Mat-1.3621 Tilastollinen päättely

10. harjoitukset / Tehtävät

Aiheet: **Suurimman uskottavuuden menetelmä, asymptoottinen teoria ja yleinen lineaarinen malli**

Avainsanat:

Asymptoottinen normaalisuus, Cramérin ja Raon alaraja, Derivaatta, Diagnostinen testi, Estimaattori, Estimointi, *F*-jakauma, *F*-testi, Fisherin informaatiomatriisi, Gradientti, Harhattomuus, Havainto, Havaintoarvo, Hessen matriisi, Homoskedastisuus, Hylkäysalue, Jakaumaoletus, Jäännösneliösumma, Jäännöstermi, Jäännösvarianssi, Khi^2-jakauma, Kokonaisneliösumma, Korrelaatio, Korreloimattomuus, Kovarianssimatriisi, Kriittinen arvo, Lagrangen kertojatesti, Lineaarinen regressiomalli, Logaritminen uskottavuusfunktio, Luottamuskerroin, Luottamustaso, Luottamusväli, Maksimointi, Minimointi, Nollahypoteesi, Otos, Osamäärätesti, Parametri, Pienimmän neliösumman menetelmä, Rajoitus, Realisaatio, Regressiofunktio, Regressiokerroin, Regressiomalli, Regressiotaso, Residuaali, Riippumattomuus, Satunnaismuuttuja, Selitettävä muuttuja, Selittäjä, Selittävä muuttuja, Selitysaste, Side-ehto, Sovite, Suurimman uskottavuuden estimaattori, Suurimman uskottavuuden menetelmä, *t*-jakauma, *t*-testi, Tarkentuvuus, Tehokas pistemäärä, Tehokkuus, Testi, Tyhjentävyys, Uskottavuusfunktio, Vaihtoehtoinen hypoteesi, Vakiotermi, Waldin testi, Vapausasteet, Varianssianalyysihajotelma, Yhteisjakauma, Yleinen lineaarinen malli

Tehtävä 11.1.

Olkoon

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\,,\ \mathbf{X}\ n\times(k+1)$$

yleinen lineaarinen malli, joka toteuttaa ns. standardioletukset ja olkoon regressiokertoimien vektorin $\boldsymbol{\beta}$ *PNS-estimaattori*

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Todista:

(a) $\mathbf{b}$ on *harhaton* eli $\mathrm{E}(\mathbf{b}) = \boldsymbol{\beta}$

(b) $\mathrm{Cov}(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

Tehtävä 11.2.

Olkoon

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\,,\ \mathbf{X}\ n\times(k+1)$$

yleinen lineaarinen malli, joka toteuttaa ns. standardioletukset ja olkoon regressiokertoimien vektorin $\boldsymbol{\beta}$ *PNS-estimaattori*

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Määritellään estimoidun regressiomallin *sovitteiden* muodostama vektori kaavalla

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$$

ja vastaavien *residuaalien* muodostama vektori kaavalla

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$$

Näytä, että $\hat{\mathbf{y}} = \mathbf{P}\mathbf{y}$, jossa matriisi

$$\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

on *symmetrinen* ja *idempotentti* eli *projektio* matriisin **X** sarakeavaruuteen eli matriisin **X** sarakkeiden virittämään vektoriavaruuteen (tasoon) ja

$$r(\mathbf{P}) = k + 1$$

Näytä, että $\mathbf{e} = \mathbf{M}\mathbf{y}$, jossa matriisi

$$\mathbf{M} = \mathbf{I} - \mathbf{P}$$

on *symmetrinen* ja *idempotentti* eli *projektio* matriisin **X** sarakeavaruutta vastaan kohtisuoraan vektoriavaruuteen ja

$$r(\mathbf{M}) = n - k - 1$$

Näytä lisäksi, että

$$\mathbf{PM} = \mathbf{MP} = \mathbf{0}$$

Todista myös seuraavat kaavat:

(a) $\hat{\mathbf{y}}'\mathbf{e} = 0$

(b) $\mathbf{1}'\mathbf{e} = 0$

(c) $\mathbf{1}'\mathbf{y} = \mathbf{1}'\hat{\mathbf{y}}$

(d) $\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \mathbf{e}'\mathbf{e}$

(e) $\mathrm{E}(\mathbf{e}) = 0$

(f) $\mathrm{Cov}(\mathbf{e}) = \sigma^2\mathbf{M}$

Tehtävä 11.3.

Todista: Jos satunnaismuuttuja $\boldsymbol{x}$ noudattaa p-ulotteista multinormaalijakaumaa $\mathrm{N}_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$, niin neliömuoto

$$(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \sim \chi^2(p)$$

Tehtävä 11.4.

Olkoon

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\,,\ \mathbf{X}\ n\times(k+1)$$

yleinen lineaarinen malli, joka toteuttaa ns. standardioletukset.

Muodostetaan nollahypoteesi

$$\mathrm{H}_0 : \boldsymbol{\beta} = \mathbf{0}$$

Todista, että F-testisuure

$$F = \frac{n-p}{p} \cdot \frac{\mathbf{y}'\mathbf{P}\mathbf{y}}{\mathbf{y}'\mathbf{M}\mathbf{y}} \sim F(p, n-p)$$

jos nollahypoteesi H_0 pätee.

Tehtävä 11.5.

Olkoon

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\,,\ \mathbf{X}\ n \times (k+1)$$

on tavanomaiset oletukset toteuttava yleinen lineaarinen malli, jonka regressiokertoimien vektoria $\boldsymbol{\beta}$ sitoo lineaarinen rajoitus

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$

jossa $\mathbf{R}$ täysiasteinen $m \times (k+1)$-matriisi, $m \leq k+1$.

Johda regressiokertoimien vektorin $\boldsymbol{\beta}$ rajoitettu PNS-estimaattori ja todista, että se on parempi kuin tavanomainen PNS-estimaattori Gaussin ja Markovin lauseen mielessä, jos rajoitukset pätevät.

Ohje: Käytä *Lagrangen keinoa* minimin määräämiseen.