Tilastollisen päättelyn jatkokurssi
5. harjoitus (4. 12. 2012)

1. Tarkastellaan muistiinpanojen yhtälössä (2.16) esitettyä lineaarista mallia $Y_i = Z_i'\beta + \varepsilon_i$, $i = 1, \dots, n$, ja sen su-estimaattoreita yhtälössä (2.18) (sivut 21–22). Oletetaan, että satunnaisvektorit Z_i ja virheet ε_i toteuttavat sivulla 27 mainitut ehdot: $\frac{1}{n}\sum_i Z_i Z_i' \xrightarrow{p} Q$ (jossa Q on kiinteä ja positiivisesti definiitti) ja $\frac{1}{n}\sum_i Z_i\varepsilon_i \xrightarrow{p} 0$, kun $n \to \infty$. Sivulla 28 on todettu, että tällöin $\hat{\beta}$ on tarkentuva. Osoita, että myös $\hat{\sigma}^2$ on tarkentuva eli $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$.

   Ohje. Aloita kirjoittamalla $\hat{\sigma}^2$:n lausekkeessa $Y_i - Z_i'\hat{\beta} = (Y_i - Z_i'\beta) - Z_i'(\hat{\beta} - \beta)$ ja korottamalla neliöön. Tarvitset oletusten lisäksi suurten lukujen lakia ja lausetta 1.1.

2. Tarkastellaan mallia $f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta)$, $\theta \in \Theta$, jossa mitkään kaksi eri θ:n arvoa eivät vastaa samaa jakaumaa (ts. θ on *identifioituva*). Olkoon $l(\theta; \boldsymbol{y})$ mallin log-uskottavuusfunktio, ja oletetaan, että odotusarvo

   $$l^*(\theta) = \mathsf{E}_{\theta_0}[l(\theta; \boldsymbol{Y})]$$

   on olemassa kaikilla $\theta \in \Theta$, kun θ_0 on todellinen parametriarvo. Todista, että $l^*(\theta)$ saa suurimman arvonsa täsmälleen pisteessä $\theta = \theta_0$.

   Ohje. Todennäköisyyslaskennan Jensenin epäyhtälö sanoo, että jos g on aidosti ylöspäin kupera (eli aidosti konkaavi) funktio ja X on satunnaismuuttuja, niin $\mathsf{E}[g(X)] \leq g(\mathsf{E}(X))$ ja yhtäsuuruus pätee vain jos X on vakio (tn:llä 1). Sovella tätä logaritmifunktioon ja muuttujaan $X = f_{\boldsymbol{Y}}(\boldsymbol{Y}; \theta)/f_{\boldsymbol{Y}}(\boldsymbol{Y}; \theta_0)$. Voit olettaa, että kyseessä on jatkuva malli eli ytf.

3. Pidetään tunnettuna alempana esitetty suurten lukujen lain tasainen versio (jonka edistyneemmät ja topologiaa tuntevat voivat yrittää lisätehtävänä todistaa). Osoita sen ja edellisen tehtävän avulla seuraava iid-tapauksen tarkentuvuutta koskeva tulos:

   Lause. Olkoot havaintoja vastaavat satunnaismuuttujat Y_i riippumattomia ja samoin jakautuneita, ja olkoon Y_i:n tf/ptf $f(y; \theta)$, joka riippuu parametrista $\theta \in \Theta$. Oletetaan:

   i) $f(\,\cdot\,; \theta_1)$ ja $f(\,\cdot\,; \theta_2)$ ovat eri jakaumia aina kun $\theta_1 \neq \theta_2$
   ii) parametriavaruus Θ on kompakti (so. suljettu ja rajoitettu joukko)
   iii) $l_1(\theta; y) = \log f(y; \theta)$ on jatkuva θ:n suhteen jokaisella y
   iv) $|l_1(\theta; y)| \leq k(y)$ kaikilla y ja $\theta \in \Theta$, jossa $\mathsf{E}_{\theta_0}[k(Y_1)] < \infty$.

   Tällöin lauseen 2.2 oletukset toteutuvat valinnalla $\bar{l}(\theta) = \mathsf{E}_{\theta_0}[l_1(\theta; Y_1)]$, jossa θ_0 on todellinen parametriarvo. Siten $\hat{\theta}_n$ on tarkentuva.

   Tasainen SLL. Olkoot $X_1, X_2, \dots$ riippumattomia ja samoin jakautuneita, ja olkoon $h(x, \theta)$ funktio, joka on jatkuva θ:n suhteen kompaktissa joukossa Θ jokaisella x. Oletetaan, että on olemassa funktio $k(x)$ siten, että $|h(x, \theta)| \leq k(x)$ kaikilla x ja $\theta \in \Theta$ sekä $\mathsf{E}[k(X_1)] < \infty$. Tällöin

   $$\frac{1}{n}\sum_{i=1}^{n} h(X_i, \theta) \xrightarrow{p} \mu(\theta) \quad \text{tasaisesti } \Theta\text{:ssa,}$$

   kun $\mu(\theta) = \mathsf{E}[h(X_1, \theta)]$.

KÄÄNNÄ!

4. (Hieman erikoisempi su-estimaattorin asymptoottinen jakauma.) Tarkastellaan mallia $Y_1, \dots, Y_n \sim \mathrm{Tas}(0, \theta) \perp\!\!\!\perp$, jossa $\theta > 0$. Palauta mieleen, että θ:n su-estimaattori on $\hat{\theta}_n = Y_{(n)} = \max(Y_1, \dots, Y_n)$. Osoita, että

$$n(\theta - \hat{\theta}_n) \xrightarrow{d} \mathrm{Exp}(1/\theta),$$

eli symbolisesti

$$\hat{\theta}_n \underset{as}{\sim} \theta - \tfrac{1}{n}\mathrm{Exp}(1/\theta),$$

jossa $\mathrm{Exp}(1/\theta)$ viittaa eksponenttijakaumaan, odotusarvona θ. Miksi lause 2.3 ei toteudu tässä mallissa?

Apu. Tutki kertymäfunktiota ja muista jakaumasuppenemisen määritelmä. Asiaan liittyviä tarkasteluja on tehty aineopintojen päättelyn kurssin monisteen harjoitustehtävässä 3.10. Muista myös, että $\lim_{n\to\infty}(1 + x/n)^n = e^x$ kaikilla $x \in \mathbb{R}$.