



# EEGs em Modelos de Regressão Beta para dados de medidas repetidas.

Prof. Dr. Juvêncio Santos Nobre

Douglas Chaves Moura

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada  
26 de fevereiro de 2025

# Visão geral

1. Introdução
2. Distribuição Beta
3. Regressão Beta
4. EEGs
5. Aplicação na Regressão Beta
6. Diagnóstico
7. Referências



# Introdução

---

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada

# Introdução

- **Estudos Longitudinais:**
  - Coleta sequencial de dados ao longo do tempo para os mesmos indivíduos.
  - Permitem acompanhar a evolução dos sujeitos.
  - Desafios analíticos devido à correlação entre observações de um mesmo indivíduo.
- **Desafios:**
  - Não aleatorização na ordem das observações.
  - Espaçamentos irregulares entre medições.
  - Ausência de dados em determinados momentos.

# Motivação

- **Equações de Estimação Generalizadas (EEGs):**
  - Proposta por Liang e Zeger (1986).
  - Baseadas em modelos de quasi-verossimilhança (Wedderburn, 1974).
  - Focam nas distribuições marginais da variável resposta.
- **Modelo de Regressão Beta:**
  - Útil para variáveis contínuas no intervalo  $(0,1)$  (Ferrari e Cribari-Neto, 2004).
  - Parametrização com um componente sistemático e um parâmetro de precisão.
  - Captura a variabilidade esperada para respostas restritas ao intervalo  $(0,1)$ .
- **Objetivo:**
  - Apresentar a distribuição beta e a regressão beta.
  - Explorar a aplicação das EEGs em modelos de regressão beta para dados longitudinais.
  - Discutir os princípios teóricos das EEGs no contexto da regressão beta.



# Distribuição Beta

---

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada

# Distribuição Beta

## Definição:

- Flexível para modelagem de proporções.
- Função densidade de probabilidade:

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

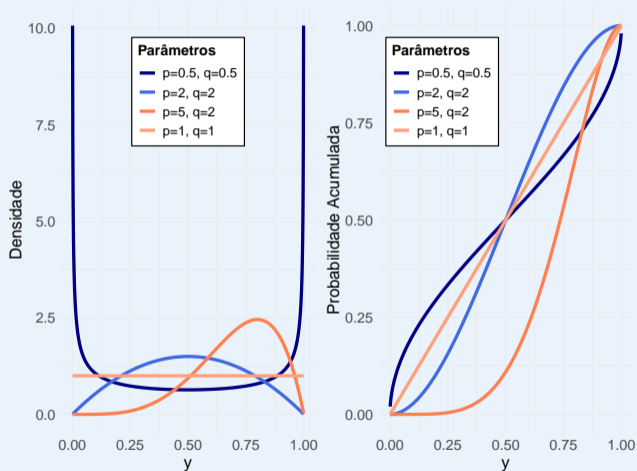
- Parâmetros:  $p > 0, q > 0$
- Função gama ( $\Gamma$ ):  $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ ,
- Para números naturais:  $\Gamma(z) = (z-1)!$

# Distribuição Beta

## Propriedades dos Parâmetros:

- Valores maiores de  $p$  resultam em maior concentração de  $y$  próximos a 1.
- Valores menores de  $p$  resultam em maior concentração de  $y$  próximos a 0.
- Distribuição simétrica quando  $p = q$ .

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



# Distribuição Beta

## 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  $y = \frac{p-1}{p+q-2}$ .

# Distribuição Beta

Medidas estatísticas:

Média

$$\mathbb{E}(y) = \frac{p}{p + q} \quad (2)$$

Variância

$$\text{Var}(y) = \frac{pq}{(p + q)^2(p + q + 1)}. \quad (3)$$

# Distribuição Beta

## Moda

Para  $(p, q > 1)$ :

$$\text{Mo}(y) = \frac{p - q}{p + q - 2}$$

Já nos casos onde  $(p, q < 1)$  se tem o que é chamado de anti-moda.



# Regressão Beta

---

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada

# Motivação/Proposta da Regressão Beta

- **Versatilidade da Distribuição Beta:**
  - Modelagem eficaz de uma ampla variedade de incertezas.
  - Utilização empírica em diversas aplicações práticas.
- **Limitações das Aplicações Tradicionais:**
  - Geralmente não exigem estrutura de regressão para a variável de interesse.
- **Proposta de Ferrari e Cribari-Neto (2004):**
  - Investigar o comportamento de uma variável Beta em relação a variáveis explicativas.
  - Construção de modelos preditivos mais simples e adequados para dados no intervalo unitário.

# Regressão Beta - Introdução e Parametrização

- **Parametrização Tradicional:**
  - Dois parâmetros de forma:  $p$  e  $q$ .
- **Nova Parametrização:**
  - Parâmetro de posição ( $\mu$ ) e parâmetro de precisão ( $\phi$ ).
  - Relação entre  $p$  e  $q$  com  $\mu$  e  $\phi$ :

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

- Expressão de  $p$  e  $q$  em função de  $\mu$  e  $\phi$ :

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

# Regressão Beta

Medidas estatísticas:

Média

$$\mathbb{E}(y) = \mu \quad (4)$$

Variância

$$\text{Var}(y) = \frac{\mu(1 - \mu)}{1 + \phi}. \quad (5)$$

**OBS.:**  $\phi$  controla a dispersão dos dados: valores maiores de  $\phi$  reduzem a variância.

# Regressão Beta - Densidade de Probabilidade

## Densidade de Probabilidade

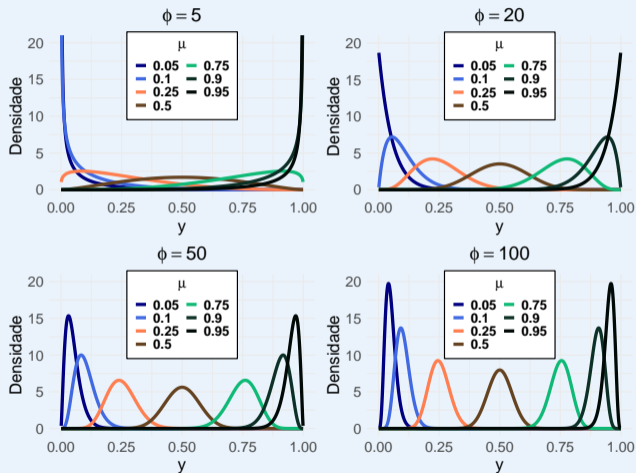
$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1 \quad (6)$$

- Para evitar problemas com valores extremos (próximos de 0 ou 1), Cribari-Neto e Zeileis (2010) sugerem a transformação:

$$y^* = (y \cdot (n - 1) + 0.5) / n$$

onde  $n$  é o tamanho da amostra, (SMITHSON; VERKUILEN, 2006).

Figura 2: Efeito do aumento da precisão ( $\phi$ ) na forma da distribuição Beta reparametrizada para diferentes valores de  $\mu$ .



# Regressão Beta - Função de Ligação

## Modelagem da Média:

- Média  $\mu_t$  como combinação linear de covariáveis:

$$g(\mu_t) = \sum_{i=1}^k x_{ti}\beta_i = \eta_t, \quad (7)$$

- $\eta_t$  é o preditor linear.
- $g(\cdot)$  deve ser monotônica e diferenciável.

# Regressão Beta - Função de Ligação

Funções de Ligação Comuns:

- Ligação *Logito*:

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

- Ligação *Probit*:

$$g(\mu_t) = \Phi^{-1}(\mu_t), \quad \mu_t = \Phi(x_t^\top \beta).$$

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

# Regressão Beta - Verossimilhança e Estimação

## Logaritmo da Função de Verossimilhança

$$\ell(\boldsymbol{\beta}, \phi) = \sum_{t=1}^n [\log \Gamma(\phi) - \log \Gamma(\mu_t \phi) - \log \Gamma((1 - \mu_t) \phi) + (\mu_t \phi - 1) \log y_t + \{(1 - \mu_t) \phi - 1\} \log(1 - y_t)] \quad (8)$$

Essa função é maximizada para obter os estimadores de máxima verossimilhança (EMV) de  $\boldsymbol{\beta}$  e  $\phi$ .

# Regressão Beta - Estimação dos Parâmetros

- Equações de Verossimilhança:

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

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

- $y_t^* = \log(y_t/(1 - y_t))$ ,  $\mu_t^* = \psi(\mu_t\phi) - \psi((1 - \mu_t)\phi)$ , e  $\psi(\cdot)$  é a função digama.
- Solução requer métodos numéricos como Newton-Raphson.

# Regressão Beta - Matriz de Informação de Fisher

- Matriz de Informação de Fisher:

$$K(\beta, \phi) = \begin{pmatrix} \phi X^T W X & X^T T c \\ c^T T^T X & \text{tr}(D) \end{pmatrix} \quad (9)$$

- Componentes da matriz:
  - $W = \text{diag}\{w_1, \dots, w_n\}$ , com:

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

- $c = (c_1, \dots, c_n)^T$ , com:

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

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

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

# Regressão Beta - Matriz Inversa de Fisher

- Matriz Inversa de Fisher:

$$K^{-1}(\beta, \phi) \quad (10)$$

mais detalhes em Ferrari e Cribari-Neto (2004).

- Permite calcular os erros padrão assintóticos dos estimadores de máxima verossimilhança.
- Sob condições usuais de regularidade para estimativa de máxima verossimilhança, quando o tamanho da amostra é grande:

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



# EEGs

---

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada

# Equações de Estimação Generalizadas (EEGs)

- **Contexto:**
  - MLGs e Regressão Beta pressupõem independência entre observações.
  - Estudos longitudinais exigem métodos que incorporem correlação intraunidade.
- **Introdução às EEGs:**
  - Introduzidas por Liang e Zeger (1986).
  - Focam nas distribuições marginais da variável resposta.
  - Baseadas em quasi-verossimilhança.
  - Produzem estimadores consistentes e assintoticamente normais.

# Formulação das Equações de Estimação

- **Função de Ligação:** Conecta a média marginal  $\mu_{ij}$  ao preditor linear  $\eta_{ij}$  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  $\phi$  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.

# Formulação das Equações de Estimação

## Sistema de Equações:

- É formalizado por:

$$\sum_{i=1}^n D_i^T V_i^{-1} (y_i - \boldsymbol{\mu}_i) = 0,$$

sendo,

- $D_i = X_i G_i$  é a matriz de sensibilidade, com  $G_i = \text{diag}(g'(\mu_{i1}), \dots, g'(\mu_{iT_i}))$ .
- $V_i^{-1}$  pondera os resíduos considerando a estrutura de dependência.

# Estrutura de Covariância de Trabalho

## Construção da Matriz de Covariância:

- A matriz  $\Omega_{W_i}$  é construída como:

$$\Omega_{W_i}(\alpha, \phi) = \phi A_i^{1/2} R_W(\alpha) A_i^{1/2}$$

onde  $R_W(\alpha)$  é uma matriz de correlação parametrizada.

- Quando  $R_W(\alpha)$  coincide com a matriz de correlação verdadeira, temos:

$$\Omega_{W_i} = \text{Var}(y_i).$$

- Eficiência vs. Consistência:
  - A escolha de  $R_W(\alpha)$  afeta a eficiência.
  - Não afeta a consistência dos estimadores.

# Estimação dos Parâmetros

## Algoritmo Iterativo:

- Os estimadores de  $\beta$  são obtidos sob a seguinte iteração:

$$\hat{\beta}^{(h+1)} = \hat{\beta}^{(h)} + \left[ \sum_{i=1}^n X_i^\top \hat{\Delta}_i^{(h)} [\Omega_{W_i}(\hat{\theta}^{(h)})]^{-1} \hat{\Delta}_i^{(h)} X_i \right]^{-1} \times \left[ \sum_{i=1}^n X_i^\top \hat{\Delta}_i^{(h)} [\Omega_{W_i}(\hat{\theta}^{(h)})]^{-1} [y_i - \mu_i(X_i^\top \hat{\beta}^{(h)})] \right] \quad (11)$$

onde,

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

# Distribuição Assintótica

- Sob condições de regularidade,

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

em que

- $A^{-1} = \lim_{n \rightarrow \infty} nS^{-1} [\sum_{i=1}^n X_i^T \Delta_i [\Omega_{W_i}(\theta)]^{-1} V(y_i) [\Omega_{W_i}(\theta)]^{-1} \Delta_i X_i] S^{-1}$

- Estimador Sanduíche:**

$$A^{-1} = S^{-1} B S^{-1}$$

- $S = \sum_{i=1}^n X_i^T \Delta_i \Omega_{W_i}^{-1} \Delta_i X_i$  representa a informação de Fisher observada.
- $B = \sum_{i=1}^n X_i^T \Delta_i \Omega_{W_i}^{-1} V(y_i) \Omega_{W_i}^{-1} \Delta_i X_i$  captura a variabilidade empírica.

# Abordagens EEG1 vs. EEG2

- **EEG1 (Liang e Zeger, 1986):**
  - Trata  $\alpha$  como parâmetro de perturbação.
  - Mantém consistência mesmo com  $R_W(\alpha)$  mal especificada.
  - Possível perda de eficiência.
- **EEG2 (Prentice e Zhao, 1991):**
  - Estima conjuntamente  $\beta$  e  $\alpha$ .
  - Exige correta especificação de  $R_W(\alpha)$  para eficiência ótima.
  - Preferível quando a estrutura de dependência é de interesse científico.



# Aplicação na Regressão Beta

---

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada

# Aplicação das EEGs em Modelos de Regressão Beta

## Modelagem do Parâmetro de Posição:

- Suponha que a resposta  $y_{ij}$  siga uma distribuição beta:

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

sua densidade é dada por:

## Densidade de Probabilidade

$$p(y_{ij}; \mu_{ij}, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu_{ij}\phi)\Gamma((1-\mu_{ij})\phi)} y_{ij}^{\mu_{ij}\phi-1} (1-y_{ij})^{(1-\mu_{ij})\phi-1}. \quad (12)$$

# Modelagem do Parâmetro de Posição

## Função de Ligação e Componente Sistemático:

- Modelagem da média  $\mu_{ij}$  por meio de uma função de ligação  $g(\cdot)$ :

$$g(\mu_{ij}) = \mathbf{x}_{ij}^{\top} \boldsymbol{\beta} = \eta_{ij} \quad (13)$$

- $\mathbf{x}_{ij}$ : vetor de covariáveis.
- $\boldsymbol{\beta}$ : vetor de parâmetros desconhecidos.

# Modelagem do Parâmetro de Posição

## Definição das Funções de Estimação:

- Definição simples:

$$b_i = y_i - \mu_i.$$

Quando as observações são independentes, os vetores  $b_i$  não pertencem à classe  $\mathcal{L}(b)$ .

- Transformações para garantir propriedades regulares:

$$b_i = y_i^* - \mu_i^*$$

onde  $y_i^*$  e  $\mu_i^*$  são transformações das observações e da média, respectivamente. Essas transformações garantem que os vetores tenham média zero, sejam independentes e incluam a função escore do modelo beta clássico.

# Modelagem do Parâmetro de Posição

## Equações de Estimação para $\beta$ :

- Dessa forma, as equações de estimação para o parâmetro  $\beta$  são formuladas a partir da densidade marginal de  $y_{ij}$  (Equação 12) e do componente sistemático (Equação 13), adotando a definição:

$$\Psi_1^o(\beta) = \sum_{i=1}^n X_i^T \Lambda_i \text{Cov}(b_i)^{-1} (y_i^* - \mu_i^*).$$

- $X_i$ : matriz das covariáveis.
- $\Lambda_i$ : matriz de sensibilidade.
- $\text{Cov}(b_i)$ : matriz de covariância dos vetores  $b_i$ .

# Modelagem do Parâmetro de Posição

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

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathcal{D}} \mathcal{N}_p \left( 0, \lim_{n \rightarrow \infty} n \left( \sum_{i=1}^n X_i^\top \Lambda_i \text{Cov}(b_i)^{-1} \Lambda_i X_i \right)^{-1} \right).$$

- 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,  $R(\alpha)$ , 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  $\alpha$  de dimensão  $s \times 1$  e não precisa coincidir com a verdadeira matriz de correlação dos  $y_i^*$ .

# Modelagem do Parâmetro de Posição

**Matriz de Correlação de Trabalho  $R(\alpha)$ :**

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

$$\Psi_1(\beta) = \sum_{i=1}^n X_i^\top \Lambda_i \Omega_i^{-1} (y_i^* - \mu_i^*) = \sum_{i=1}^n X_i^\top W_i \Lambda_i^{-1} b_i, \quad (14)$$

- $\Omega_i = A_i^{1/2} R(\alpha) A_i^{1/2}$  é a matriz de covariância de trabalho.
- $W_i = \Lambda_i \Omega_i^{-1} \Lambda_i$  são os pesos.

# Modelagem do Parâmetro de Posição

- Nota-se que, uma vez que a especificação de  $R(\alpha)$  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  $\alpha$  e para  $\phi$ . Quando tais parâmetros adicionais são incorporados ao procedimento, o estimador de  $\beta$  mantém propriedades desejáveis (como consistência e normalidade assintótica), desde que as condições do teorema 1 sejam satisfeitas (VENEZUELA, 2008).

# Modelagem do Parâmetro de Posição

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

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

- Quando a matriz de correlação de trabalho  $R(\alpha)$  coincide com a verdadeira matriz de correlação dos  $y_i^*$ , o estimador robusto se reduz à forma «*naive*» ou «*model-based*»:

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

# Modelagem do Parâmetro de Posição

## Processo Iterativo para Estimação de $\beta$ , $\phi$ e $\alpha$ :

- O processo iterativo proposto para a estimação dos parâmetros  $\beta$ ,  $\phi$  e  $\alpha$  combina o método *scoring* de Fisher para  $\beta$  com o método dos momentos para  $\alpha$  e  $\phi$ .

# Modelagem do Parâmetro de Posição

## Estimação de $\beta$ :

- Expandindo a função de estimação  $\Psi_1(\beta)$  em torno de um valor inicial  $\hat{\beta}^{(0)}$ , obtém-se a atualização iterativa dada por:

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} - \left\{ E \left[ \frac{\partial \Psi_1(\hat{\beta}^{(m)})}{\partial \beta^\top} \right] \right\}^{-1} \Psi_1(\hat{\beta}^{(m)}) = \hat{\beta}^{(m)} + \left\{ \left[ \sum_{i=1}^n X_i^\top \hat{W}_i X_i \right]^{-1} \left[ \sum_{i=1}^n X_i^\top \hat{W}_i \hat{\Lambda}_i^{-1} \hat{b}_i \right] \right\}, \quad (15)$$

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  $\beta$ ,  $\phi$  e  $\alpha$ .

# Modelagem do Parâmetro de Posição

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

$$\hat{\beta}^{(m+1)} = \left\{ \left[ \sum_{i=1}^n X_i^T \hat{W}_i X_i \right]^{-1} \left[ \sum_{i=1}^n X_i^T \hat{W}_i z_i \right] \right\}^{(m)}, \quad (16)$$

em que a variável dependente modificada é definida por

$$z_i = \hat{\eta}_i + \hat{\Lambda}_i^{-1} \hat{b}_i.$$

# Modelagem do Parâmetro de Posição

## Estimação de $\phi$ :

- Adicionalmente, utilizando o método dos momentos e considerando que

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

a estimação de  $\phi$  no  $m$ -ésimo passo é dada por:

$$\hat{\phi}^{(m)} = \left( \left\{ \sum_{i=1}^n \sum_{j=1}^t \left( \frac{y_{ij} - \hat{\mu}_{ij}^{(m)}}{\sqrt{\hat{\mu}_{ij}^{(m)}(1 - \hat{\mu}_{ij}^{(m)})}} \right)^2 / (nt - p) \right\}^{-1} - 1 \right). \quad (17)$$

# Modelagem do Parâmetro de Posição

## Estimação de $\alpha$ :

- O procedimento permite especificar diferentes estruturas de correlação para observações repetidas, com estimadores desenvolvidos para a matriz de correlação  $R(\alpha)$ . As configurações incluem:

# Modelagem do Parâmetro de Posição

## 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  $\alpha$  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  $\alpha$  é:

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

### Matriz de Correlação Não Estruturada:

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

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

# Modelagem do Parâmetro de Posição

## Procedimento Prático para Estimação:

- Inicialmente, supondo independência entre as observações dentro de cada unidade experimental, ajusta-se um modelo de regressão linear de  $g(y)$  sobre  $X$  mediante mínimos quadrados ordinários. Nesta etapa, os parâmetros  $\alpha$  e  $\phi$  não são considerados.
- Com as estimativas iniciais de  $\beta$  obtidas na etapa anterior, procede-se à estimação iterativa de  $\beta$  e  $\phi$  utilizando as equações (16) e a fórmula para  $\hat{\phi}$ .
- Por fim, especifica-se uma estrutura de correlação de trabalho,  $R(\alpha)$ . Se essa estrutura for, por exemplo, a de independência, os valores finais de  $\beta$  e  $\phi$  correspondem aos obtidos na etapa anterior. Caso contrário, estima-se os parâmetros  $\alpha$  e, em seguida, reestima-se  $\beta$  e  $\phi$  de forma iterativa até que haja convergência.

# Aplicação das EEGs em Modelos de Regressão Beta

## Modelagem Conjunta dos Parâmetros de Posição e Precisão:

- Suponha que a resposta  $y_{ij}$  siga uma distribuição beta:

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

sua densidade é dada por:

### Densidade de Probabilidade

$$p(y_{ij}; \mu_{ij}, \phi_{ij}) = \frac{\Gamma(\phi_{ij})}{\Gamma(\mu_{ij}\phi_{ij})\Gamma((1-\mu_{ij})\phi_{ij})} y_{ij}^{\mu_{ij}\phi_{ij}-1} (1-y_{ij})^{(1-\mu_{ij})\phi_{ij}-1}. \quad (18)$$

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Componentes Sistemáticos para Média e Precisão:

- Consideramos que o componente sistemático que modela a média permaneça o mesmo, conforme a Equação (13):

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

- Para que o parâmetro de precisão também seja modelado, define-se um componente sistemático para  $\phi_{ij}$  como:

$$f(\phi_{ij}) = \mathbf{q}_{ij}^T \boldsymbol{\gamma} = \delta_{ij} \quad (19)$$

- $\mathbf{q}_{ij}$ : vetor de covariáveis.
- $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^T$ : vetor de parâmetros desconhecidos.

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Funções de Estimação para Modelagem Conjunta:

- A construção das equações de estimação para o modelo conjunto leva em consideração a densidade marginal em (18), os componentes sistemáticos (Equações (13) e (19)) e utiliza os vetores

$$b_i = y_i^* - \mu_i^*,$$

onde  $y_i^*$  e  $\mu_i^*$  são as versões transformadas que garantem as propriedades necessárias para a construção das funções de estimação.

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Construção das Equações de Estimação:

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

$$\psi_2^o(\theta) = \sum_{i=1}^n \begin{pmatrix} X_i^\top G_i \Phi_i A_i \\ Q_i^\top F_i C_i \end{pmatrix} \left[ A_i^{1/2} R(y_i^*) A_i^{1/2} \right]^{-1} (y_i^* - \mu_i^*), \quad (20)$$

sendo:

- $X_i = (x_{i1}, \dots, x_{it})^\top$ ;
- $\Phi_i = \text{diag}(\phi_{i1}, \dots, \phi_{it})$ ;
- $A_i = \text{diag}(a_{i1}, \dots, a_{it})$ , com  $a_{ij} = \psi'(\mu_{ij}\phi_{ij}) + \psi'((1 - \mu_{ij})\phi_{ij})$ ;
- $G_i = \text{diag}\left(\frac{\partial g^{-1}(\eta_{i1})}{\partial \eta_{i1}}, \dots, \frac{\partial g^{-1}(\eta_{it})}{\partial \eta_{it}}\right)$ ;
- $Q_i = (q_{i1}, \dots, q_{it})^\top$ ;
- $C_i = \text{diag}(c_{i1}, \dots, c_{it})$ , com  $c_{ij} = \mu_{ij}\psi'(\mu_{ij}\phi_{ij}) - (1 - \mu_{ij})\psi'((1 - \mu_{ij})\phi_{ij})$ ;
- $F_i = \text{diag}\left(\frac{\partial f^{-1}(\delta_{i1})}{\partial \delta_{i1}}, \dots, \frac{\partial f^{-1}(\delta_{it})}{\partial \delta_{it}}\right)$ ;
- $R(y_i^*)$  é a verdadeira matriz de correlação associada a  $y_i^*$ .

## Modelagem Conjunta dos Parâmetros de Posição e Precisão

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

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}_{p+q} \left( 0, \lim_{n \rightarrow \infty} n \left\{ \sum_{i=1}^n \begin{pmatrix} X_i^\top G_i \Phi_i A_i \\ Q_i^\top F_i C_i \end{pmatrix} \text{Cov}(b_i)^{-1} \begin{pmatrix} X_i^\top G_i \Phi_i A_i \\ Q_i^\top F_i C_i \end{pmatrix}^\top \right\}^{-1} \right).$$

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

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

Matriz de Correlação de Trabalho  $R(\alpha)$ :

- Logo,  $\theta$  tem sua função de estimação generalizada dada por:

$$\Psi_2(\theta) = \sum_{i=1}^n M_i^\top \Lambda_i \Omega_i^{-1} b_i = \sum_{i=1}^n M_i^\top W_i \Lambda_i^\top b_i \quad (21)$$

- $M_i = \begin{pmatrix} X_i & 0 \\ 0 & Q_i \end{pmatrix}$
- $\Lambda_i = \begin{pmatrix} G_i \Phi_i A_i \\ F_i C_i \end{pmatrix}$
- $\Omega_i = A_i^{1/2} R(\alpha) A_i^{1/2}$
- $W_i = A_i \Omega_i^{-1} A_i^\top$

## Modelagem Conjunta dos Parâmetros de Posição e Precisão

- O estimador robusto da matriz de covariância de  $\hat{\theta}$  é dado por

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

obtida substituindo  $\theta$  e  $\alpha$  por suas estimativas consistentes.

- Quando  $R(\alpha)$  coincide com a verdadeira matriz de correlação, o estimador robusto se reduz à forma «naive»:

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

## Modelagem Conjunta dos Parâmetros de Posição e Precisão

Processo Iterativo para Estimação de  $\beta$ ,  $\gamma$  e  $\alpha$ :

- O processo iterativo para calcular  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$  e  $\hat{\alpha}$ , combina o método *scoring* de Fisher para os parâmetros de regressão  $\theta = (\beta^\top, \gamma^\top)^\top$  com o método dos momentos para a estimação dos parâmetros de correlação  $\alpha$ .

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Estimação de $\theta$ :

- Expandindo a EEGs dada na Equação (21) em torno de um valor inicial  $\hat{\theta}^{(0)} = (\hat{\beta}^{(0)\top}, \hat{\gamma}^{(0)\top})^\top$ , o processo iterativo para atualizar  $\theta$  é dado por:

$$\hat{\theta}^{(m+1)} = \hat{\theta}^{(m)} + \left\{ \left[ \sum_{i=1}^n M_i^\top \hat{W}_i M_i \right]^{-1} \left[ \sum_{i=1}^n M_i^\top \hat{W}_i \hat{A}_i^{-\top} \hat{b}_i \right] \right\}^{(m)} \quad (22)$$

para  $m = 0, 1, 2, \dots$ , onde as matrizes  $M_i$  apresentam as derivadas dos preditores lineares em relação a  $\theta$ , e  $\hat{W}_i$  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**:

$$\hat{\theta}^{(m+1)} = \left\{ \left[ \sum_{i=1}^n M_i^T \hat{W}_i M_i \right]^{-1} \left[ \sum_{i=1}^n M_i^T \hat{W}_i z_i \right] \right\}^{(m)}, \quad (23)$$

onde a variável modificada  $z_i$  é definida por

$$z_i = \hat{\nu}_i + \hat{A}_i^{-T} \hat{b}_i, \quad \text{com} \quad \nu_i = \begin{pmatrix} \eta_i \\ \delta_i \end{pmatrix}.$$

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Estimação de $\alpha$ :

- Para a estimação dos parâmetros de correlação,  $\alpha$ , que determinam as correlações entre  $y_{ij}^*$  e  $y_{ij}^*$  (para  $j \neq I$ ), utiliza-se o procedimento descrito anteriormente.

## Modelagem Conjunta dos Parâmetros de Posição e Precisão

**Procedimento Prático para Estimação Conjunta dos Parâmetros  $\beta$ ,  $\gamma$  e  $\alpha$ :**

- Supondo independência entre as observações, ajusta-se um modelo de regressão linear de  $g(y)$  sobre  $X$  e outro de  $f(\check{\phi})$  sobre  $Q$  por mínimos quadrados ordinários.
- Para o parâmetro de precisão, utiliza-se:

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

sendo:

$$\check{\mu}_{ij} = g^{-1} \left( x_{ij}^T \hat{\beta}_{MQO} \right) \quad \text{e} \quad \check{\sigma}_{ij}^2 = \frac{\check{e}^T \check{e} \check{G}_{ij}^2}{n - p}.$$

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Estimativas Iniciais e Estrutura de Correlação:

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

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Atualização Iterativa dos Parâmetros:

- Atualiza-se  $\beta$  por:

$$\hat{\beta}^{(m+1)} = \left\{ \left[ \sum_{i=1}^n X_i^T \hat{W}_{\beta i} X_i \right]^{-1} \left[ \sum_{i=1}^n X_i^T \hat{W}_{\beta i} z_{\beta i} \right] \right\}^{(m)}, \quad (24)$$

com

$$\hat{W}_{\beta i} = \hat{G}_i \hat{\Phi}_i \hat{A}_i \hat{\Omega}_i^{-1} \hat{A}_i \hat{\Phi}_i \hat{G}_i \quad \text{e} \quad z_{\beta i} = \hat{\eta}_i + \left( \hat{A}_i \hat{\Phi}_i \hat{G}_i \right)^{-1} \hat{b}_i,$$

- E atualiza  $\gamma$  por:

$$\hat{\gamma}^{(m+1)} = \left\{ \left[ \sum_{i=1}^n \mathbf{Q}_i^T \hat{\mathbf{W}}_{\gamma_i} \mathbf{Q}_i \right]^{-1} \left[ \sum_{i=1}^n \mathbf{Q}_i^T \hat{\mathbf{W}}_{\gamma_i} \mathbf{z}_{\gamma_i} \right] \right\}^{(m)}, \quad (25)$$

com

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

# Modelagem Conjunta dos Parâmetros de Posição e Precisão

## Iteração até Convergência:

- As atualizações de  $\beta$ ,  $\gamma$  e  $\alpha$  prosseguem iterativamente até a convergência dos estimadores.
- Utiliza-se o processo iterativo conjunto com as estimativas iniciais  $\hat{\beta}^{(0)}$ ,  $\hat{\gamma}^{(0)}$  e  $\hat{\alpha}^{(0)}$ , repetindo até a convergência de  $\beta$  e  $\gamma$ .



# Diagnóstico

---

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada

## Análise de Diagnóstico Para Dados de Medidas Repetidas

- Verificação das suposições assumidas no ajuste.
- Identificação de possíveis violações.
- Detecção de observações influentes.
- Complexidade aumentada devido à estrutura de correlação intragrupo.

# Medidas de Diagnóstico

- Baseadas em um processo iterativo reponderado.
- Estimacão dos parâmetros de regressão  $\theta$ :
  - Modelagem da média:  $\theta = \beta$ .
  - Modelagem conjunta da média e dispersão:  $\theta = (\beta^T, \gamma^T)^T$ .
- As medidas de diagnóstico resultantes serão vetores de dimensão:
  - $N \times 1$  quando  $\theta = \beta$ ;
  - $2N \times 1$  quando  $\theta = (\beta^T, \gamma^T)^T$ ;(com  $N = nt$ ).

# Processo Iterativo Reponderado

- Processo iterativo para estimação de  $\theta$ :

$$\hat{\theta}^{(m+1)} = \left\{ \left[ \sum_{i=1}^n M_i^T \hat{W}_i M_i \right]^{-1} \left[ \sum_{i=1}^n M_i^T \hat{W}_i z_i \right] \right\}^{(m)} \quad (26)$$

- $M_i = X_i$  se  $\theta = \beta$ , ou  $M_i = \begin{bmatrix} X_i & 0 \\ 0 & Q_i \end{bmatrix}$  se  $\theta = (\beta^T, \gamma^T)^T$
- $\hat{W}_i = \hat{\Lambda}_i \hat{Q}_i^{-1} \hat{\Lambda}_i$
- $z_i = \hat{\nu}_i + \hat{\Lambda}_i^{-1} \hat{b}_i$ , com  $\hat{\nu}_i = M_i \hat{\theta}$

# Convergência do Processo Iterativo

- Convergência do processo iterativo leva à estimativa de  $\theta$ :

$$\hat{\theta} = (M^T \hat{W} M)^{-1} M^T \hat{W} z \quad (27)$$

- $M = (M_1^T, \dots, M_n^T)^T$
- $\hat{W} = \text{diag}(\hat{W}_1, \dots, \hat{W}_n)$
- $z = (z_1^T, \dots, z_n^T)^T$

# Ponto Alavanca Baseado na Matriz de Projeção

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

$$r_O = \hat{W}^{1/2}(z - \hat{\nu}) = (I - H)\hat{W}^{1/2}z, \quad (28)$$

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

$$H_i = \hat{W}_i^{1/2}M_i(M^\top \hat{W}M)^{-1}M_i^\top \hat{W}_i^{1/2}. \quad (29)$$

# Ponto Aberrante Baseado no Resíduo Padronizado

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

$$\begin{aligned}
 \text{Cov}(r_O) &= \text{Cov} \left( \hat{W}^{1/2} z - \hat{W}^{1/2} M \hat{\theta} \right) \\
 &= \text{Cov} \left( \hat{W}^{1/2} z - \hat{W}^{1/2} M (M^T \hat{W} M)^{-1} M^T \hat{W} z \right) \\
 &= (I - H) \hat{W}^{1/2} \text{Cov}(z) \hat{W}^{1/2} (I - H) \cong (I - H).
 \end{aligned}$$

- Como os elementos de  $r_O$  possuem variâncias distintas, o resíduo padronizado para a observação  $y_{ij}$  é definido por:

$$(r_{PD})_{ij} = \frac{e_{(ij)}^\top \widehat{W}^{1/2} (z - \widehat{\nu})}{\sqrt{1 - h_{ij}}} = \frac{e_{(ij)}^\top \widehat{W}^{1/2} \widehat{\Lambda}^{-1} \widehat{b}}{\sqrt{1 - h_{ij}}}, \quad (30)$$

em que:

- $\widehat{A} = \text{diag}(\widehat{A}_1, \dots, \widehat{A}_n)$ ,
- $\widehat{b} = (\widehat{b}_1^\top, \dots, \widehat{b}_n^\top)^\top$ ,
- $e_{(ij)}$  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  $H_i$ , com  $i = 1, \dots, n$  e  $j = 1, \dots, t'$ .

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

# Ponto Influyente Baseado na Distância de Cook

- Um ponto influente possui um perfil distinto na variável resposta e um valor elevado na matriz de projeção  $H$ .
- Para identificar essa influência utiliza-se a distância de Cook (COOK, 1977), que mede o afastamento entre:
  - As estimativas do vetor paramétrico com todas as observações ( $\hat{\theta}$ ).
  - As estimativas sem a observação  $y_{ij}$  ( $\hat{\theta}_{(ij)}$ ), onde  $i = 1, \dots, n$  e  $j = 1, \dots, t'$ .
- Como não é viável obter  $\hat{\theta}_{(ij)}$  de forma fechada, utiliza-se uma aproximação.
- Aproxima-se  $\hat{\theta}_{(ij)}$  com a primeira iteração do método *scoring* de Fisher, iniciando em  $\hat{\theta}$ .
- Essa abordagem foi introduzida por Pregibon (1981) e é adaptada para modelos de regressão com medidas repetidas.

- Tal adaptação é expressa da seguinte forma,

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

- Conseqüentemente, a distância de Cook, ao excluir a observação  $y_{ij}$ , é definida por:

$$(DC)_{ij} = \frac{1}{d} (\hat{\theta} - \hat{\theta}_{(ij)})^T \mathbf{M}^T \hat{\mathbf{W}} \mathbf{M} (\hat{\theta} - \hat{\theta}_{(ij)}) = (r_{PD})_{ij}^2 \frac{h_{ij}}{d(1 - h_{ij})}. \quad (31)$$

- 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.

## Seleção de Modelos e de Matriz de Correlação de Trabalho

- Critério de Informação de Akaike (AIC)
  - Seleção do modelo mais parcimonioso.
  - Minimização da discrepância entre modelo candidato e verdadeiro modelo.
- Definição do AIC:

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

sendo,

- $\ell(\cdot)$  o logaritmo da função de verossimilhança associada aos dados;
- $\hat{\beta}$  o estimador de máxima verossimilhança do modelo candidato;
- $p$  a dimensão do vetor de parâmetros  $\beta$ .

## QIC - Quasi-likelihood under the Independence model Criterion

- **Limitação do AIC em EEGs:**
  - Não se baseia diretamente em verossimilhanças.
- **Motivação para o QIC:**
  - Ao assumir independência entre as observações, as equações de estimação de Liang e Zeger (1986) são equivalentes à função quasi-escore (McCullagh e Nelder, 1989).
  - No caso da família exponencial, equivalem à função escore correspondente.
- **Proposta de Pan (2001) - QIC:**
  - Método útil para seleção de modelos e escolha da matriz de correlação de trabalho.

# QIC - Quasi-likelihood under the Independence model Criterion

- Definição do QIC:

$$QIC(R) \equiv -2\ell(\hat{\beta}(R)) + 2\text{tr}(\hat{S}_I \hat{J}_{nR}^{-1}) \quad (32)$$

sendo,

- $\ell(\cdot)$ : Logaritmo da função de quasi-verossimilhança.
- $\hat{\beta}(R)$ : Estimador de  $\beta$  sob a matriz de correlação  $R$ .
- $\hat{S}_I$ : Matriz de sensibilidade.
- $\hat{J}_{nR}^{-1}$ : Inversa da matriz de variância robusta.

## Simplificação do QIC

- Quando especificações do modelo via EEGs estão corretas:
  - $\hat{S}_I$  e  $\hat{J}_{nR}^{-1}$  são assintoticamente equivalentes.
  - $\text{tr}(\hat{S}_I \hat{J}_{nR}^{-1}) \approx 2$ .
- Medida simplificada QIC:

$$QIC_s(R) \equiv -2\ell(\hat{\beta}(R)) + 2p \quad (33)$$

- Adequada para seleção de covariáveis.

# Validade das Medidas QIC e $QIC_S$

- Medidas QIC e  $QIC_S$  são válidas quando:
  - Parâmetro de dispersão  $\phi^{-1}$  é conhecido e único.
- Caso  $\phi^{-1}$  desconhecido:
  - QIC é calculado utilizando o maior valor estimado para  $\phi^{-1}$  dentre os modelos candidatos.
  - $QIC_S$  é calculado utilizando valor estimado de  $\phi^{-1}$  do modelo ajustado com todas as covariáveis.











# Referências





---

Universidade Federal do Ceará, Depart. de Estatística e Matemática Aplicada

## Referências

-  Venezuela, M. K. (2008).  
Equação de estimação generalizada e influência local para modelos de regressão beta com medidas repetidas.  
*Tese de Doutorado em Estatística*, Instituto de Matemática e Estatística, USP, São Paulo.
-  Ferrari, S. L. P., & Cribari-Neto, F. (2004).  
Beta Regression for Modelling Rates and Proportions.  
*Journal of Applied Statistics*, 31(7), 799–815.
-  Cribari-Neto, F., & Zeileis, A. (2010).  
Beta Regression in R.  
*Journal of Statistical Software*, 34(2), 1–24.
-  Liang, K. Y., & Zeger, S. L. (1986).  
Longitudinal Data Analysis Using Generalized Linear Models.  
*Biometrika*, 73(1), 13–22.

-  Wedderburn, R. W. M. (1974).  
Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss-Newton Method.  
*Biometrika*, 61(3), 439–447.
-  Prentice, R. L., & Zhao, L. P. (1991).  
Estimating Equations for Parameters in Means and Covariances of Multivariate Discrete and Continuous Responses.  
*Biometrics*, 47(3), 825–839.
-  Cook, R. D. (1977).  
Detection of Influential Observations in Linear Regression.  
*Technometrics*, 19(1), 15–18.
-  Pan, W. (2001).  
Akaike's Information Criterion in Generalized Estimating Equations.  
*Biometrics*, 57(1), 120–125.

-  Pregibon, D. (1981).  
Logistic Regression Diagnostics.  
*The Annals of Statistics*, 9(4), 705–724.
-  McCullagh, P., & Nelder, J. A. (1989).  
Generalized Linear Models.  
*Springer*.
-  Smithson, M., & Verkuilen, J. (2006).  
A Better Lemon Squeezer? Maximum-Likelihood Regression with Beta-Distributed Dependent Variables.  
*Psychological Methods*, 11(1), 54–71.
-  Cox, D. R., & Snell, E. J. (1989).  
Analysis of Binary Data (2<sup>a</sup> ed.).  
*Chapman & Hall*, London.



# UFC



## Grato pela atenção!

Douglas Chaves Moura

Depart. de Estatística e Matemática Aplicada - DEMA

26 de fevereiro de 2025