

Bayesiansk statistik, 732g43, 7.5 hp

Moment 5 - Mer om Multilevelmodeller

Bertil Wegmann

STIMA, IDA, Linköpings universitet

- Bayesiansk multilevelmodell med gruppsspecifika förklaringsvariabler
- Bayesiansk heteroskedastisk multilevelmodell
- Bayesiansk multilevelmodell med 3 nivåer (flera nivåer)
- Bayesiansk multilevelmodell med varierande intercept och lutning

■ 2 nivåer:

$$y_i \stackrel{iid}{\sim} N(\mu_i, \sigma_y)$$

$$\mu_i = \alpha_j + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i},$$

$$\alpha_j \sim N(\mu_{\alpha j}, \sigma_{\alpha}),$$

$$\mu_{\alpha j} = \alpha_0 + \alpha z_j,$$

$$\alpha_0 \sim N(0, 10),$$

$$\alpha \sim N(0, 10),$$

$$\sigma_{\alpha} \sim \text{halfcauchy}(0, 1),$$

$$\beta_k \sim N(0, 10),$$

$$\sigma_y \sim \text{halfcauchy}(0, 1),$$

där $y = IQ$, x_1 är antalet äldre syskon, x_2 är antalet yngre syskon, $x_3 = 1$ för pojke, $x_4 = SES$, x_5/x_6 är åldersklass för moderns/faderns ålder vid födseln och z_j är en vektor med grupp-specifika förklaringsvariabler. Index $j = 1, \dots, 46$ är för skolklass j och $k = 1, \dots, 6$ är för förklaringsvariabel k .

- Heteroskedastisk multilevelmodell med 2 nivåer:

$$y_i \stackrel{iid}{\sim} N(\mu_i, \sigma_{yi})$$

$$\mu_i = \alpha_j + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i},$$

$$\alpha_j \sim N(\alpha, \sigma_\alpha),$$

$$\alpha \sim N(0, 10),$$

$$\sigma_\alpha \sim \text{halfcauchy}(0, 1),$$

$$\beta_k \sim N(0, 10),$$

$$\ln \sigma_{yi} = \gamma_0 + \gamma z_i, \quad (\gamma_0, \gamma) \stackrel{iid}{\sim} N(0, 10)$$

där $y = IQ$, x_1 är antalet äldre syskon, x_2 är antalet yngre syskon, $x_3 = 1$ för pojke, $x_4 = SES$, x_5/x_6 är åldersklass för moderns/faderns ålder vid födseln och z_i är en vektor med förklaringsvariabler för variansen på elevnivå. Index $j = 1, \dots, 46$ är för skolklass j och index $k = 1, \dots, 6$ är för förklaringsvariabel k .

■ Multilevelmodell med 3 nivåer:

$$y_i \stackrel{iid}{\sim} N(\mu_i, \sigma_y)$$

$$\mu_i = \alpha_{jk} + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i},$$

$$\alpha_{jk} \sim N(\alpha_j, \sigma_2),$$

$$\alpha_j \sim N(\alpha, \sigma_3),$$

$$\alpha \sim N(0, 10),$$

$$\sigma_2 \sim \text{halfcauchy}(0, 1),$$

$$\sigma_3 \sim \text{halfcauchy}(0, 1),$$

$$\beta_l \sim N(0, 10),$$

$$\sigma_y \sim \text{halfcauchy}(0, 1),$$

där $y = IQ$, x_1 är antalet äldre syskon, x_2 är antalet yngre syskon, $x_3 = 1$ för pojke, $x_4 = SES$, x_5/x_6 är åldersklass för moderns/faderns ålder vid födseln. Index $j = 1, \dots, 7$ är för skola j , $k = 1, \dots, 46$ för skolklass k och $l = 1, \dots, 6$ för förklaringsvariabel l .

- Graduate school ansökningar för män och kvinnor till 6 olika institutioner på UC Berkley (ett av bokens datamaterial). Data kolumner innehåller för respektive kön och institution antal ansökningar ($n = \text{applications}$) samt antal beviljade ($y = \text{admit}$) och icke-beviljade ($n-y = \text{reject}$) ansökningar.
- Totalt antal ansökningar är 4526, varav 1755 stycken eller 38.8 % beviljades.
- Mål med binomialregression: anpassa en multilevelmodell med varierande intercept och lutning för att förklara sannolikheten att en ansökan beviljas.
- Varierande intercept modellerar allmänna skillnader i sannolikheten för beviljad ansökan mellan institutioner.
- Varierande lutning modellerar variation mellan institutioners beviljande av manliga och kvinnliga ansökningar.

- Modell med y = antal beviljade per kön och institution och $x_1 = 1$ om ansökan kom från en man och 0 annars:

$$y_i \stackrel{iid}{\sim} \text{bin}(n_i, p_i)$$

$$\log \left(\frac{p_i}{1 - p_i} \right) = \text{logit}(p_i) = \alpha_j + \beta x_{1i},$$

$$\alpha_j \stackrel{iid}{\sim} N(0, 10),$$

$$\beta \sim N(0, 10),$$

där n_i är antalet ansökningar (applications), p_i är sannolikheten för att en ansökan beviljas för en kombination i av kön och institution och index $j = 1, \dots, 6$, är för varje institution j .

- Modell med y = antal beviljade per kön och institution och $x_1 = 1$ om ansökan kom från en man och 0 annars:

$$y_i \stackrel{iid}{\sim} \text{bin}(n_i, p_i)$$

$$\log \left(\frac{p_i}{1 - p_i} \right) = \text{logit}(p_i) = \alpha_j + \beta x_{1i},$$

$$\alpha_j \sim N(\alpha, \sigma_\alpha),$$

$$\alpha \sim N(0, 10),$$

$$\sigma_\alpha \sim \text{halfcauchy}(0, 1),$$

$$\beta \sim N(0, 10),$$

där n_i är antalet ansökningar (applications), p_i är sannolikheten för att en ansökan beviljas för en kombination i av kön och institution och index $j = 1, \dots, 6$, är för varje institution j .

■ Modell:

$$\begin{aligned}
 y_i &\overset{iid}{\sim} \text{bin}(n_i, p_i) \\
 \log \left(\frac{p_i}{1 - p_i} \right) &= \text{logit}(p_i) = \alpha_j + \beta_j x_{1i}, \\
 \begin{bmatrix} \alpha_j \\ \beta_j \end{bmatrix} &\sim N_2 \left(\begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \mathbf{S} \right), \\
 \mathbf{S} &= \begin{pmatrix} \sigma_\alpha & 0 \\ 0 & \sigma_\beta \end{pmatrix} \mathbf{R} \begin{pmatrix} \sigma_\alpha & 0 \\ 0 & \sigma_\beta \end{pmatrix}, \\
 \alpha &\sim N(0, 10), \\
 \beta &\sim N(0, 10), \\
 \sigma_\alpha &\sim \text{halfcauchy}(0, 1), \\
 \sigma_\beta &\sim \text{halfcauchy}(0, 1), \\
 \mathbf{R} &\sim \text{LKJcorr}(2),
 \end{aligned}$$

där index $j = 1, \dots, 6$, är för varje institution j .

- Bayesiansk multilevelmodell med varierande intercept och lutning kallas även Bayesiansk varierande-lutningar modell som implicit inkluderar varierande intercept.
- Priorfördelningen (Hyperprior) för intercept och lutning, (α, β) , följer en bivariat normalfördelning.
- Standard att ta hänsyn till korrelation mellan intercept och lutning. Detta ger en regulariserande prior för grupp-specifika intercept, lutningar och korrelationen mellan dem.
- Kovariansmatrisen $\mathbf{S}$ för kovariansen mellan intercept och lutning ges som

$$\mathbf{S} = \begin{pmatrix} \sigma_{\alpha}^2 & \sigma_{\alpha}\sigma_{\beta}\rho \\ \sigma_{\alpha}\sigma_{\beta}\rho & \sigma_{\beta}^2 \end{pmatrix}.$$

- Uppdelningen av kovariansmatrisen $\mathbf{S}$ i standardavvikelser och korrelationsmatris ger enklare struktur på de varierande effekterna.
- Korrelationsmatrisen $\mathbf{R}$ ges som

$$\mathbf{R} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

- Priorn för korrelationsmatrisen $\mathbf{R}$, $LKJcorr(2)$, är en svagt informativ prior för ρ som har liten sannolikhetsmassa för extrema korrelationer nära ± 1 .
- $LKJcorr(1)$ ger en uniform prior för ρ mellan -1 och 1 . Högre värden på η i $LKJcorr(\eta)$ ger tightare prior för ρ kring 0 .

- Modellen lär sig om varje grupp-specifikt intercept och lutning (α_j, β_j) , $j = 1, \dots, J$, samtidigt som den lär sig om populationen av intercept och lutningar från den regulariserande priorn (hyperpriorn).
- Populationen av intercept och lutningar består av hyperpriorns hyperparametrar $(\alpha, \beta, \sigma_\alpha, \sigma_\beta, \rho)$.
- Hyperpriorn uppdateras från data och krymper alla intercept och lutningar mot deras gemensamma medelvärde (α, β) (*shrinkage effect*).
- Grupper med lite data påverkas mer av den poolade informationen från hyperpriorn.