

Matematiikan ja tilastotieteen laitos
Alueittaisten aineistojen spatiaalinen analyysi

Loppukoe 3.4.2008

Kirjoita vastauspaperiin **nimesi** lisäksi **opiskelijanumero** tai **henkilötunnus**

1. Oletetaan, että satunnaisvektori $\mathbf{x} = [x_1 \ x_2 \ x_3]^T$ noudattaa 3-ulotteista normaalijakaumaa odotusarvovektorilla $[0 \ 0 \ 0]^T$ ja tarkkuusmatriisilla

$$\mathbf{Q} = \begin{bmatrix} 1 & -\rho & 0 \\ -\rho & 2 & -\rho \\ 0 & -\rho & 1 \end{bmatrix},$$

missä $0 < \rho < 1$, eli \mathbf{x} :n jakauman tiheysfunktio on

$$p(\mathbf{x}) = C(\rho) \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{Q} \mathbf{x}\right),$$

missä $C(\rho)$ on \mathbf{x} :n arvosta riippumaton vakio. Johda satunnaismuuttujan x_2 ehdollinen jakauma $p(x_2 | x_1, x_3)$.

2. Brookin lemma voidaan esittää kaavana

$$\frac{p_{\mathbf{Y}}(\mathbf{y})}{p_{\mathbf{Y}}(\mathbf{y}_0)} = \prod_{i=1}^n \frac{p_i(y_i | Y_j = y_{j0}, j < i; Y_j = y_j, j > i)}{p_i(y_{i0} | Y_j = y_{j0}, j < i; Y_j = y_j, j > i)},$$

missä $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$ ja $\mathbf{y}_0 = [y_{10} \ y_{20} \ \dots \ y_{n0}]^T$.

- Mitä tässä kaavassa tarkoittaa $p_{\mathbf{Y}}(\mathbf{y})$?
 - Entäpä $p_i(y | Y_j = y_j, j \neq i)$?
 - Mitä tarkoittaa CAR-malli?
 - Miten Brookin lemma liittyy CAR-malleihin?
3. Kuvaile Monte Carlo Newton–Raphson -menetelmän ja Monte Carlo likelihood -menetelmän periaatteet ja olennaiset eroavaisuudet.

4. n alueelta havaittiin muuttujien x ja y arvot (x_i, y_i) , $i = 1, 2, \dots, n$, ja havaittuun aineistoon sovitettiin malli

$$y_i = a + bx_i + u_i$$

- i) olettaen jäännökset u_i samoin jakautuneiksi ja keskenään riippumattomiksi odotusarvoilla $Eu_i = 0$, $i = 1, \dots, n$, ja
 ii) olettaen, että

$$u_i = \rho \sum_{j=1}^n w_{ij} u_j + e_i,$$

missä e_i :t ovat samoin jakautuneita ja keskenään riippumattomia odotusarvoilla $Ee_i = 0$, $i = 1, \dots, n$,

$$w_{ij} = \begin{cases} 1 & \text{jos alueet } i \text{ ja } j \text{ ovat rajanaapureita,} \\ 0, & \text{jos eivät ole,} \end{cases}$$

ja ρ on aineistosta estimoitava parametri.

Parametrien a , b ja ρ estimaatit, niiden keskivirheet ja p -arvot nollahypoteesien $H_0 : a = 0$, $H_0 : b = 0$ $H_0 : \rho = 0$ testeissä olivat

	Parametri	estimaatti	keskivirhe	p -arvo
i)	a	-2.5328	0.4104	1.16e-06
	b	0.3018	0.1368	0.0358
	Parametri	estimaatti	keskivirhe	p -arvo
ii)	a	-0.65610	0.92594	0.4786
	b	0.25252	0.15959	0.1136
	ρ	0.44226	0.026412	< 2.22e-16

- a) Tulkitse oletuksella i) saadut tulokset.
 b) Tulkitse oletuksella ii) saadut tulokset.
 c) Mistä arvelet erilaisten tulosten johtuvan.
 d) Kumman mallin valitsisit? Miksi?
 e) Mikä on johtopäätöksesi x :n ja y :n välisestä yhteydestä?
 f) Millä nimellä oletuksista ii) johdettu malli tunnetaan?