculadas conforme mostrado na subseção "Perturbação Individual das Covariáveis".

Conforme a proposta de Ospina (2007), quando os parâmetros de posição e de precisão são modelados simultaneamente, destacamos quatro cenários que relacionam as matrizes que modelam a média e a precisão:

1. $\mathbf{X} \neq \mathbf{Q}$: matrizes totalmente diferentes;
2. $\mathbf{X} = \mathbf{Q}$: matrizes totalmente iguais;
3. $\mathbf{X} \neq \mathbf{Q}$: com a k' -ésima coluna da matriz \mathbf{Q} igual à k -ésima coluna da matriz \mathbf{X} ;
4. $\mathbf{X} \neq \mathbf{Q}$: com a k' -ésima coluna da matriz \mathbf{Q} sendo uma função da k -ésima coluna da matriz \mathbf{X} .

A seguir, para cada um desses quatro cenários, apresentamos as medidas de influência local com perturbação em uma covariável.

8.4.4 Matriz \mathbf{Q} Totalmente Diferente da Matriz \mathbf{X}

Considerando que a matriz de covariáveis \mathbf{Q} , que modela a precisão, é totalmente diferente da matriz \mathbf{X} , a qual tem a k -ésima coluna perturbada segundo o esquema definido em (8.14), a matriz de covariáveis \mathbf{M} modificada é dada por:

$$\mathbf{M}_\omega^\top = \begin{pmatrix} \mathbf{X}_\omega^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^\top \end{pmatrix},$$

onde o índice ω indica que a matriz \mathbf{X} depende da perturbação definida em (8.14). Logo, temos:

$$\frac{\partial \mathbf{M}_\omega^\top}{\partial \boldsymbol{\omega}^\top} = \begin{pmatrix} \frac{\partial \mathbf{X}_\omega^\top}{\partial \boldsymbol{\omega}^\top} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

sendo uma matriz $(p+q) \times 2N$, onde a derivada de \mathbf{X}_ω^\top em relação a $\boldsymbol{\omega}^\top$ é uma matriz $p \times N$ de zeros, exceto a k -ésima linha, que é composta pela constante s_{x_k} , com $N = nt$.

A seguir, são apresentadas as derivadas de $\boldsymbol{\Lambda}_\omega$, $\boldsymbol{\Omega}_\omega$ e \mathbf{b}_ω para o modelo Beta.

Modelo Beta De acordo com a equação de estimação definida em (5.11), temos:

$$\boldsymbol{\Lambda}_{\omega i} = \begin{pmatrix} \mathbf{G}_{\omega i} \boldsymbol{\Phi}_i \mathbf{A}_{\omega i} \\ \mathbf{F}_i \mathbf{C}_{\omega i} \end{pmatrix}, \quad \boldsymbol{\Omega}_{\omega i} = \mathbf{A}_{\omega i}^{1/2} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{A}_{\omega i}^{1/2}, \quad \text{e} \quad \mathbf{b}_{\omega i} = \mathbf{y}_i^* - \boldsymbol{\mu}_{\omega i}^*, \quad (8.25)$$

sendo:

- $\mathbf{G}_{\omega i} = \text{diag} \left(\frac{\partial g^{-1}(\eta_{\omega i1})}{\partial \eta_{\omega i1}}, \dots, \frac{\partial g^{-1}(\eta_{\omega it})}{\partial \eta_{\omega it}} \right)$,
- $\mathbf{A}_{\omega i} = \text{diag} (\mathbf{a}_{\omega i1}, \dots, \mathbf{a}_{\omega it})$,
- $\mathbf{C}_{\omega i} = \text{diag} (\mathbf{c}_{\omega i1}, \dots, \mathbf{c}_{\omega it})$,
- $\boldsymbol{\mu}_{\omega i}^* = (\mu_{\omega i1}^*, \dots, \mu_{\omega it}^*)^\top$,

com:

$$\begin{aligned} \mathbf{a}_{\omega ij} &= \psi'(\mu_{\omega ij} \phi_{ij}) + \psi'((1 - \mu_{\omega ij}) \phi_{ij}), \\ \mathbf{c}_{\omega ij} &= \mu_{\omega ij} \psi'(\mu_{\omega ij} \phi_{ij}) - (1 - \mu_{\omega ij}) \psi'((1 - \mu_{\omega ij}) \phi_{ij}), \end{aligned}$$

e

$$\mu_{\omega ij}^* = \psi(\mu_{\omega ij} \phi_{ij}) - \psi((1 - \mu_{\omega ij}) \phi_{ij}),$$

para $i = 1, \dots, n$ e $j = 1, \dots, t$.

Assim, é obtido as seguintes derivadas:

$$\begin{aligned} \frac{\partial \boldsymbol{\Lambda}_{\omega i}}{\partial \boldsymbol{\omega}_i^\top} &= \beta_k s_{x_k} \begin{pmatrix} \boldsymbol{\Phi}_i \left[\dot{\mathbf{G}}_{\omega i} \mathbf{A}_{\omega i} + \mathbf{G}_{\omega i}^2 \boldsymbol{\mathcal{E}}_{1\omega i} \right] \\ \mathbf{F}_i \mathbf{G}_{\omega i} \left[\mathbf{A}_{\omega i} + \boldsymbol{\Phi}_i \boldsymbol{\mathcal{E}}_{2\omega i} \right] \end{pmatrix}, \\ \frac{\partial \boldsymbol{\Omega}_{\omega i}}{\partial \boldsymbol{\omega}_i^\top} &= \frac{1}{2} \beta_k s_{x_k} \left[\mathbf{A}_{\omega i}^{1/2} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{A}_{\omega i} + \mathbf{A}_{\omega i} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{A}_{\omega i}^{1/2} \right], \\ \frac{\partial \mathbf{b}_{\omega i}}{\partial \boldsymbol{\omega}_i^\top} &= -\beta_k s_{x_k} \mathbf{G}_{\omega i} \boldsymbol{\Phi}_i \mathbf{A}_{\omega i}, \end{aligned}$$

sendo:

- $\dot{\mathbf{G}}_{\omega i} = \text{diag} \left(\frac{\partial^2 g^{-1}(\eta_{\omega i 1})}{\partial \eta_{\omega i 1}^2}, \dots, \frac{\partial^2 g^{-1}(\eta_{\omega i t})}{\partial \eta_{\omega i t}^2} \right),$
- $\mathbf{A}_{\omega i} = \mathbf{A}_{\omega i}^{-1/2} \mathbf{G}_{\omega i} \mathcal{E}_{1\omega i},$
- $\mathcal{E}_{1\omega i} = \text{diag} (\mathcal{E}_{1\omega i 1}, \dots, \mathcal{E}_{1\omega i t}),$
- $\mathcal{E}_{2\omega i} = \text{diag} (\mathcal{E}_{2\omega i 1}, \dots, \mathcal{E}_{2\omega i t}),$

com:

$$\mathcal{E}_{1\omega ij} = \phi_{ij} [\psi''(\mu_{\omega ij} \phi_{ij}) - \psi''((1 - \mu_{\omega ij}) \phi_{ij})], \quad (8.26)$$

e

$$\mathcal{E}_{2\omega ij} = [\mu_{\omega ij} \psi''(\mu_{\omega ij} \phi_{ij}) + (1 - \mu_{\omega ij}) \psi''((1 - \mu_{\omega ij}) \phi_{ij})], \quad (8.27)$$

para $i = 1, \dots, n$ e $j = 1, \dots, t$.

8.4.5 Matriz \mathbf{Q} Totalmente Igual à Matriz \mathbf{X}

Considerando que a matriz de covariáveis \mathbf{Q} que modela a precisão é totalmente igual à matriz \mathbf{X} , a qual tem a k -ésima coluna perturbada segundo o esquema definido em (8.14), a matriz de covariáveis \mathbf{M} modificada é dada por:

$$\mathbf{M}_{\omega}^{\top} = \begin{pmatrix} \mathbf{X}_{\omega}^{\top} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{\omega}^{\top} \end{pmatrix}, \quad (8.28)$$

onde o índice ω indica que a matriz \mathbf{X} depende da perturbação definida em (6.20). Logo, temos:

$$\frac{\partial \mathbf{M}_{\omega}^{\top}}{\partial \boldsymbol{\omega}^{\top}} = \begin{pmatrix} \frac{\partial \mathbf{X}_{\omega}^{\top}}{\partial \boldsymbol{\omega}^{\top}} & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{X}_{\omega}^{\top}}{\partial \boldsymbol{\omega}^{\top}} \end{pmatrix}, \quad (8.29)$$

sendo uma matriz $(2p) \times 2N$, onde a derivada de $\mathbf{X}_{\omega}^{\top}$ com relação a $\boldsymbol{\omega}^{\top}$ é uma matriz $p \times N$ de zeros, exceto a k -ésima linha que é composta pela constante s_{z_k} , com $N = nt$.

A seguir, apresentamos as derivadas de \mathbf{A}_{ω} , $\boldsymbol{\Omega}_{\omega}$ e \mathbf{b}_{ω} todas com relação a $\boldsymbol{\omega}^{\top}$, para a Beta.

BETA De acordo com a equação de estimação definida em (5.11), temos:

$$\boldsymbol{\Lambda}_{\omega i} = \begin{pmatrix} \mathbf{G}_{\omega i} \boldsymbol{\Phi}_{\omega i} \mathbf{A}_{\omega i} \\ \mathbf{F}_{\omega i} \mathbf{C}_{\omega i} \end{pmatrix}, \quad \boldsymbol{\Omega}_{\omega i} = \mathbf{A}_{\omega i}^{1/2} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{A}_{\omega i}^{1/2}, \quad \text{e} \quad \mathbf{b}_{\omega i} = \mathbf{y}_i^* - \boldsymbol{\mu}_{\omega i}^*, \quad (8.30)$$

sendo $\mathbf{G}_{\omega i} = \text{diag}(\partial g^{-1}(\eta_{\omega i 1})/\partial \eta_{\omega i 1}, \dots, \partial g^{-1}(\eta_{\omega i t})/\partial \eta_{\omega i t})$, $\boldsymbol{\Phi}_{\omega i} = \text{diag}(\phi_{\omega i 1}, \dots, \phi_{\omega i t})$, $\mathbf{A}_{\omega i} = \text{diag}(\mathbf{a}_{\omega i 1}, \dots, \mathbf{a}_{\omega i t})$, $\mathbf{F}_{\omega i} = \text{diag}(\partial f^{-1}(\delta_{\omega i 1})/\partial \delta_{\omega i 1}, \dots, \partial f^{-1}(\delta_{\omega i t})/\partial \delta_{\omega i t})$, $\mathbf{C}_{\omega i} = \text{diag}(c_{\omega i 1}, \dots, c_{\omega i t})$ e $\boldsymbol{\mu}_{\omega i}^* = (\mu_{\omega i 1}^*, \dots, \mu_{\omega i t}^*)^{\top}$, em que $\mathbf{a}_{\omega ij} = \psi'(\mu_{\omega ij} \phi_{\omega ij}) + \psi'((1 - \mu_{\omega ij}) \phi_{\omega ij})$, $c_{\omega ij} = \mu_{\omega ij} \psi'(\mu_{\omega ij} \phi_{\omega ij}) - (1 - \mu_{\omega ij}) \psi'((1 - \mu_{\omega ij}) \phi_{\omega ij})$ e $\mu_{\omega ij}^* = \psi(\mu_{\omega ij} \phi_{\omega ij}) - \psi((1 - \mu_{\omega ij}) \phi_{\omega ij})$, com $i = 1, \dots, n$ e $j = 1, \dots, t$. Assim, obtêm-se:

$$\frac{\partial \boldsymbol{\Lambda}_{\omega i}}{\partial \boldsymbol{\omega}_i^{\top}} = \begin{pmatrix} \beta_k s_{z_k} \boldsymbol{\Phi}_{\omega i} \left[\dot{\mathbf{G}}_{\omega i} \mathbf{A}_{\omega i} + \mathbf{G}_{\omega i}^2 \mathcal{E}_{1\omega i} \right] + \gamma_k s_{z_k} \mathbf{G}_{\omega i} \mathbf{F}_{\omega i} \left[\mathbf{A}_{\omega i} + \boldsymbol{\Phi}_{\omega i} \mathcal{E}_{2\omega i} \right] \\ \beta_k s_{z_k} \mathbf{F}_{\omega i} \mathbf{C}_{\omega i} \left[\mathbf{A}_{\omega i} + \mathcal{M}_{\omega i} \mathcal{E}_{1\omega i} \right] + \gamma_k s_{z_k} \left[\dot{\mathbf{F}}_{\omega i} \mathbf{C}_{\omega i} - \mathbf{F}_{\omega i}^2 (\mathbf{I}_t - \mathcal{M}_{\omega i}) \mathcal{E}_{2\omega i} \right] \end{pmatrix},$$

$$\frac{\partial \boldsymbol{\Omega}_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} = \frac{1}{2} \left[\mathbf{A}_{\omega_i}^{1/2} \mathbf{R}(\boldsymbol{\alpha}) \mathcal{A}_{\omega_i} + \mathcal{A}_{\omega_i} \mathbf{R}(\boldsymbol{\alpha}) \mathbf{A}_{\omega_i}^{1/2} \right]$$

e

$$\frac{\partial \mathbf{b}_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} = - [\beta_k s_{z_k} \mathbf{G}_{\omega_i} \boldsymbol{\Phi}_{\omega_i} \mathbf{A}_{\omega_i} + \gamma_k s_{z_k} \mathbf{F}_{\omega_i} \boldsymbol{\mathcal{E}}_{3\omega_i}],$$

sendo $\mathcal{A}_{\omega_i} = \mathbf{A}_{\omega_i}^{-1/2} [\beta_k s_{z_k} \mathbf{G}_{\omega_i} \boldsymbol{\mathcal{E}}_{1\omega_i} + \gamma_k s_{z_k} \mathbf{F}_{\omega_i} \boldsymbol{\mathcal{E}}_{2\omega_i}]$, $\boldsymbol{\mathcal{E}}_{1\omega_i} = \text{diag}(\boldsymbol{\mathcal{E}}_{1\omega_i1}, \dots, \boldsymbol{\mathcal{E}}_{1\omega_it})$, $\boldsymbol{\mathcal{E}}_{2\omega_i} = \text{diag}(\boldsymbol{\mathcal{E}}_{2\omega_i1}, \dots, \boldsymbol{\mathcal{E}}_{2\omega_it})$, $\boldsymbol{\mathcal{E}}_{3\omega_i} = \text{diag}(\boldsymbol{\mathcal{E}}_{3\omega_i1}, \dots, \boldsymbol{\mathcal{E}}_{3\omega_it})$, $\mathcal{M}_{\omega_i} = \text{diag}(\mu_{\omega_i1}, \dots, \mu_{\omega_it})$, $\dot{\mathbf{G}}_{\omega_i} = \text{diag}(\partial^2 g^{-1}(\eta_{\omega_i1})/\partial \eta_{\omega_i1}^2, \dots, \partial^2 g^{-1}(\eta_{\omega_it})/\partial \eta_{\omega_it}^2)$, $\dot{\mathbf{F}}_{\omega_i} = \text{diag}(\partial^2 f^{-1}(\delta_{\omega_i1})/\partial \delta_{\omega_i1}^2, \dots, \partial^2 f^{-1}(\delta_{\omega_it})/\partial \delta_{\omega_it}^2)$ e \mathbf{I}_t uma matriz identidade de dimensão $t \times t$, em que

$$\boldsymbol{\mathcal{E}}_{1\omega_it} = \phi_{\omega_ij} [\psi''(\mu_{\omega_ij} \phi_{\omega_ij}) - \psi''((1 - \mu_{\omega_ij}) \phi_{\omega_ij})],$$

$$\boldsymbol{\mathcal{E}}_{2\omega_it} = [\mu_{\omega_ij} \psi''(\mu_{\omega_ij} \phi_{\omega_ij}) + ((1 - \mu_{\omega_ij}) \psi''((1 - \mu_{\omega_ij}) \phi_{\omega_ij}))],$$

e

$$\boldsymbol{\mathcal{E}}_{3\omega_it} = [\mu_{\omega_ij} \psi'(\mu_{\omega_ij} \phi_{\omega_ij}) - ((1 - \mu_{\omega_ij}) \psi'((1 - \mu_{\omega_ij}) \phi_{\omega_ij}))],$$

com $i = 1, \dots, n$ e $j = 1, \dots, t$.

8.4.6 A k' -ésima coluna da matriz \mathbf{Q} igual à k -ésima coluna da matriz \mathbf{X}

Considerando que a matriz de covariáveis \mathbf{Q} que modela a precisão tem a k' -ésima coluna igual a k -ésima coluna da matriz \mathbf{X} , a qual é perturbada segundo o esquema definido em (8.14), a matriz de covariáveis \mathbf{M} modificada é dada por

$$\mathbf{M}_\omega^\top = \begin{pmatrix} \mathbf{X}_\omega^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_\omega^\top \end{pmatrix},$$

em que o índice ω indica que as matrizes \mathbf{X} e \mathbf{Q} dependem da perturbação definida em (8.14). Logo, temos

$$\frac{\partial \mathbf{M}_\omega^\top}{\partial \boldsymbol{\omega}^\top} = \begin{pmatrix} \frac{\partial \mathbf{X}_\omega^\top}{\partial \boldsymbol{\omega}^\top} & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{Q}_\omega^\top}{\partial \boldsymbol{\omega}^\top} \end{pmatrix},$$

sendo uma matriz $(p + q) \times 2N$, em que a derivada de \mathbf{X}_ω^\top com relação à $\boldsymbol{\omega}^\top$ é uma matriz $p \times N$ de zeros exceto a k -ésima linha que é composta pela constante s_{x_k} e a derivada de \mathbf{Q}_ω^\top com relação à $\boldsymbol{\omega}^\top$ é uma matriz $q \times N$ de zeros exceto a k' -ésima linha que é composta pela constante s_{x_k} , com $N = nt$.

BETA De acordo com a equação de estimação definida em (5.11), as matrizes $\boldsymbol{\Lambda}_{\omega_i}$ e $\boldsymbol{\Omega}_{\omega_i}$ e o vetor \mathbf{b}_{ω_i} dados anteriormente. Assim, da Seção 8.4.5, obtemos

$$\frac{\partial \mathbf{A}_{\omega_i}}{\partial \boldsymbol{\omega}_i^\top} = \begin{pmatrix} \beta_k s_{x_k} \boldsymbol{\Phi}_{\omega_i} \left[\dot{\mathbf{G}}_{\omega_i} \mathbf{A}_{\omega_i} + \mathbf{G}_{\omega_i}^2 \boldsymbol{\mathcal{E}}_{1\omega_i} \right] + \gamma_{k'} s_{x_k} \mathbf{G}_{\omega_i} \mathbf{F}_{\omega_i} \left[\mathbf{A}_{\omega_i} + \boldsymbol{\Phi}_{\omega_i} \boldsymbol{\mathcal{E}}_{2\omega_i} \right] \\ \beta_k s_{x_k} \mathbf{F}_{\omega_i} \mathbf{C}_{\omega_i} \left[\mathbf{A}_{\omega_i} + \mathcal{M}_{\omega_i} \boldsymbol{\mathcal{E}}_{1\omega_i} \right] + \gamma_{k'} s_{x_k} \left[\dot{\mathbf{F}}_{\omega_i} \mathbf{C}_{\omega_i} - \mathbf{F}_{\omega_i}^2 (\mathbf{I}_i - \mathcal{M}_{\omega_i}) \boldsymbol{\mathcal{E}}_{2\omega_i} \right] \end{pmatrix},$$

$$\frac{\partial \Omega_{\omega_i}}{\partial \omega_i^\top} = \frac{1}{2} \left[\mathbf{A}_{\omega_i}^{1/2} \mathbf{R}(\alpha) \mathcal{A}_{\omega_i} + \mathcal{A}_{\omega_i} \mathbf{R}(\alpha) \mathbf{A}_{\omega_i}^{1/2} \right] \quad \text{e}$$

e

$$\frac{\partial \mathbf{b}_{\omega_i}}{\partial \omega_i^\top} = - [\beta_k s_{x_k} \mathbf{G}_{\omega_i} \Phi_{\omega_i} \mathbf{A}_{\omega_i} + \gamma_{k'} s_{x_k} \mathbf{F}_{\omega_i} \mathcal{E}_{3\omega_i}],$$

em que $\mathcal{A}_{\omega_i} = \mathbf{A}_{\omega_i}^{-1/2} [\beta_k s_{x_k} \mathbf{G}_{\omega_i} \mathcal{E}_{1\omega_i} + \gamma_{k'} s_{x_k} \mathbf{F}_{\omega_i} \mathcal{E}_{2\omega_i}]$ e $\mathcal{E}_{1\omega_i}$, $\mathcal{E}_{2\omega_i}$, $\mathcal{E}_{3\omega_i}$, \mathcal{M}_{ω_i} , $\dot{\mathbf{G}}_{\omega_i}$ e $\dot{\mathbf{F}}_{\omega_i}$ estão descritos na Seção 8.4.5 para o caso Beta, com $i = 1, \dots, n$ e $j = 1, \dots, t$.

8.4.7 A k' -ésima coluna da matriz \mathbf{Q} função da k -ésima coluna da matriz \mathbf{X}

Considere que a matriz de covariáveis \mathbf{Q} que modela a precisão tem a k' -ésima coluna função da k -ésima coluna da matriz \mathbf{X} , ou seja, $q_{wij k'} = h(x_{wij k})$, com $i = 1, \dots, n$, $j = 1, \dots, t$, $k = 2, \dots, p$ e $k' = 2, \dots, q$, em que h é uma função contínua e diferenciável.

Assim, a matriz de covariáveis \mathbf{M} modificada segundo o esquema definido em (8.14) pode ser expressa como

$$\frac{\partial \mathbf{M}_{\omega}^\top}{\partial \omega^\top} = \begin{pmatrix} \frac{\partial \mathbf{X}_{\psi}^\top}{\partial \omega^\top} & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{Q}_{\psi}^\top}{\partial \omega^\top} \end{pmatrix},$$

sendo uma matriz $(p+q) \times 2N$, em que a derivada de \mathbf{X}_{ω}^\top com relação a ω^\top e uma matriz $p \times N$ de zeros exceto a k -ésima linha que é composta pela constante s_{x_k} e a derivada de \mathbf{Q}_{ω}^\top com relação a ω^\top e uma matriz $q \times N$ de zeros exceto a k' -ésima linha que é composta pelo vetor $\dot{h}(\mathbf{x}_{\omega k}) s_{x_k}$, em que $N = nt$ e $\dot{h}(\mathbf{x}_{\omega k}) = (\dot{h}(\mathbf{x}_{\omega 11k}), \dot{h}(\mathbf{x}_{\omega 12k}), \dots, \dot{h}(\mathbf{x}_{\omega t k}))$, com $\dot{h}(\mathbf{x}_{\omega i j k}) = \partial h(\mathbf{x}_{\omega i j k}) / \partial x_{\omega i j k}$, com $i = 1, \dots, n$, $j = 1, \dots, t$ e $k = 2, \dots, p$.

BETA De acordo com a equação de estimação definida em (5.11), as matrizes Λ_{ω_i} e Ω_{ω_i} e o vetor \mathbf{b}_{ω_i} são dados anteriormente. Assim, da Seção 8.4.5, obtemos

$$\frac{\partial \Lambda_{\omega_i}}{\partial \omega_i^\top} = \begin{pmatrix} \beta_k s_{x_k} \Phi_{\omega_i} \left[\dot{\mathbf{G}}_{\omega_i} \mathbf{A}_{\omega_i} + \mathbf{G}_{\omega_i}^2 \mathcal{E}_{1\omega_i} \right] + \gamma_{k'} s_{x_k} \dot{\mathcal{H}}_{\omega_i} \mathbf{G}_{\omega_i} \mathbf{F}_{\omega_i} [\mathbf{A}_{\omega_i} + \Phi_{\omega_i} \mathcal{E}_{2\omega_i}] \\ \beta_k s_{x_k} \mathbf{F}_{\omega_i} \mathbf{C}_{\omega_i} [\mathbf{A}_{\omega_i} + \mathcal{M}_{\omega_i} \mathcal{E}_{1\omega_i}] + \gamma_{k'} s_{x_k} \dot{\mathcal{H}}_{\omega_i} \left[\dot{\mathbf{F}}_{\omega_i} \mathbf{C}_{\omega_i} - \mathbf{F}_{\omega_i}^2 (\mathbf{I}_t - \mathcal{M}_{\omega_i}) \mathcal{E}_{2\omega_i} \right] \end{pmatrix},$$

$$\frac{\partial \Omega_{\omega_i}}{\partial \omega_i^\top} = \frac{1}{2} \left[\mathbf{A}_{\omega_i}^{1/2} \mathbf{R}(\alpha) \mathcal{A}_{\omega_i} + \mathcal{A}_{\omega_i} \mathbf{R}(\alpha) \mathbf{A}_{\omega_i}^{1/2} \right] \quad \text{e}$$

$$\frac{\partial \mathbf{b}_{\omega_i}}{\partial \omega_i^\top} = - \left[\beta_k s_{x_k} \mathbf{G}_{\omega_i} \Phi_{\omega_i} \mathbf{A}_{\omega_i} + \gamma_{k'} s_{x_k} \dot{\mathcal{H}}_{\omega_i} \mathbf{F}_{\omega_i} \mathcal{E}_{3\omega_i} \right],$$

em que $\mathcal{A}_{\omega_i} = \mathbf{A}_{\omega_i}^{-1/2} [\beta_k s_{x_k} \mathbf{G}_{\omega_i} \mathcal{E}_{1\omega_i} + \gamma_{k'} s_{x_k} \dot{\mathcal{H}}_{\omega_i} \mathbf{F}_{\omega_i} \mathcal{E}_{2\omega_i}]$ e $\mathcal{E}_{1\omega_i}$, $\mathcal{E}_{2\omega_i}$, $\mathcal{E}_{3\omega_i}$, \mathcal{M}_{ω_i} , $\dot{\mathbf{G}}_{\omega_i}$ e $\dot{\mathbf{F}}_{\omega_i}$ estão descritos na Seção 8.4.5 para o caso Beta, com $i = 1, \dots, n$ e $j = 1, \dots, t$.

Referências

- BATES, D. M.; WATTS, D. G. Relative curvature measures of nonlinearity. *Journal of the Royal Statistical Society: Series B (Methodological)*, v. 42, n. 1, p. 1–25, 1980.
- BURY, K. *Statistical Distributions in Engineering*. Cambridge: Cambridge University Press, 1999.
- CADIGAN, N. G. A generalized linear model for fishery data. *Canadian Journal of Fisheries and Aquatic Sciences*, v. 52, n. 1, p. 1–10, 1995.
- CADIGAN, N. G.; FARRELL, P. J. Fishery stock assessment: Statistical developments. *Canadian Journal of Statistics*, v. 30, n. 4, p. 567–584, 2002.
- COLLETT, D. *Modelling Binary Data*. 2. ed. Boca Raton: Chapman & Hall/CRC, 2003.
- COOK, R. D. Detection of influential observation in linear regression. *Technometrics*, v. 19, n. 1, p. 15–18, 1977.
- COOK, R. D. Assessment of local influence. *Journal of the Royal Statistical Society: Series B (Methodological)*, v. 48, n. 2, p. 133–169, 1986.
- COX, D. R.; SNELL, E. J. *Analysis of Binary Data*. 2. ed. London: Chapman & Hall, 1989.
- CRIBARI-NETO, F.; ZEILEIS, A. Beta regression in R. *Journal of Statistical Software*, v. 34, n. 2, p. 1–24, 2010.
- FERRARI, S. L. P.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, v. 31, n. 7, p. 799–815, 2004.
- HARVILLE, D. A. *Matrix Algebra from a Statistician's Perspective*. [S.l.]: Springer, 1997.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous Univariate Distributions*. 2. ed. New York: John Wiley & Sons, 1995. v. 2.
- JOHNSTON, J.; DINARDO, J. *Econometric Methods*. 4. ed. New York: McGraw-Hill, 1996.
- KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. *The Annals of Mathematical Statistics*, v. 22, n. 1, p. 79–86, 1951.
- LIANG, K. Y.; ZEGER, S. L. Longitudinal data analysis using generalized linear models. *Biometrika*, v. 73, n. 1, p. 13–22, 1986.
- MCCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. [S.l.]: Springer, 1989.
- NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, v. 135, n. 3, p. 370–384, 1972.
- OSPINA, R. *Metodologias para Modelagem de Dados Proporcionalis*. Tese (Doutorado) — Universidade de São Paulo, 2007.
- PAN, W. Akaike's information criterion in generalized estimating equations. *Biometrics*, v. 57, n. 1, p. 120–125, 2001.

- PREGIBON, D. Logistic regression diagnostics. *The Annals of Statistics*, v. 9, n. 4, p. 705–724, 1981.
- PRENTICE, R. L.; ZHAO, L. P. Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics*, v. 47, n. 3, p. 825–839, 1991.
- SMITHSON, M.; VERKUILEN, J. A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, v. 11, n. 1, p. 54–71, 2006.
- THOMAS, W.; COOK, R. D. Assessing influence on predictions from generalized linear models. *Technometrics*, v. 32, n. 1, p. 61–68, 1990.
- VENEZUELA, M. K. *Equação de estimação generalizada e influência local para modelos de regressão beta com medidas repetidas*. Tese (Doutorado em Estatística) — Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2008. Doi:10.11606/T.45.2008.tde-10072008-210246.
- WEDDERBURN, R. W. M. Quasi-likelihood functions, generalized linear models, and the gauss-newton method. *Biometrika*, v. 61, n. 3, p. 439–447, 1974.