Modelos hierárquicos (multiníveis) de dois níveis

Prof. Caio Azevedo

Introdução

- Como já discutido anteriormente, modelos (dados) hierárquicos podem apresentar dois ou mais níveis.
- Quando consideramos apenas um nível, o respectivo modelo (hierárquico) simplifica-se em algum correspondente modelo usual de regressão (pe, MRNLH).
- Note, contudo, que não, necessariamente, há uma equivalência entre um modelo hierárquico (de dois níveis) com um modelo competidor, de um nível (veja os exemplos vistos em aula e a introdução).

Exemplo 1: Vestibular Unicamp (COMVEST)

- O vestibular da Unicamp é uma das formas de acesso à essa Universidade (+ ENEM, Vestibular Indígena, Vagas Olímpícas, Vagas remanescentes, PROFIS, Cursos de Graduação).
- No ano de 2018 houve um total de 73.487 inscritos para 3.340 vagas distribuídas em 70 cursos de graduação (link).
- O vestibular é dividido em duas fases (+ a prova de habilidades específicas para os candidatos aos cursos de: Arquitetura, Artes Cênicas, Artes Visuais, Música) (próximo slide):

Exemplo 1: Vestibular Unicamp

- Fase 1 (prova de conhecimentos gerais - Matemática, Física, Química, Biologia, História, Geografia, Língua Estrangeira): 90 itens (questões) de múltipla escolha com quatro alternativas
- Fase 2 (para os selecionados na Fase 1): As provas são compostas por itens dissertativos (abertos). Depende da área ao qual o curso (1^a opção) pretendido pertence + a redação, prova de Língua Portuguesa e Literaturas de Língua Portuguesa e Interdisciplinares com Língua Inglesa (próximo slide):

Exemplo 1: Vestibular Unicamp (Fase 2)

- Área de Ciências Biológicas / Saúde: Interdisciplinares de Ciências Humanas e Ciências da Natureza, Matemática, Física, Biologia e Química.
- Área de Ciências Exatas / Tecnológicas: Interdisciplinares de Ciências Humanas e Ciências da Natureza, Matemática, Física, Química.
- Área de Ciências Humanas / Artes: Interdisciplinares de Ciências Humanas e Ciências da Natureza, Matemática, História, Geografia.

Exemplo 1 (Cont.)

- Há várias variáveis relacionadas aos candidatos (C): curso pretendido (CP) (1^a e 2^a escolhas) e informações disponíveis no questionário sócio-econômico.
- Podemos pensar em uma estrutura hierárquica de dois níveis: nível 1 (C) e nível 2 (CP).
- Variável resposta: nota bruta (NB) $\rightarrow$ número de acertos (crítica: neste caso a reposta é uma contagem binomial - número de sucessos em 90 tentativas). Assim, assumir normalidade não é uma hipótese razoável (contagem binomial e contagem de Poisson - Lee & Nelder (1996)). Se o candidato não tiver respondido a determinado item, é considerado como resposta incorreta.

Exemplo 2: SARESP (site)

- O Saresp é aplicado pela Secretaria da Educação do Estado de São Paulo com a finalidade de produzir um diagnóstico da situação da escolaridade básica paulista, visando orientar os gestores do ensino no monitoramento das políticas voltadas para a melhoria da qualidade educacional.
- Os alunos do 3º, 5º, 7º e 9º anos do Ensino Fundamental e da 3ª série do Ensino Médio têm seus conhecimentos avaliados por meio de provas com questões de Língua Portuguesa, Matemática, Ciências Humanas, Ciências da Natureza e redação. Os resultados são utilizados para orientar ações e também integram o cálculo do Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp).

Exemplo 2 (cont.)

- O conhecimento (nota) dos alunos é avaliado através da Teoria da Resposta ao Item (TRI).
- Além das notas, escolas e municípios (delegacias) de ensino, há outras informações relativas aos alunos e escolas, levantadas através de questionários apropriados (link).
- Como mencionado, podemos pensar (ao menos) em estruturas de dois e três níveis:
  - Dois níveis: aluno (nível 1) e escola (nível 2)
  - Três níveis: aluno (nível 1), escola (nível 2) e município (nível 3).

Exemplo 3: Desempenho cognitivo (Achievement data)

- Disponível no site do livro Multilevel Modeling using R (link). Mais referências (site do curso).
- Relativo ao "Indiana's Prime Time" que consite em um mecanismo de financiamento projetado para reduzir o tamanho das turmas (classes) e/ou reduzir o tamanho da razão aluno/professor razão nas primeiras séries do ensino fundamental (Estados Unidos).
- Várias variáveis foram medidas: conhecimento (desempenho) em várias áreas, idade, sexo, turma, escola entre outras.
- Podemos considerar (ao menos), estruturas de dois e três níveis:
  - Dois níveis: aluno (nível 1) e escola (nível 2)
  - Três níveis: aluno (nível 1), turma (nível 2) e escola (nível 3).

Exemplo 4: Produção de aveia (dados no R - "oats")

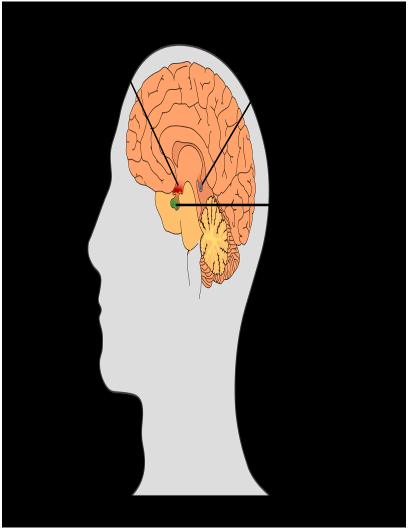
- O rendimento da aveia em um ensaio de campo em parcelas subdivididas, usando três variedades (V: victory, golden rain, marvellous) e quatro níveis de tratamento com manjares (0,0 wct; 0,2 wct; 0,4 wct; 0,6 wct).
- O experimento foi desenvolvido em 6 blocos de 3 parcelas principais, cada uma dividida em 4 subparcelas. As variedades foram aplicadas nas parcelas principais e os manjares, nas subparcelas.
- Dois níveis: nível 1 (subparcela) e nível 2 (parcela principal) (em não se considerando os Blocos).
- Três níveis: nível 1 (subparcela), nível 2 (parcela principal), nível 3 (Blocos).

Exemplo 5: Distância do centro da glândula pituitária para a fissura pterigomaxilar (Potthoff and Roy (1964))

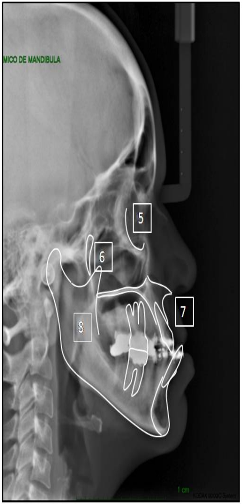
- Esta pesquisa está relacionada aos famosos dados de Potthoff-Roy, usados para demonstrar a utilização da MANOVA em análise de medidas repetidas (comparação entre grupos, embora comparação entre variáveis seja possível).
- O estudo considerou 16 meninos e 11 meninas, nos quais, nas idades 8, 10, 12 e 14 anos tiveram a distância (mm) do centro da glândula pituitária para a fissura pterigomaxilar medidas.

Exemplo 5 (Cont.)

- Mudanças nas distâncias pituitária-pterigomaxilar durante o crescimento são importantes na terapia ortodôntica.
- Os objetivos do estudo foram descrever a distância em função da idade e comparar esse desenvolvimento (temporal) entre sexos.
- Nível 1 (medida repetida ao longo dos anos), Nível 2 (crianças).
- Disponível no programa R, pacote "mice" sob o nome de "potthoffroy".



- 5 – Borda inferior da órbita
- 6 – Fissura pterigomaxilar
- 7 – Maxila
- 8 – Mandíbula
- 9 – Dentes (incisivos e 1º. molares)



Mais exemplos

- Na literatura sugerida (site do curso e programa da disciplina).
- Em pacotes (que lidam com modelos multiníveis) do R: veja o site do curso.
- Do livro Multilevel Analysis: Techniques and Applications, Third Edition (veja aqui o site do livro).
- Conjuntos de dados, do site do Centre for Multilevel Modelling
- Do livro Multilevel Statistical Models, conjuntos de dados.

Níveis dos modelos

- Como já mencionado, nos concentraremos (pelo menos por enquanto) em modelos com distribuição normal para os erros e efeitos aleatórios. Dentro desse contexto:
  - Modelos de um nível (MH1N) são, essencialmente, os modelos de regressão normal linear homocedástico ME613.
  - Modelos de dois níveis (MH2N) (como vistos na aula Introdutória slides). Neste caso, os coeficiente de regressão escolhidos variam, através de uma estrutura de regressão, entre os elementos do segundo nível.

Níveis dos modelos (Cont.)

- Cont.:
  - Mais adiante veremos modelos de três níveis (MH3N) (na prática, via de regra, não se utilizam modelos com mais de três níveis). Nesses modelos, os coeficientes de regressão do segundo nível, (que nos modelos de dois níveis, não variam), podem variar, através de um modelo de regressão, entre os elementos do terceiro nível.

(Etapas de) Construção/utilização de um modelo de regressão

1. Compreender os objetivos do problema e levantar todas as informações possíveis, junto aos respectivos responsáveis (consulentes/pesquisadores).
2. Rever a literatura pertinente (da área em questão e da Estatística)
3. Realizar uma adequada análise descritiva dos dados.
4. Com base nos itens 1), 2), 3) e na experiência do analista de dados (você!), elencar e ajustar todos os modelos que se considerar pertinentes.

(Etapas de) Construção/utilização de um modelo de regressão (cont.)

5. Realizar análises de diagnóstico apropriadas para os modelos do item 4) e, entre aqueles (se houver mais de um) que se ajustaram de forma adequada, comparar seus ajustes.
6. Com o modelo selecionado no item 5) (eventualmente uma versão reduzida ou seja, com menos parâmetros): extrair todas as conclusões apropriadas realizando, se for o caso, análises preditivas; fornecer estimativas (pontuais e intervalares) para os parâmetros de interesse e realizar pertinentes testes de hipótese.

Modelo geral (MRLNH) de um nível sem covariável

- Corresponde ao MRNLH usual (link), sob alguma parametrização (aula). Neste caso, assim como anteriormente (introdução), escolheremos como casela de referência o grupo 1 (j=1).

$$Y_{ji} = \mu + \alpha_j + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j$$

- Erros $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, μ, α_j são não aleatórios.
- $\mathcal{E}_{\xi_{ji}}(Y_{ji}) \equiv \mathcal{E}(Y_{ji}) = \mu_j$.
- $\mathcal{V}_{\xi_{ji}}(Y_{ji}) \equiv \mathcal{V}(Y_{ji}) = \sigma^2$ (modelo homocedástico).
- $\mu_j = \mu + \alpha_j$: média (populacional) da resposta do j-ésimo grupo, em que $\alpha_1 = 0$ (o grupo 1 é a casela de referência). Assim, $\mu = \mu_1$.

Cont. modelo 1

- $\alpha_j = \mu_j - \mu_1$ é o incremento/diferença (positivo(a) ou negativo(a)), entre a média da resposta do j-ésimo grupo e do grupo 1. Assim, se for positivo (negativo) indica que o grupo j possui média maior (menor) do que o grupo 1.
- Além disso, temos que: $Y_{ji} \overset{ind.}{\sim} N(\mu + \alpha_j, \sigma^2)$.
- Defina: $n = \sum_{j=1}^{J} n_j$.
- Vamos também explorar formas matriciais dos modelos. Estas são úteis para: estimação, implementação computacional, obtenção de propriedades (dos modelos e das ferramentas inferencias), análise residual (diagnóstico), análise de influência, entre outros.

Cont. modelo 1 (forma matricial por grupo)

- Para este modelo em particular, temos que:

$$\boldsymbol{Y}_{j(n_j \times 1)} = \boldsymbol{X}_{j(n_j \times p)} \boldsymbol{\beta}_{(p \times 1)} + \boldsymbol{\xi}_{j(n_j \times 1)}, \text{ em que}$$

- $\boldsymbol{Y}_j = (Y_{j1}, ..., Y_{jn_j})'$ (vetor de respostas do j-ésimo grupo - desconhecido, observável, aleatório), $\boldsymbol{\beta} = (\mu, \alpha_2, ..., \alpha_J)'$ (vetor de parâmetros de regressão - desconhecido, não-observável, não-aleatório), $\boldsymbol{X}_j$ (matriz de planejamento-conhecida, observável, não-aleatória).
- Erros $\boldsymbol{\xi}_j \overset{iid}{\sim} N_{n_j}(0, \boldsymbol{\Sigma}_j)$. Assim $\boldsymbol{Y}_j \overset{iid}{\sim} N_{n_j}(\boldsymbol{X}_j\boldsymbol{\beta}, \boldsymbol{\Sigma}_j)$
- Nos próximos slides, apresenta-se mais detalhes.

Cont. modelo 1 (forma matricial por grupo)

$$\boldsymbol{X}_j = \begin{bmatrix} 1 & 0 & \ldots & 0 & \overbrace{1}^{\text{posição j}} & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 & \vdots & \vdots \\ 1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_j = \begin{bmatrix} \sigma^2 & 0 & \ldots & 0 \\ 0 & \sigma^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma^2 \end{bmatrix} = \sigma^2 \boldsymbol{I}_{n_j}$$

em que $\boldsymbol{I}_{n_j}$ é uma matriz identidade de ordem n_j, $j = 1, 2, ..., J$.

Cont. modelo 1 (forma matricial geral)

$$\boldsymbol{Y}_{(n\times 1)} = \boldsymbol{X}_{(n\times p)}\boldsymbol{\beta}_{(p\times 1)} + \boldsymbol{\xi}_{(n\times 1)}$$

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \\ \vdots \\ \boldsymbol{Y}_J \end{bmatrix} ; \boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_J \end{bmatrix} ; \boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \vdots \\ \boldsymbol{\xi}_J \end{bmatrix}$$

OBS: diversos modelos (MRNLH) podem ser escritos dentro dessa estrutura (não somente o modelo em questão). Veja ME613 e ME623.

Modelo hierárquico de dois níveis sem covariáveis (MH1)

- Neste caso, o modelo é bem semelhante àquele utilizado no Exemplo artificial 1 (introdução).

$$
\begin{aligned}
Y_{ji} &= \beta_{0j} + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j \, (\text{nível } 1) \\
\beta_{0j} &= \gamma_{00} + u_{0j} \, (\text{nível } 2)
\end{aligned}
$$

- Erros e efeitos aleatórios: $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, $u_{0j} \overset{iid}{\sim} N(0, \psi)$, $\xi_{ji} \perp u_{0j}, \forall i, j$ e γ_{00} é não aleatório.

- (momentos condicionais) $\mathcal{E}(Y_{ji}|u_{0j}) = \gamma_{00} + u_{0j}$ - valor esperado da resposta para as UAE's pertencentes ao grupo j; $\mathcal{V}(Y_{ji}|u_{0j}) = \sigma^2$ - variância da resposta para as UAE's pertencentes ao grupo j.

Modelo hierárquico de dois níveis sem covariáveis (MH1) (cont.)

- (momentos marginais) $\mathcal{E}(Y_{ji}) = \gamma_{00}$ - valor esperado global da resposta (média das médias), $\mathcal{V}(Y_{ji}) = \sigma^2 + \psi$ - variância global da resposta (considerando todos os grupos).
- $u_{0j} = \mathcal{E}(Y_{ji}|u_{0j}) - \mathcal{E}(Y_{ji})$: representa o quanto a média (condicional) de cada UAE no nível 2 difere da média geral (de todas as UAE's no nível 2).
- $\text{Cov}(Y_{ji}, Y_{ji'}) = \psi$ e $\text{Corre}(Y_{ji}, Y_{ji'}) = \frac{\psi}{\psi + \sigma^2}, \forall i \neq i'$.
- Distribuições condicionais e marginais $Y_{ji}|u_{0j} \overset{ind.}{\sim} N(\gamma_{00} + u_{0j}, \sigma^2)$ e $Y_{ji} \sim N(\gamma_{00}, \sigma^2 + \psi)$ (estes últimos, **não** são independentes).

Modelo hierárquico de dois níveis sem covariáveis (MH1) (cont.)

- De uma forma mais geral, temos que

$\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})' \sim N_{n_j}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, $j = 1, 2, ..., J$, em que

$$\boldsymbol{\mu}_j = \gamma_{00} 1_{n_j}, 1_{n_j} = (1, 1, ..., 1)'_{(n_j \times 1)}, \text{ e}$$

$$\boldsymbol{\Sigma}_j = \begin{bmatrix} \sigma^2 + \psi & \psi & \dots & \psi \\ \psi & \sigma^2 + \psi & \dots & \psi \\ \vdots & \vdots & \ddots & \vdots \\ \psi & \psi & \dots & \sigma^2 + \psi \end{bmatrix}_{n_j \times n_j}$$

Modelo hierárquico de dois níveis sem covariáveis (MH1) (cont.)

- Em modelos de dois níveis podemos considerar covariáveis tanto no nível 1 (resposta) quanto no nível 2 [coeficientes (aleatórios)].
- Quando covariáveis são consideradas para a resposta, procura-se explicar a variabilidade da resposta (condicional/marginal). Por outro lado, quando covariáveis são consideradas para os coeficientes, busca-se explicar a variabilidade entre eles (coeficientes e grupos).

Modelo hierárqico de dois níveis sem covariáveis (MH1) (cont.)

- Pode-se considerar diversas possibilidades de modelagem (em termos das covariáveis, como utilizá-las e da estrutura dos efeitos aleatórios), as quais podem gerar modelos com diversas características: estruturas de dependência, estruturas de heterocedasticidade, distribuição marginal da resposta, entre outras .
- Certamente, além das características do problemas, dos dados e do conhecimento de especialistas, devemos procurar manter parcimônia (A Modern Approach to Regression with R, Regression Modeling Strategies, Regression and Other Stories).

Comparação entre os modelos

Característica	MRNLH	MRNLH2N
$\mathcal{E}(Y_{ji}\|u_{0j})$	$\mu + \alpha_j$	$\gamma_{00} + u_{0j}$
$\mathcal{E}(Y_{ji})$	$\mu + \alpha_j$	γ_{00}
$\mathcal{V}(Y_{ji}\|u_{0j})$	σ^2	σ^2
$\mathcal{V}(Y_{ji})$	σ^2	$\sigma^2 + \psi$
Parâmetros (efeitos fixos)	J+1	3
Efeitos aleatórios	0	J
Testes de Hipótese (H_0)	$\alpha_2 = ... = \alpha_J = 0$	$\psi = 0$

Comparação entre os modelos (cont.)

Característica	**MRNLH**	**MRNLH2N**
Homocedástico	Sim	Sim
Dependência entre as obs.	Não	Sim
Estimação Freq.	MQO	MVR

OBS: Os parâmetros σ^2 não são os mesmos entre os modelos. MQO: mínimos quadrados ordinários. MVR: máxima verossimilhança restrita.

Um modelo hierárquico de dois níveis com uma covariável no nível 1 (MH2)

- Neste caso, um modelo é dado por:

$$
\begin{aligned}
Y_{ji} &= \beta_{0j} + \beta_{1j}x_{ji} + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j \,\text{(nível 1)} \\
\beta_{0j} &= \gamma_{00} + u_{0j} \,\text{(nível 2)}
\end{aligned}
$$

- Erros e efeitos aleatórios: $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, $u_{0j} \overset{iid}{\sim} N(0, \psi)$, $\xi_{ji} \perp u_{0j}, \forall i, j$ e γ_{00} é não aleatório.
- (momentos condicionais) $\mathcal{E}(Y_{ji}|u_{0j}) = \gamma_{00} + u_{0j} + \beta_{1j}x_{ji}$ - valor esperado da resposta para as UAE's pertencentes ao grupo j, com valor da covariável igual a x_{ji}. Note que se $x_{ji} = 0$, então $\mathcal{E}(Y_{ji}|u_{0j}) = \gamma_{00} + u_{0j}$.

Um modelo hierárquico de dois níveis com uma covariável no nível 1 (MH2) (cont.)

- (momentos condicionais) $\mathcal{V}(Y_{ji}|u_{0j}) = \sigma^2$ - variância da resposta para as UAE's pertencentes ao grupo j.
- (momentos marginais) $\mathcal{E}(Y_{ji}) = \gamma_{00} + \beta_{1j}x_{ji}$ - valor esperado global (considerando todos os grupos) da resposta para UAE's pertecentes ao grupo j e com valor da coviariável igual x_{ji}. Note que se $x_{ji} = 0$, então $\mathcal{E}(Y_{ji}) = \gamma_{00}$. $\mathcal{V}(Y_{ji}) = \sigma^2 + \psi$ - variância global da resposta (considerando todos os grupos).

Um modelo hierárquico de dois níveis com uma covariável no nível 1 (MH2) (cont.)

- $u_{0j} = \mathcal{E}(Y_{ji}|u_{0j}) - \mathcal{E}(Y_{ji})$: representa o quanto a média (condicional) de cada UAE no nível 2 difere da média geral (marginal) (de todas as UAE's no nível 2), que apresentam o mesmo valor de x_{ji}.
- $\text{Cov}(Y_{ji}, Y_{ji'}) = \psi$ e $\text{Corre}(Y_{ji}, Y_{ji'}) = \frac{\psi}{\psi + \sigma^2}$, $\forall i \neq i'$.
- Distribuições condicionais e marginais $Y_{ji}|u_{0j} \overset{ind.}{\sim} N(\gamma_{00} + u_{0j} + \beta_{1j} x_{ji}, \sigma^2)$ e $Y_{ji} \sim N(\gamma_{00} + \beta_{1j} x_{ji}, \sigma^2 + \psi)$ (estes últimos, **não** são independentes).

Um modelo hierárquico de dois níveis com uma covariável no nível 1 (MH2) (cont.)

- De uma forma mais geral, temos que

$$\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})' \sim N_{n_j}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j),\ j = 1, 2, ..., J, \text{ em que}$$

$$\boldsymbol{\mu}_j = \gamma_{00} 1_{n_j} + \beta_1 \boldsymbol{x}_j, 1_{n_j} = (1, 1, ..., 1)'_{(n_j \times 1)}, \boldsymbol{x}_j = (x_{j1}, x_{j2}, ..., x_{jn})' \text{ e}$$

$$\boldsymbol{\Sigma}_j = \begin{bmatrix} \sigma^2 + \psi & \psi & \dots & \psi \\ \psi & \sigma^2 + \psi & \dots & \psi \\ \vdots & \vdots & \ddots & \vdots \\ \psi & \psi & \dots & \sigma^2 + \psi \end{bmatrix}_{n_j \times n_j}$$

Um modelo hierárquico de dois níveis com uma covariável no nível 1 (MH3)

- Neste caso, o modelo é bem semelhante aquele utilizado no Exemplo artificial 2 (introdução).

$$
\begin{aligned}
Y_{ji} &= \beta_0 + \beta_{1j} x_{ji} + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j \, (\text{nível } 1) \\
\beta_{1j} &= \gamma_{10} + u_{1j} \, (\text{nível } 2)
\end{aligned}
$$

- Erros e efeitos aleatórios: $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, $u_{1j} \overset{iid}{\sim} N(0, \psi)$, $\xi_{ji} \perp u_{1j}, \forall i, j$ e γ_{10} é não aleatório.
- (momentos condicionais) $\mathcal{E}(Y_{ji}|u_{1j}) = \beta_0 + (\gamma_{10} + u_{1j}) x_{ji}$ - valor esperado da resposta para as UAE's pertecentes ao grupo j e com valor da coviariável igual x_{ji}.

Um modelo hierárquico de dois níveis com uma covariável no nível 1 (MH3) (cont.)

- (momentos condicionais) Note que se $x_{ji} = 0$, então $\mathcal{E}(Y_{ji}|u_{1j}) = \beta_0$. $\mathcal{V}(Y_{ji}|u_{1j}) = \sigma^2$ variância da resposta para as UAE's pertencentes ao grupo j.
- (marginais) $\mathcal{E}(Y_{ji}) = \beta_0 + \gamma_{10}x_{ji}$ valor esperado global (considerando todos os grupos) da resposta para UAE's com valor da coviariável igual a x_{ji}. Note que se $x_{ji} = 0$, então $\mathcal{E}(Y_{ji}) = \beta_0 = \mathcal{E}(Y_{ji}|u_{1j})$ (interpretar). $\mathcal{V}(Y_{ji}) = \sigma^2 + \psi x_{ji}^2$ variância global da resposta (considerando todos os grupos) - modelo heterocedástico. Também, se $x_{ji} = 0$, então $\mathcal{V}(Y_{ji}) = \sigma^2 = \mathcal{V}(Y_{ji}|u_{1j})$. O que isso significa?

Um modelo hierárquico de dois níveis com uma covariável no nível 1 (MH3) (cont.)

- $u_{1j} = \frac{1}{x_{ji}} \left(\mathcal{E}(Y_{ji}|u_{1j}) - \mathcal{E}(Y_{ji}) \right)$: representa o quanto a média (condicional) de cada UAE no nível 2 difere da média geral (de todas as UAE's no nível 2), para UAE's, ponderada por x_{ji} (se $x_{ji} \neq 0$).
- $Cov(Y_{ji}, Y_{ji'}) = x_{ji}x_{ji'}\psi$, $Corre(Y_{ji}, Y_{ji'}) = \frac{x_{ji}x_{ji'}\psi}{\sqrt{\sigma^2+\psi x_{ji}}\sqrt{\sigma^2+\psi x_{ji'}}}, \forall i \neq i'$
- Para pares de UAE's em que pelo menos um delas apresenta $x_{ji} = 0$, teremos covariâncias e correlações nulas. Interpretar.
- Distribuições condicionais e marginais $Y_{ji}|u_{1j} \overset{ind.}{\sim} N(\beta_0 + (u_{1j} + \gamma_{10})x_{ji}, \sigma^2)$ e $Y_{ji} \sim N(\beta_0 + \gamma_{10}x_{ji}, \sigma^2 + \psi x_{ji}^2)$ (estes últimos, **não** são independentes).

Um modelo geral de dois níveis com uma covariável no nível 1 (MH3) (cont.)

- De uma forma mais geral, temos que

$$\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})' \sim N_{n_j}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j),\ j = 1, 2, ..., J, \text{ em que}$$

$$\boldsymbol{\mu}_j = \beta_0 1_{n_j} + \gamma_{10}\boldsymbol{x}_j, 1_{n_j} = (1, 1, ..., 1)'_{(n_j \times 1)}, \boldsymbol{x}_j = (x_{j1}, x_{j2}, ..., x_{jn})', \text{ e}$$

$$\boldsymbol{\Sigma}_j = \sigma^2 \boldsymbol{I}_{n_j} + \psi \boldsymbol{x}_j \boldsymbol{x}_j' = \begin{bmatrix} \sigma^2 + \psi x_{j1}^2 & \psi x_{j1}x_{j2} & \dots & \psi x_{j1}x_{jn_j} \\ \psi x_{j1}x_{j2} & \sigma^2 + \psi x_{j2}^2 & \dots & \psi x_{j2}x_{jn_j} \\ \vdots & \vdots & \ddots & \vdots \\ \psi x_{j1}x_{jn_j} & \psi x_{j2}x_{jn_j} & \dots & \sigma^2 + \psi x_{jn_j}^2 \end{bmatrix}_{n_j \times n_j}$$

Um modelo hierárquico geral de dois níveis com uma covariável no nível 1 (MH4)

- Neste caso, o modelo é dado por:

$$
\begin{aligned}
Y_{ji} &= \beta_{0j} + \beta_{1j} x_{ji} + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j \, \text{(nível 1)} \\
\beta_{0j} &= \gamma_{00} + u_{0j} \, \text{(nível 2)} \\
\beta_{1j} &= \gamma_{10} + u_{1j} \, \text{(nível 2)}
\end{aligned}
$$

- Erros e efeitos aleatórios: $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, $\boldsymbol{u}_j = (u_{0j}, u_{1j})' \overset{iid}{\sim} N_2(0, \boldsymbol{\Psi})$, $\xi_{ji} \perp \boldsymbol{u}_j, \forall i, j$ e $(\gamma_{00}, \gamma_{10})'$ são não aleatórios e $\boldsymbol{\Psi} = \begin{bmatrix} \psi_{00} & \psi_{01} \\ \psi_{01} & \psi_{11} \end{bmatrix}$

Um modelo geral de dois níveis com uma covariável no nível 1 (MH4) (cont.)

- (momentos condicionais) $\mathcal{E}(Y_{ji}|\boldsymbol{u}_j) = \gamma_{00} + u_{0j} + (\gamma_{10} + u_{1j})x_{ji}$ - valor esperado da resposta para as UAE's pertecentes ao grupo j e com valor da coviariável igual x_{ji}. $\mathcal{V}(Y_{ji}|\boldsymbol{u}_j) = \sigma^2$.
- (momentos marginais) $\mathcal{E}(Y_{ji}) = \gamma_{00} + \gamma_{10}x_{ji}$ - valor esperado da resposta para as UAE's pertecentes ao grupo j e com valor da coviariável igual x_{ji}. $\mathcal{V}(Y_{ji}) = \psi_{00} + \psi_{11}x_{ji}^2 + \psi_{01}x_{ji} + \sigma^2$.
- $Cov(Y_{ji}, Y_{ji'}) = \psi_{00} + \psi_{11}x_{ji}x_{ji'} + \psi_{01}x_{ji} + \psi_{01}x_{ji'}$, $i \neq i'$

Um modelo geral de dois níveis com uma covariável no nível 1 (MH4) (cont.)

- $Corre(Y_{ji}, Y_{ji'}) = \frac{\psi_{00}+\psi_{11}x_{ji}x_{ji'}+\psi_{01}x_{ji}+\psi_{01}x_{ji'}}{\sqrt{\psi_{00}+\psi_{11}x_{ji}^2+\psi_{01}x_{ji}+\sigma^2}\sqrt{\psi_{00}+\psi_{11}x_{ji'}^2+\psi_{01}x_{ji'}+\sigma^2}}$
- $u_{0j} = \mathcal{E}(Y_{ji}|\boldsymbol{u}_j, x_{ji} = 0) - \mathcal{E}(Y_{ji}|x_{ji} = 0)$ é a diferença entre o valor esperado para o grupo j e a esperança global, para UAE's com $x_{ji} = 0$. Outra interpretação: é o quanto o intercepto do grupo j se difere do intercepto comum a todos os grupos.

Um modelo geral de dois níveis com uma covariável no nível 1 (MH4) (cont.)

- $u_{1j} = \frac{1}{x_{ji}} \left(\mathcal{E}(Y_{ji}|u_{1j}) - \mathcal{E}(Y_{ji}) \right), \forall x_{ji} \neq 0$, é a diferença entre o valor esperado para o grupo j, em relação a distribuição de u_{0j} e a esperança global. Outra interpretação: é o quanto o coeficiente angular do grupo j se difere do coeficiente angular comum a todos os grupos.
- Distribuições condicionais e marginais:
  $Y_{ji}|\boldsymbol{u}_j \overset{ind}{\sim} N(\beta_0 + u_{0j} + (\gamma_{10} + u_{1j})x_{ji}, \sigma^2)$ e
  $Y_{ji} \sim N(\beta_0 + \gamma_{10}x_{ji}, \psi_{00} + \psi_{11}x_{ji}^2 + \psi_{01}x_{ji} + \sigma^2)$ (estes últimos não são independentes).

Um modelo geral de dois níveis com uma covariável no nível 1 (MH4) (cont.)

- De uma forma mais geral, temos que

$\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})' \sim N_{n_j}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, $j = 1, 2, ..., J$, em que

$$\boldsymbol{\mu}_j = \gamma_{00}1_{n_j} + \gamma_{10}\boldsymbol{x}_j, 1_{n_j} = (1, 1, ..., 1)'_{(n_j \times 1)}, \boldsymbol{x}_j = (x_{j1}, x_{j2}, ..., x_{jn})' \text{ e}$$

$$\boldsymbol{\Sigma}_j = \psi_{00}1_{n_j \times n_j} + \psi_{01}(1_{n_j}\boldsymbol{x}_j' + \boldsymbol{x}_j 1_{n_j}' - \text{diag}(\boldsymbol{x}_j)) + \boldsymbol{x}_j\boldsymbol{x}_j'\psi_{11} + \sigma^2\boldsymbol{I}_{n_j}$$

em que $1_{(n_j \times n_j)}$ é uma matriz de uns, de dimensão n_j.

Comentários

1. Podemos considerar covariáveis também para os coeficientes do nível 1 [nos casos vistos, seriam (β_{0j}, β_{1j})]. Offsets também podem ser incluídos.
2. Podemos considerar covariáveis qualitativas: sexo, grau de formação, etc (não, necessariamente, associados aos níveis da estrutura hierárquica). Nesse caso, podemos precisar de mais índices na resposta e/ou erros e/ou efeitos aleatórios e/ou efeitos fixos .
3. Podemos centralizar (ou padronizar) as covariáveis quantitativas. Nesse caso, alguns dos resultados e interpretações podem mudar.
4. Também, diferentes parametrizações podem ser consideradas o que, por sua vez, pode modificar resultados e/ou interpretações.

Exercícios

- Repetir as contas anteriores para os seguintes modelos:

  Modelo 1

$$
\begin{aligned}
Y_{ji} &= \beta_0 + \beta_{1j}x_{ji} + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j \text{ (nível 1)} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01}w_j + u_{0j} \text{ (nível 2)} \\
\beta_{1j} &= \gamma_{10} + u_{1j} \text{ (nível 2)}
\end{aligned}
$$

- Erros e efeitos aleatórios: $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, $\boldsymbol{u}_j = (u_{0j}, u_{1j})' \overset{iid}{\sim} N_2(0, \boldsymbol{\Psi})$, $\xi_{ji} \perp \boldsymbol{u}_j, \forall i, j$ e $(\gamma_{00}, \gamma_{01}, \gamma_{10})'$ são não aleatórios e $\boldsymbol{\Psi} = \begin{bmatrix} \psi_{00} & \psi_{01} \\ \psi_{01} & \psi_{11} \end{bmatrix}$.

Exercícios

- Modelo 2

$$
\begin{aligned}
Y_{ji} &= \beta_0 + \beta_{1j}x_{ji} + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j \text{ (nível 1)} \\
\beta_{0j} &= \gamma_{00} + u_{0j} \text{ (nível 2)} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11}w_j + u_{1j} \text{ (nível 2)}
\end{aligned}
$$

- Erros e efeitos aleatórios: $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, $\boldsymbol{u}_j = (u_{0j}, u_{1j})' \overset{iid}{\sim} N_2(0, \boldsymbol{\Psi})$, $\xi_{ji} \perp \boldsymbol{u}_j, \forall i, j$ e $(\gamma_{00}, \gamma_{10}, \gamma_{11})'$ são não aleatórios e $\boldsymbol{\Psi} = \begin{bmatrix} \psi_{00} & \psi_{01} \\ \psi_{01} & \psi_{11} \end{bmatrix}$.

Exercícios

- Modelo 3

$$
\begin{aligned}
Y_{ji} &= \beta_{0j} + \beta_{1j} x_{ji} + \xi_{ji}, j = 1, 2, ..., J; i = 1, 2, ..., n_j \text{ (nível 1)} \\
\beta_{0j} &= \gamma_{00} + \gamma_{01} z_j + u_{0j} \text{ (nível 2)} \\
\beta_{1j} &= \gamma_{10} + \gamma_{11} w_j + u_{1j} \text{ (nível 2)}
\end{aligned}
$$

- Erros e efeitos aleatórios: $\xi_{ji} \overset{iid}{\sim} N(0, \sigma^2)$, $\boldsymbol{u}_j = (u_{0j}, u_{1j})' \overset{iid}{\sim} N_2(0, \boldsymbol{\Psi})$, $\xi_{ji} \perp \boldsymbol{u}_j, \forall i, j$ e $(\gamma_{00}, \gamma_{01}, \gamma_{10}, \gamma_{11})'$ são não aleatórios e $\boldsymbol{\Psi} = \begin{bmatrix} \psi_{00} & \psi_{01} \\ \psi_{01} & \psi_{11} \end{bmatrix}$.

Exercícios

- Verificar como ficam os resultados anteriores, quando $\psi_{01} = 0$.
- Qual o significado da suposição anterior?
- Repetir os desenvolvimentos anteriores, para todos os modelos (constantes neste slide), centrando todas as covariáveis (quando pertinente) em suas respectivas médias amostrais (de forma adequada).
- Propor um MH apropriado, com as devidas interpretações e resultados, para cada um dos exemplos vistos nestes slides.

Modelo com dois níveis e várias covariáveis: formulação matricial geral

- Essencialmente, qualquer modelo hierárquico de dois níveis pode ser escrito, matricialmente, da seguinte forma (em que j = 1,...,J definem os grupos - nível 2 da hierarquia).

$$\boldsymbol{Y}_{j(n_j \times 1)} = \boldsymbol{X}_{j(n_j \times p)} \boldsymbol{\beta}_{j(p \times 1)} + \boldsymbol{\xi}_{j(n_j \times 1)} \qquad (1)$$

$$\boldsymbol{\beta}_{j(p \times 1)} = \boldsymbol{W}_{j(p \times q)} \gamma_{(q \times 1)} + \boldsymbol{u}_{j(p \times 1)}, \qquad (2)$$

em que $\boldsymbol{Y}_j$ é o vetor de respostas do j ésimo conglomerado, $\boldsymbol{X}_j$ e $\boldsymbol{W}_j$ sao matrizes de planejamento (conhecidas e não aleatórias), $\boldsymbol{\xi}_j \overset{ind.}{\sim} N_{n_j}(0, \boldsymbol{\Sigma}_j)$, $\boldsymbol{u}_j \overset{ind.}{\sim} N_p(0, \boldsymbol{\Psi})$, $\boldsymbol{\xi}_j \perp \boldsymbol{u}_j, \forall j$, $\boldsymbol{\gamma}$ é um vector de efeitos fixos.

Modelo com dois níveis e várias covariáveis: formulação matricial geral

- Do slide anterior nota-se que quaisquer estruturas lineareas, ou seja, que podem ser representadas por $\boldsymbol{X}_j\boldsymbol{\beta}_j$ e $\boldsymbol{W}_j\boldsymbol{\gamma}$, são passíveis de serem consideradas.
- Assim, não somente retas, mas também polinômios de qualquer grau, polinômios fracionários, funções trigonométricas, logaritmos entre outras, podem ser consideradas.
- Modelos intrinsicamente não lineares constituem uma outra classe, a qual veremos, mais adiante, de forma introdutória.

Formulação matricial de alguns modelos

- MH1: $\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})'$, $\boldsymbol{X}_j = 1_{n_j}$, $\boldsymbol{\beta}_j = \beta_{0j}$, $\boldsymbol{W}_j = 1$, $\boldsymbol{\gamma} = \gamma_{00}$, $\boldsymbol{u}_j = u_{0j}$
- MH2: $\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})'$, $\boldsymbol{X}_j = [1_{n_j}\ \boldsymbol{x}_j]$, $\boldsymbol{\beta}_j = [\beta_{0j}\ \beta_{1j}]'$, $\boldsymbol{W}_j = 1$, $\boldsymbol{\gamma} = \gamma_{00}$, $\boldsymbol{u}_j = u_{0j}$
- MH3: $\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})'$, $\boldsymbol{X}_j = [1_{n_j}\ \boldsymbol{x}_j]$, $\boldsymbol{\beta}_j = [\beta_0\ \beta_{1j}]'$, $\boldsymbol{W}_j = 1$, $\boldsymbol{\gamma} = \gamma_{10}$, $\boldsymbol{u}_j = u_{1j}$
- MH4: $\boldsymbol{Y}_j = (Y_{j1}, Y_{j2}, ..., Y_{jn_j})'$, $\boldsymbol{X}_j = [1_{n_j}\ \boldsymbol{x}_j]$, $\boldsymbol{\beta}_j = [\beta_{0j}\ \beta_{1j}]'$, $\boldsymbol{W}_j = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}_2$, $\boldsymbol{\gamma} = [\gamma_{00}\ \gamma_{10}]$, $\boldsymbol{u}_j = [u_{0j}\ u_{1j}]'$

Resultados importantes

- Distribuições de interesse. Pelas equações (1) e (2), temos que:

$$\boldsymbol{Y}_j|\boldsymbol{\beta}_j \overset{ind.}{\sim} N_{n_j}(\boldsymbol{X}_j\boldsymbol{\beta}_j, \boldsymbol{\Sigma}_j)$$
$$\boldsymbol{\beta}_j \overset{ind.}{\sim} N_p(\boldsymbol{W}_j\gamma, \boldsymbol{\Psi})$$

- Além disso, temos que $\boldsymbol{Y}_j \overset{ind.}{\sim} N_{n_j}(\boldsymbol{X}_j\boldsymbol{W}_j\gamma, \boldsymbol{\Sigma}_j + \boldsymbol{X}_j\boldsymbol{\Psi}\boldsymbol{X}_j')$.

- As verossimilhanças completa (aumentada) e marginal são dadas, respectivamente, por:

$$L(\boldsymbol{\beta}, \gamma, \boldsymbol{\Sigma}, \boldsymbol{\Psi}) = \prod_{i=1}^{J} f(\boldsymbol{y}_j|\boldsymbol{\beta}_j)f(\boldsymbol{\beta}_j) \;;\; L(\gamma, \boldsymbol{\Sigma}, \boldsymbol{\Psi}) = \prod_{j=1}^{J} f(\boldsymbol{y}_j)$$

em que $\boldsymbol{\beta} = (\beta_1, ..., \boldsymbol{\beta}_J)'$, $\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_1, ...\boldsymbol{\Sigma}_J)$ e $\boldsymbol{\Psi} = (\boldsymbol{\Psi}_1, ...\boldsymbol{\Psi}_J)$

Exercícios

- Coloque os modelos MH3 e MH4 na forma matricial, assim como os modelos de 1 a 3.
- Considere um modelo como o MH4, mas considerando, no primero nível: $Y_{ji} = \beta_{0j} + \beta_{1j}x_{ji} + \beta_{2j}x_{ji}^2$ (com estruturas de regressão para cada um dos parâmetros). Inteprete todos os parâmetros, bem como realize as contas como nos exemplos anteriores.
- Proponha ao menos dois modelos hierárquicos, diferentes daqueles que foram apresentados até agora, intepretando todos os parâmetros, bem como realizando as contas como nos exemplos anteriores.

Relação entre modelos mistos (MMI) e modelos hierárquicos (MH)

- Podemos escrever a formulação (1) - (2) da seguinte forma:

$$
\begin{aligned}
\boldsymbol{Y}_j &= \boldsymbol{X}_j(\boldsymbol{W}_j\gamma + \boldsymbol{u}_j) + \xi_j \\
&= \boldsymbol{X}_j\boldsymbol{W}_j\gamma + \boldsymbol{X}_j\boldsymbol{u}_j + \xi_j \\
&= \boldsymbol{W}_j^*\gamma + \boldsymbol{X}_j\boldsymbol{u}_j + \xi_j \qquad (3)
\end{aligned}
$$

em que $\boldsymbol{W}_j^* = \boldsymbol{X}_j\boldsymbol{W}_j$. A Equação (3) **assemelha-se** à estrutura (matricial) de modelos mistos veja (link). Contudo, existem diferenças conceituais e metodológicas entre eles (próximo slide):

Relação entre modelos mistos e modelos hierárquicos

- De fato, ambos objetivam (ainda que por motivos diferentes) modelar heterocedasticidade e dependência.
- Contudo, enquanto que os MMI surgem da necessidade de se considerar efeitos aleatórios e modelar dependência em medidas repetidas, os MH buscam introduzir semelhanças/dependência intra-grupos (clusters).
- Além disso, nos MH os níveis de hierarquia são explicitados na modelagem, além do que há uma especificação da estrutura de regressão para os coeficientes (de interesse).

Relação entre modelos mistos e modelos hierárquicos

- Ademais, podemos ter dados longitudinais com estrutura hierárquica, o que demanda a utilização de MH.
- Também, nos MH permitimos a inclusão de covariáveis para os efeitos aleatórios o que, em princípio, não ocore para os MMI.
- De uma certa forma, os MH são uma generalização dos MMI.
- Contudo, existem semelhanças nos processos inferencias para ambas as classes. Com efeitos, utilizaremos tal semelhança na apresentação dos procedimentos inferenciais para os MH.