MAT02035 - Modelos para dados correlacionados

Modelos lineares de efeitos mistos (continuação)

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Departamento de Estatística

Porto Alegre, 2023

Estrutura de covariância de efeitos aleatórios

Estrutura de covariância de efeitos aleatórios

- No modelo linear de efeitos mistos,

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i,$$

$R_i = \mathrm{Cov}\,(\epsilon_i)$ descreve a covariância entre as observações longitudinais ao focar no perfil de resposta média condicional de um indivíduo **específico**.

- Ou seja, é a covariância dos desvios do i-ésimo indivíduo com respeito ao seu perfil de resposta média,

$$\mathrm{E}\,(Y_i | b_i) = X_i\beta + Z_i b_i.$$

Estrutura de covariância de efeitos aleatórios

- É usualmente assumido que R_i é uma matriz diagonal, $\sigma^2 I_{n_i}$, em que I_{n_i} denota uma matriz identidade $n_i \times n_i$.
- Esta suposição é comumente referida como a **"suposição de independência condicional"**.
- Ou seja, dado os efeitos aleatórios b_i, os erros de medição são distribuídos independentemente com uma variância comum σ^2.

Estrutura de covariância de efeitos aleatórios

- Como comentamos anteriormente, no modelo linear de efeitos mistos podemos distinguir a média condicional de Y_i, dado b_i,

$$\mathrm{E}\left(Y_i|b_i\right) = X_i\beta + Z_i b_i,$$

da **média marginal** de Y_i,

$$\mathrm{E}\left(Y_i\right) = X_i\beta.$$

Estrutura de covariância de efeitos aleatórios

- De forma similar podemos distinguir entre as **covariância condicional e marginal**.
- A covariância condicional de Y_i, dado b_i, é

$$\text{Cov}\left(Y_i | b_i\right) = \text{Cov}\left(\epsilon_i\right) = R_i,$$

enquanto a covariância marginal de Y_i é

$$\begin{aligned}
\text{Cov}\left(Y_i\right) &= \text{Cov}\left(Z_i b_i\right) + \text{Cov}\left(\epsilon_i\right) \\
&= Z_i \text{Cov}\left(b_i\right) Z_i' + \text{Cov}\left(\epsilon_i\right) \\
&= Z_i G Z_i' + R_i.
\end{aligned}$$

Estrutura de covariância de efeitos aleatórios

- Mesmo quando $R_i = \text{Cov}(\epsilon_i) = \sigma^2 I_{n_i}$ (uma matriz diagonal com todas as correlações duas-a-duas iguais a zero), a matriz $\text{Cov}(Y_i)$ **possui elementos fora da diagonal diferentes de zero**, deste modo **levando em consideração a correlação entre as observações repetidas no mesmo indivíduo** em um estudo longitudinal.
- Isto é, **a introdução de efeitos aleatórios**, b_i, **induz correlação entre os componentes** de Y_i.

Estrutura de covariância de efeitos aleatórios

Comentários

- O modelo linear de efeitos mistos permite a análise explícita das fontes de variação nas respostas:
  - entre indivíduos (G);
  - e intra-indivíduo (R_i).
- A covariância marginal de Y_i é uma função do tempo de medição.
- A estrutura de covariância induzida por efeitos aleatórios [$\text{Cov}\,(Y_i) = Z_i G Z_i' + \sigma^2 I_{n_i}$] pode ser contrastada com os modelos de padrão de covariância apresentados na aula anterior.
  - Modelos de padrão de covariância não distinguem as diferentes fontes de variabilidade, enquanto que modelos lineares de efeitos mistos distinguem as fontes de variabilidade entre indivíduos e intra-indivíduo.

Estrutura de covariância de efeitos aleatórios

Comentários (continuação)

- Para os modelos lineares **com respostas contínuas**, as duas abordagens (padrão de covariância e efeitos mistos) produzem o mesmo modelo para a média marginal de Y_i [$\mathrm{E}(Y_i) = X_i\beta$], e diferem somente em termos do modelo assumido para a covariância.
- A estrutura de covariância de efeitos aleatórios **não requer** delineamento balanceado.
- Ainda, o número de parâmetros de covariância é o mesmo independente do número e as ocasiões de medições.
- Finalmente, ao contrário de muitos dos modelos de padrão de covariância que fazem suposições fortes sobre a homogeneidade da variância ao longo do tempo, a estrutura de covariância de efeitos aleatórios permite que a variância e a covariância aumentem ou diminuam em função dos tempos de medição.

Estimação via máxima verossimilhança

Estimação via máxima verossimilhança

- Note, que pelas propriedades da distribuição normal, temos que

$$Y_i \sim N(X_i\beta, Z_iGZ_i' + \sigma^2 I_{n_i}).$$

- Logo, podemos escrever a **função de verossimilhança** com base no modelo normal multivariado.
- Como esperado, o estimador de máxima verossimilhança de β é o estimador de **mínimos quadrados generalizados** (MQG) e depende da covariância marginal entre as medidas repetidas [$\text{Cov}\,(Y_i) = Z_iGZ_i' + \sigma^2 I_{n_i}$]

$$\hat{\beta} = \left\{\sum_{i=1}^{N}(X_i'[\text{Cov}\,(Y_i)]^{-1}X_i)\right\}^{-1}\sum_{i=1}^{N}(X_i'[\text{Cov}\,(Y_i)]^{-1}y_i).$$

Estimação via máxima verossimilhança

- Em geral, não há expressão simples para o estimador de máxima verossimilhança dos componentes de covariância [G e σ^2 (ou R_i)] e requer **técnicas iterativas**.
- Porque a estimativa de covariância de máxima verossimilhança é enviesada em amostras pequenas, usa-se a **estimação de máxima verossimilhança restrita (REML)**;
  - e a resultante estimativa REML de β é dada por $\hat{\beta}$ substituindo $\text{Cov}(Y_i)$ pela sua estimativa REML. # Inferência para o modelo linear de efeitos mistos

Inferência para o modelo misto

Considere o modelo

$$Y_i = X_i\beta + Z_i b_i + \epsilon_i,$$

em que, $b_i \sim N_q(0, G(\alpha))$ e $\epsilon_{ij} \sim N(0, \sigma^2)$, b_i e ϵ_{ij} independentes.

- Tem-se: p efeitos fixos e $\frac{q(q+1)}{2} + 1$ efeitos aleatórios.
- Inferência estatística para $\theta = (\beta, \alpha, \sigma^2)$:
  1. Máxima verossimilhança.
  2. Máxima verossimilhança restrita.

Inferência para o modelo misto

A função de verossimilhança é dada por:

$$
\begin{aligned}
L(\theta|y) &= \prod_{i=1}^{N} p(y_i|\theta) \\
&= \prod_{i=1}^{N} \int p(y_i, b_i|\theta) db_i \\
&= \prod_{i=1}^{N} \int p(y_i|b_i, \theta) p(b_i|\theta) db_i,
\end{aligned}
$$

em que $p(y_i|b_i, \theta) \sim N_{n_i}(X_i\beta + Z_i b_i, \sigma^2 I_{n_i})$ e $p(b_i|\theta) \sim N_q(0, G)$.

- Note que $p(y_i|\theta) \sim N_{n_i}(X_i\beta, Z_i G Z_i' + \sigma^2 I_{n_i})$.

Escolha entre modelos de covariância de efeitos aleatórios

Escolha entre modelos de covariância

- Embora o modelo linear de efeitos mistos assume que as respostas longitudinais dependem em uma combinação dos efeitos populacionais e indivíduo-específicos, quando tomamos a média com respeito a distribuição dos efeitos aleatórios

$$\mathrm{E}\,(Y_i) = X_i\beta,$$

e a covariância entre as respostas tem a estrutura distinta de efeitos aleatórios

$$\mathrm{Cov}\,(Y_i) = Z_i G Z_i' + \sigma^2 I_{n_i}.$$

Escolha entre modelos de covariância

- Da perspectiva de modelar a covariância, a estrutura de efeitos aleatórios é atraente porque o **número de parâmetros de covariância**, $q \times (q+1)/2 + 1$, é o mesmo, independentemente do número e do momento das ocasiões de medição.
- Em muitas aplicações, será suficiente incluir apenas interceptos e inclinações aleatórios para o tempo (um total de $2 \times (2+1)/2 + 1 = 4$ parâmetros de covariância), **permitindo** assim a **heterogeneidade nas variâncias** e **correlações** que podem ser expressas **como funções do tempo**.
- Em outras aplicações, uma estrutura de efeitos aleatórios mais complexa pode ser necessária.

Escolha entre modelos de covariância

- Na escolha de um modelo para a covariância, muitas vezes será interessante **comparar** dois modelos aninhados, um com q efeitos aleatórios correlacionados, outro com $q + 1$ efeitos aleatórios correlacionados.
- A diferença no número de parâmetros de covariância entre esses dois modelos é $q + 1$, pois há uma variância adicional e q covariâncias adicionais no modelo "completo".
- Conforme mencionado em aulas anteriores, o **teste da razão de verossimilhança** fornece um método válido para **comparar modelos aninhados** para a covariância.
- No entanto, em certos casos, a distribuição nula usual para o teste da razão de verossimilhança não é mais válida.

Escolha entre modelos de covariância

- Estes testes, usualmente, estão na **fronteira do espaço de parâmetros**.
  - Neste caso, a estatística da RV, sob H_0, não tem uma distribuição qui-quadrado.
- A distribuição neste caso é uma **mistura de distribuições qui-quadrado**.
  - Ou seja, por exemplo, para $H_0 : \sigma_{b_2} = 0$

$$RV \sim 0.5\chi_q + 0.5\chi_{q+1}.$$

Exemplo

- Modelo completo: $q = 2$ (intercepto e inclinação aleatórios)
- Modelo restrito: $q = 1$ (somente intercepto aleatório)
  - Teste usual **(errado)**: nível de significância 5%, o valor crítico é dado por 5,99.
  - Teste correto: $RV \sim 0.5\chi_1 + 0.5\chi_2$ nível de significância 5%, o valor crítico é dado por 5,14 **(Tabela, Apend. C, Fitzmaurice et al.)**.

Predição de efeitos aleatórios

Predição de efeitos aleatórios

- **Objetivo:** predizer perfis individuais ou identificar indivíduos acima ou abaixo do perfil médio.
- Deseja-se:

$$\widehat{Y}_i = \widehat{\mathrm{E}}(Y_i|b_i) = Xi\widehat{\beta} + Z_i\widehat{b}_i,$$

e para tal é necessário $\widehat{b}_i$, o chamado \structure{Estimador BLUP, *"Best Linear Unbiased Predictor"* de b_i.

Predição de efeitos aleatórios

- No modelo linear misto, Y_i e b_i tem uma distribuição conjunta normal multivariada.
- Usando conhecidas propriedades da normal multivariada, temos que

$$\mathrm{E}\left(b_i | Y_i, \widehat{\beta}\right) = G Z_i' \Sigma_i^{-1} (Y_i - X_i \widehat{\beta})$$

- Usando as estimativas de máxima verossimilhança dos componentes de variância,

$$\widehat{b}_i = \widehat{G} Z_i' \widehat{\Sigma}_i^{-1} (Y_i - X_i \widehat{\beta}),$$

o BLUP de b_i.

- (Abordagem **empirical Bayes**)

Predição de efeitos aleatórios

$$\widehat{Y}_i = X_i\widehat{\beta} + Z_i\widehat{b}_i = (\widehat{R}_i\widehat{\Sigma}_i^{-1})X_i\widehat{\beta} + (I_{n_i} - \widehat{R}_i\widehat{\Sigma}_i^{-1})Y_i,$$

em que $\text{Var}\,(\epsilon_i) = R_i$, e
$\widehat{\Sigma}_i\widehat{\Sigma}_i^{-1} = I_{n_i} = (Z_i\widehat{G}Z_i' + \widehat{R}_i)\widehat{\Sigma}_i^{-1} = Z_i\widehat{G}Z_i'\widehat{\Sigma}_i^{-1} + \widehat{R}_i\widehat{\Sigma}_i^{-1}$.

- **Interpretação:** média ponderada entre a média populacional $X_i\widehat{\beta}$ e o i-ésimo perfil observado.
  - Isto significa que o perfil predito é **"encolhido"** na direção da média populacional.

Predição de efeitos aleatórios

- A quantidade de "encolhimento" (***shrinkage***) depende da magnitude de R_i e Σ_i.
  - R_i: variância intra-indivíduo;
  - Σ_i: variância total (entre e intra-indivíduo).
- Quando R_i é relativamente grande, e a variabilidade intra indivíduo é maior que a variabilidade entre indivíduos, mais peso é atribuído a $X_i\widehat{\beta}$, a média populacional estimada, do que à resposta individual observada.
- Por outro lado, quando a variabilidade entre indivíduos é grande em relação à variabilidade intra-indivíduos, mais peso é dado à resposta observada Y_i.

Predição de efeitos aleatórios

- Finalmente, o grau de "encolhimento" em direção à média populacional também depende de n_i.
- Em geral, há maior encolhimento em direção à curva média populacional quando n_i é pequeno.
- Intuitivamente, isso faz sentido já que menos peso deve ser dado à trajetória observada do indivíduo quando menos dados estão disponíveis.

Avisos

- **Próxima aula:** Modelos lineares de efeitos mistos - exemplos e implementação computacional.
- **Para casa:** ler o Capítulo 8 do livro "**Applied Longitudinal Analysis**".
  - Caso ainda não tenha lido, leia também os Caps. 1, 2, 3, 4, 5, 6 e 7.

Bons estudos!