

Ei-parametrinen regressio

Harjoitus 2

1.4.2008

1. Osoita, että $SS = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ on esitettävissä muodossa

$$SS = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\mathbf{X}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}),$$

missä $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.

2. Osoita, että sekamallissa

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}$$

parametrien $\boldsymbol{\beta}$ ja \mathbf{u} BLUP on kirjoitettavissa muodossa

$$\begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{pmatrix} = (\mathbf{C}'\mathbf{R}^{-1}\mathbf{C} + \mathbf{B})^{-1}\mathbf{C}'\mathbf{R}^{-1}\mathbf{y},$$

missä $\mathbf{C} = (\mathbf{X}, \mathbf{Z})$ ja

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1} \end{pmatrix},$$

missä $Var(\mathbf{u}) = \mathbf{D}$ ja $Var(\boldsymbol{\epsilon}) = \mathbf{R}$.

3. Osoita lisäksi, että

$$Var(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$$

ja

$$Var\left\{\begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{pmatrix} \mid \mathbf{u}\right\} = (\mathbf{C}'\mathbf{R}^{-1}\mathbf{C} + \mathbf{B})^{-1}\mathbf{C}'\mathbf{R}^{-1}\mathbf{C}(\mathbf{C}'\mathbf{R}^{-1}\mathbf{C} + \mathbf{B})^{-1}.$$

4. Olkoon maidontuotantoaineistossa (seuraava sivu) kaksi vaikutusta:
- isän geneettinen arvo (SIRE), satunnainen
- lauma (HERD), kiinteä

Estimoi muuttujalle *Yield* sekamalli (kiinteät ja satunnaiset vaikutukset), kun oletetaan, että

$$Var(\boldsymbol{\epsilon}) = \mathbf{I}, \quad Var(\mathbf{u}) = 0.1\mathbf{I} \quad \text{ja} \quad Cov(\boldsymbol{\epsilon}, \mathbf{u}) = \mathbf{0}.$$

5. Osoita, että

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}\hat{\mathbf{u}} = \mathbf{C}(\mathbf{C}'\mathbf{R}^{-1}\mathbf{C} + \mathbf{B})^{-1}\mathbf{C}'\mathbf{R}^{-1}\mathbf{y}$$

ja laske muuttujan *Yield* ennustearvot kaikille yksilöille.

Herd	Sire	Yield
1	A	110
1	D	100
2	B	110
2	D	100
2	D	100
3	C	110
3	C	110
3	D	100
3	D	100