

Helsingin yliopisto
Matematiikan ja tilastotieteen laitos
Latenttien muuttujien hierarkkiset mallit
Uusintatentti 15.6.2005

VASTAA KOLMEEN KYSYMYKSEEN. LISÄSUORITUKSISTA LISÄPISTEITÄ

1. Tarkastellaan lineaarista satunnaistermimallia

$$y_j = \mathbf{x}'_{ij}\beta + \mathbf{z}'_{ij}\eta_j + \epsilon_{ij},$$

jossa \mathbf{x}_{ij} on kiinteiden selittäjien vektori havainnolle/klusterille $j, j = 1, \dots, K$, β niitä vastaava regressiokerroinvektori ja \mathbf{z}_{ij} latentteja muuttujia η_j vastaava kiinteä design-muuttujien vektori. Oletetaan, että $\eta_j \sim N(0, \psi)$, $\epsilon_{ij} \sim N(0, \theta)$.

a) Kirjoita satunnaistasomallin (random intercept model) hierarkkinen esitys ja siihen liittyvät oletukset havaintojen ja latenttien muuttujien keskinäisestä ja välisestä riippuvuudesta.

b) Mikä on mallin parametrien tulkinta (sekä regressioparametrien β että 2-tason parametrien)?

2. Havaintojen yhteisjakauma $p(y_1, y_2, \dots, y_n)$ kuvaa niiden yhteistodennäköisyyttä. Havaintoja kutsutaan *vaihdettaviksi* (exchangeable), jos uuden havainnon havaitsemisjärjestyksellä ehdolla aikaisemmat havainnot ei ole merkitystä yhteisjakauman kannalta, ts. havainnot ovat tässä mielessä riippumattomia.

a) Millä ehdolla hierarkkisen mallin havainnot ovat vaihdettavia? Esitä tämä ehto tehtävän 1 satunnaistasomallissa (sanallinen esitys ei riitä)

b) Miten mittaat hierarkkisessa mallissa 'ei-vaihdettavien' havaintojen riippuvuutta? Käytä jälleen tehtävän 1 määrittelyjä (sanallinen esitys ei riitä)

3. Tarkastellaan yleistä latenttien muuttujien lineaarista mallia, jossa

$$y = X\beta + Z\zeta + \epsilon, \quad \epsilon \sim N(0, \Theta), \quad \zeta \sim N(0, \Psi)$$

a) Johda latenttien residuaalimuuttujien ζ posterioriodotusarvo $E(\zeta|y)$ ehdolla havainnot käyttäen allaolevaa opastusta. Miten tulokset tulosta?

b) Mikä on latentin residuaalin ζ_j posterioriodotusarvo tehtävän 1 satunnaistasomallissa, jossa design-matriisin Z alkiot saavat vain arvoja 0 tai 1 toisin kuin kohdassa a)?

Opastus: käytä multinormaalijakauman ehdollisia jakaumia koskevia tuloksia havainnoille ja latenteille muuttujille. Yleisesti: jos satunnaismuuttujien V ja W yhteisjakauma on $(V, W) \sim N(\mu, \Sigma)$ niin $V|W \sim N(\mu_{V|W}, \Sigma_{V|W})$, jossa $\mu_{V|W} = \mu_V + \Sigma_{VW}\Sigma_{WW}^{-1}(W - \mu_W)$ ja $\Sigma_{V|W} = \Sigma_{VV} - \Sigma_{VW}\Sigma_{WW}^{-1}\Sigma_{WV}$.

4. Latenttien rakenteiden merkitys ja käyttö mallintamisessa