Konfidenčné oblasti pre regresné parametre v lineárnom zmiešanom modeli pre longitudinálne dáta

(ROBUST 2010)

Gejza Wimmer

Králíky
3. Február 2010

Locally best linear-quadratic unbiased estimators of the covariance matrix elements in a special heteroscedastic regression model

Gejza Wimmer

ÚMS MU Brno, MÚ SAV Bratislava, FPV UMB Banská Bystrica
wimmer@mat.savba.sk

Explicit formulas are given for the locally best linear-quadratic estimators of the covariance matrix elements in the regression model $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, 2, \dots, n$, where $\varepsilon_i \sim N(0, \sigma^2(a + b(\beta_0 + \beta_1 x_i))^2)$ are independent, σ^2, a, b known constants, $x_i \neq x_j$ for $i \neq j$, $n \geq 4$. The variances of the derived estimators are also given and investigated in a special case of increasing number of measuring points x_i.

Model description

Let

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n, \tag{1}$$

where $\varepsilon_i \sim N(0, \sigma^2(a + b(\beta_0 + \beta_1 x_i))^2)$ are independent, σ^2, a, b are known positive constants, $x_i \neq x_j$ for $i \neq j$, $n \geq 4$. Model (1) is a special heteroscedastic regression model $(\mathbf{Y}, \mathbf{X}\beta, \sigma^2\Sigma)$ with observation vector $\mathbf{Y} = (Y_1, \dots, Y_n)' \sim N(\mathbf{X}\beta, \sigma^2\Sigma)$, design matrix $\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$ and covariance matrix $\mathrm{cov}(\mathbf{Y}) = \sigma^2\Sigma(\beta)$, where

$$\Sigma(\beta) = \begin{pmatrix} (a + b(\beta_0 + \beta_1 x_1))^2 & 0 & 0 & \dots & 0 \\ 0 & (a + b(\beta_0 + \beta_1 x_2))^2 & 0 & \dots & 0 \\ 0 & 0 & (a + b(\beta_0 + \beta_1 x_3))^2 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & (a + b(\beta_0 + \beta_1 x_n))^2 \end{pmatrix}$$

$\beta^{(0)}$-LBLQUE in model (1)

In the model (1) does not exist the uniformly best linear-quadratic estimator of $\sigma^2(a + b(\beta_0 + \beta_1 x_j))^2$ for any $j \in \{1, 2, \dots, n\}$. Therefore we are searching for $\beta^{(0)}$-locally best linear-quadratic unbiased estimator ($\beta^{(0)}$-LBLQUE) of $\sigma^2(a + b(\beta_0 + \beta_1 x_j))^2$ i.e. we are looking for a linear-quadratic estimator

$$\mathbf{a}'_{(j)}\mathbf{Y} + \mathbf{Y}'\mathbf{A}_{(j)}\mathbf{Y}$$

satisfying two conditions

(i) $$E_\beta(\mathbf{a}'_{(j)}\mathbf{Y} + \mathbf{Y}'\mathbf{A}_{(j)}\mathbf{Y}) = \sigma^2(a + b(\beta_0 + \beta_1 x_j))^2 \quad \forall \beta = (\beta_0, \beta_1)' \in \mathbb{R}^2 \quad \text{(unbiasedness)}$$

and

(ii) $$D_{\beta^{(0)}}(\mathbf{a}'_{(j)}\mathbf{Y} + \mathbf{Y}'\mathbf{A}_{(j)}\mathbf{Y})$$

is minimal of all unbiased linear-quadratic estimators of $\sigma^2(a + b(\beta_0 + \beta_1 x_j))^2$.

According to Theorem 2.1 in [1] there exists a unique $\beta^{(0)}$-LBLQUE of $\sigma^2(a + b(\beta_0 + \beta_1 x_j))^2$ for each $j \in \{1, 2, \dots, n\}$ and according to Lema 3.2 in [1] its form is

$$\mathbf{a}'_{(j)}\mathbf{Y} + \mathbf{Y}'\mathbf{A}_{(j)}\mathbf{Y} = (Y_j^2 - 2Y_j\mathbf{e}'_j\mathbf{X}\beta^{(0)}) - \sum_{i=1}^{n}\sum_{k=1}^{n} \frac{(Y_iY_k - (Y_i\mathbf{e}'_k\mathbf{X}\beta^{(0)} + Y_k\mathbf{e}'_i\mathbf{X}\beta^{(0)}))}{(a + b\mathbf{e}'_i\mathbf{X}\beta^{(0)})^2(a + b\mathbf{e}'_k\mathbf{X}\beta^{(0)})^2}$$

$$\cdot [w^{11} + w^{12}(2x_j + x_i + x_k) + w^{13}(x_j^2 + x_ix_k) + (w^{22} + w^{23})x_j(x_i + x_k) + w^{23}x_j(x_ix_j + x_kx_j - 2x_ix_k) + w^{33}x_ix_kx_j^2],$$

where $\mathbf{e}'_j$ is the j-th unit vector and w^{ij} are

$$w^{11} = \frac{1}{\Delta}\left\{\left[\left(\sum_{i=1}^{n}\frac{x_i^2}{(a + b\mathbf{e}'_i\mathbf{X}\beta^{(0)})^2}\right)^2 - \sum_{i=1}^{n}\frac{x_i^4}{(a + b\mathbf{e}'_i\mathbf{X}\beta^{(0)})^4}\right] \cdot \dots \right\}$$

[illegible]

$$w^{12} = \frac{1}{\Delta}\left\{\dots\right\}$$

[illegible]

$$w^{22} + w^{13} = \frac{1}{\Delta}\left\{\dots\right\}$$

[illegible]

$$w^{23} = \frac{1}{\Delta}\left\{\dots\right\}$$

[illegible]

$$w^{33} = \frac{1}{\Delta}\left\{\dots\right\}$$

[illegible]

with

$$\Delta = \dots$$

[illegible]

Dispersion of the $\beta^{(0)}$-LBLQUE in model (1)

According to Lemma 3.4 in [1] the dispersion of the $\beta^{(0)}$-LBLQUE of $\sigma^2(a + b(\beta_0 + \beta_1 x_j))^2$ is

$$D_{\beta^{(0)}}(\mathbf{a}'_{(j)}\mathbf{Y} + \mathbf{Y}'\mathbf{A}_{(j)}\mathbf{Y}) =$$

$$= 2\sigma^4[(a + b\mathbf{e}'_j\mathbf{X}\beta^{(0)})^4 + w^{11} + 4x_jw^{12} + 2x_j^2(w^{13} + w^{22} + w^{23}) + 4x_j^3w^{23} + x_j^4w^{33}].$$

Special limiting case

In the case $x_i = 2^i + \frac{2}{i}x_i$, $i = 1, 2, \dots$, $a = \sigma^2 = 1$, $b > 0$, $\beta^{(0)} = (0, 1)'$ after tedious but straightforward calculation (see Appendix 2 in [1]) we obtain that the dispersion of the $\beta^{(0)}$-LBLQUE of $\sigma^2(a + b(\beta_0 + \beta_1 x_j))^2$ at $\beta^{(0)}$ tends to $2\sigma^4[(1 + bx_j)^4 + \dots(b + 1)^4]$ for $i = 1, 2, \dots$

References

[1] Wimmer G. (1997) *Covariance Matrix Elements Estimation: Special Linear Model Without and With Repeated Measurements.* Mathematica Slovaca 47, 299–315.

Úvod

- longitudinálne dáta
- lineárny zmiešaný model (LZM)
  - zameranie sa na model s AR(1) chybami
- odhad regresných parametrov
  - kovariančné parametre sú známe
  - kovariančné parametre sú neznáme
    - odhadnúť neznáme kovariančné parametre - REML
- charakterisiky odhadu regresných parametrov
- konfidenčné oblasti pre lineárnu kombináciu regresných parametrov
  - "malý" počet pozorovaní
  - je nevyhnutné odhadovať i kovariančné parametre modelu

Ďakujem za pozornosť