Tilastollinen päättely III, kevät 2016
Laskuharjoitukset 4, palautus 18.4.2016 klo 23:55 mennessä kurssin Moodle-sivulle.

1. Tarkastellaan Poisson-regressiomallia jossa $Y_1, \ldots, Y_n \perp\!\!\!\perp$, $Y_i \sim P(\mu_i(\alpha, \beta))$, kun $\mu_i(\alpha, \beta) = \exp(\alpha + \beta x_i)$ kaikilla $i = 1, \ldots, n$ ja $x_1, \ldots, x_n$ ovat kiinteitä tunnettuja lukuja. Johda tilastollisen mallin lauseke $f_{\boldsymbol{Y}}(\boldsymbol{y}; \alpha, \beta)$ sekä vastaava log-uskottavuusfunktio ja pistemääräfunktio.

2. Jatkoa edelliseen tehtävään. Varmista martingaalin määritelmän avulla, että lauseen 2.1(ii) tulos pätee: Pistemäärä $s_n(\alpha, \beta; \boldsymbol{Y}_n)$ on martingaali informaatiojoukon $\boldsymbol{Y}_n = (Y_1, \ldots, Y_n)$ suhteen. *Huomaa: Satunnaisvektorijono on martingaali jos vain jos sen jokainen komponentti on.*

3. Jatkoa harjoituksen 3 tehtäviin 2 ja 3 (autoregressiivinen aikasarja). Johda suurimman uskottavuuden estimaattorin $\hat{\phi}$ lauseke. Järkeile ilman tarkkaa todistusta, että $\hat{\phi}$ ei yleensä ole harhaton, so. ei päde $\mathsf{E}(\hat{\phi}) = \phi$. Vertaa tilannetta tavalliseen regressiomalliin $Y_i = \beta x_i + \varepsilon_i$ ja sen parametrin β estimointiin.

   Opetus. Yleisemmin muistiinpanojen sivujen 21–22 lineaarisessa mallissa estimaattori $\hat{\beta}$ (ks. yhtälö (2.18)) ei ole välttämättä harhaton, mikäli selittävien muuttujien vektorissa Z_i on mukana vastemuuttujan aikaisempia arvoja.

4. Olkoot X ja Z jatkuvasti jakautuneita k-ulotteisia satunnaisvektoreita siten, että $X \perp\!\!\!\perp Z$. Osoita, että satunnaisvektorin $Y = X + Z$ ehdollinen jakauma ehdolla $X = x$ on sama kuin satunnaisvektorin $x + Z$ jakauma eli symbolisesti

$$Y \mid (X = x) \overset{d}{=} x + Z.$$

   Jos esimerkiksi $Z \sim \mathsf{N}_k(\mu, \Sigma)$, niin $Y \mid (X = x) \sim \mathsf{N}_k(x + \mu, \Sigma)$.

   Tätä (kenties intuitiivisesti selvää) tulosta käytettiin mm. sivulla 18 autoregressiivisen mallin yhteydessä ja sivulla 22 pääteltäessä yhtälöstä (2.16) tulos (2.15).

5. Tarkastellaan vielä luennoilla ja monisteessa mainittua autoregressiivista mallia, johon on lisätty vakiotermi ja "ulkopuolinen" satunnainen selittäjä:

$$Y_i = \alpha + \phi Y_{i-1} + \gamma X_i + \varepsilon_i, \qquad i = 1, \ldots, n.$$

   Oletetaan, että satunnaisvektorin $\boldsymbol{X} = (X_1, \ldots, X_n)$ jakauma riippuu jostakin parametrista λ ja $\boldsymbol{X} \perp\!\!\!\perp (\varepsilon_1, \ldots, \varepsilon_n)$, jossa tavalliseen tapaan $\varepsilon_1, \ldots, \varepsilon_n \sim \mathsf{N}(0, \sigma^2) \perp\!\!\!\perp$. Lisäksi $Y_0 = y_0$ on tunnettu vakio. Kiinnostava parametri on $\psi = (\alpha, \phi, \gamma, \sigma^2)$.

   Sijoitetaan tämä malli monisteen sivuilla 20–22 kuvatun ehdollisen mallin ja marginaalimallin viitekehykseen. Perustele, että

   a) marginaalimalli on $\boldsymbol{X}$:n yhteisjakauma, joka riippuu vain parametrista λ, ja

   b) ehdollinen malli voidaan tulkita seuraavien toisistaan riippumattomien muuttujien yhteisjakaumana:

$$Y_i \mid (Y_{i-1} = y_{i-1}, X_i = x_i) \overset{d}{=} \alpha + \phi y_{i-1} + \gamma x_i + \varepsilon_i, \qquad i = 1, \ldots, n,$$

   joka riippuu vain parametrista ψ.

   c) Mikä on ehdollinen uskottavuusfunktio $L^{(c)}(\psi; \boldsymbol{w})$? Tässä $\boldsymbol{w}$ koostuu kaikista havainnoista y_i ja x_i kuten monisteessa.

Vihjeitä seuraavalla sivulla!

Vihjeitä:

4. Parin (X, Z) ytf on helppo. Johda muunnoskaavan avulla parin (X, Y) ytf. Voit halutessasi olettaa, että $k = 1$.

5. Ideoita: X_i:n ehdollistaminen $\boldsymbol{W}_{i-1}$:n suhteen merkitsee samaa kuin sen ehdollistaminen vektorin $(X_1, \ldots, X_{i-1}, \varepsilon_1, \ldots, \varepsilon_{i-1})$ suhteen ja edelleen vektorin $(X_1, \ldots, X_{i-1})$ suhteen. Käytä tuloesitystä $\boldsymbol{X}$:n jakaumalle. Uskottavat perustelut riittävät; ei tarvitse todistaa kaikkea ehdollisen jakauman määritelmästä lähtien.