**Tilastollisen päättelyn jatkokurssi, sl 2011, HT 4, viikko 40**

**1.** Oletetaan, että havaintoja vastaavat satunnaismuuttujat $Y_1, ..., Y_n$ ovat riippumattomia ja $Y_i \sim P(\mu_i(\theta))$, jossa odotusarvo

$$\mu_i(\theta) = \exp\{\alpha + \beta x_i\}$$

riippuu kiinteästä selittävästä muuttujasta $x_i$ ja parametrista $\theta = (\alpha, \beta) \in \mathbb{R}^2$. Johda aineistoa kuvaava tilastollinen malli ja parametrin $\theta$ uskottavuusfunktio, pistemääräfunktio, havaittu informaatiomatriisi ja Fisherin informaatiomatriisi.

*Huom.*: Satunnaismuuttujan $Y \sim P(\mu)$ pistetodennäköisyysfunktio on $f(y;\mu) = \frac{1}{y!}\mu^y e^{-\mu}$, $y = 0, 1, ...,$ $\mu > 0$. Lisäksi $\mathsf{E}(Y) = \mu$ ja $\mathsf{Var}(Y) = \mu$.

**2.** (Jatkoa HT:lle 3.4). (i) Johda parametrin $\theta = (\phi, \sigma^2)$ pistemääräfunktio ja havaittu informaatiomatriisi.

(ii) Osoita, että parametrit $\phi$ ja $\sigma^2$ ovat ortogonaaliset.

**3.** Olkoon aineistoa $\boldsymbol{w} = (w_1, ..., w_n)$ vastaava satunnaisvektori $\boldsymbol{W} = (W_1, ..., W_n)$. Ositetaan $W_i = (Y_i, X_i)$ $(i = 1, ..., n)$ ja tarkastellaan monisteen yhtälössä (2.16) (s. 21) määritellyn lineaarisen regressiomallin epälineaarista yleistystä

$$Y_i = g(Z_i; \beta) + \varepsilon_i, \quad \beta \in B \subseteq \mathbb{R}^p, \ \ \sigma^2 > 0,$$

jossa $Z_i$ $(p \times 1)$ on vektorin $(\boldsymbol{W}_{i-1}, X_i)$ osavektori, $\varepsilon_1, ..., \varepsilon_n \sim \mathsf{N}(0, \sigma^2)$ $\perp\!\!\!\perp$ ja $(\boldsymbol{W}_{i-1}, X_i)$ $\perp\!\!\!\perp$ $\varepsilon_i$ kaikilla $i = 1, ..., n$. Funktiolla $\beta \mapsto g(z; \beta)$ oletetaan olevan jatkuvat toisen kertaluvun osittaisderivaatat kaikilla $z \in \mathbb{R}^k$ (epälineaarisessa tapauksessa selittävien muuttuja vektorin $Z_i$ ja parametrivektorin $\beta$ ei tarvitse olla samaa dimensiota).

(i) Johda parametrin $\theta = (\beta, \sigma^2)$ ehdollinen uskottavuusfunktio $L^{(c)}(\theta; \boldsymbol{w})$, kun ehdollisen mallin ja marginaalimallin yhteydessä tehdyt yleiset oletukset ovat voimassa (ks. moniste s. 20-21).

(ii) Osoita, että parametrin $\beta$ SU–estimaatti $\hat{\beta}$ saadaan minimoimalla epälineaarinen jäännösneliösummafunktio

$$S(\beta) = \sum_{i=1}^{n} (y_i - g(z_i; \beta))^2$$

ja että parametrin $\sigma^2$ SU–estimaatti $\hat{\sigma}^2$ saadaan kaavalla $\hat{\sigma}^2 = n^{-1} S(\hat{\beta})$ (minimipisteen olemassaolo oletetaan).

*Vihje*: Kohdassa (i) voi käyttää HT:n 3.3 tulosta.

**4.** Jatkoa edelliselle. Johda parametrin $\theta$ pistemääräfunktio ja havaittu informaatiomatriisi.

(jatkuu seuraavalla sivulla)

**5.** Edelleen jatkoa tehtävälle 3. Johda esitys parametrin $\theta$ Fisherin informaatiomatriisille (olettaen tarvittavien momenttien äärellisyys) ja totea erityisesti parametrien $\beta$ ja $\sigma^2$ ortogonaalisuus.

*Huom.* 1: Olettamalla tehtävässä 3 $g(Z_i; \beta) = Z_i' \beta$ saadaan vastaavat lineaarista mallia koskevat tulokset.

*Huom.* 2: Muuta monisteen s. 24, rivillä 6 alhaalta, oleva teksti muotoon "Edellä mainittu parametrin $\beta$ pistemäärän odotusarvoa koskeva tulos voidaan perustella myös ... " ja korvaa rivillä 2 alhaalta sana "Pistemäärä" sanoilla "Parametrin $\beta$ pistemäärä".