Tilastollisen päättelyn jatkokurssi
5. harjoitus (3. 12. 2013)

1. Tarkastellaan muistiinpanojen yhtälössä (2.16) esitettyä lineaarista mallia $Y_i = Z_i'\beta + \varepsilon_i$, $i = 1, \ldots, n$, ja sen su-estimaattoreita yhtälössä (2.18) (sivut 21–22). Oletetaan, että satunnaisvektorit Z_i ja virheet ε_i toteuttavat sivulla 28 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. Satunnaismuuttujan (tai -vektorin) Y ptf/tf $f(y;\theta)$ riippuu parametrista $\theta \in \Theta$ ja mitkään kaksi eri θ:n arvoa eivät vastaa samaa jakaumaa. Olkoon $l(\theta; y) = \log f(y;\theta)$, ja oletetaan, että odotusarvo $l^*(\theta) = \mathsf{E}_{\theta_0}\big[l(\theta; Y)\big]$ 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 $\mathsf{E}[g(X)] \le g(\mathsf{E}(X))$ ja yhtäsuuruus pätee vain jos X on vakio (tn:llä 1). Sovella tätä logaritmifunktioon ja muuttujaan $X = f(Y;\theta)/f(Y;\theta_0)$. Voit olettaa, että kyseessä on jatkuva tapaus eli tf.

3. Oletetaan, että havaintoja vastaavat satunnaismuuttujat $Y_1, \ldots, Y_n$ ovat riippumattomia ja samoin jakautuneita (iid-tapaus). Todista, että jos muistiinpanojen sivun 30 alussa mainitut ehdot ovat voimassa, niin lauseen 3.1 oletukset (i) ja (ii) toteutuvat ja siten mallin parametrin su-estimaattori on tarkentuva. Pidetään tunnettuna alla esitetty SLL:n tasainen versio sekä edellisen tehtävän tulos.

   Tasainen SLL. Olkoot $X_1, X_2, \ldots$ 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 $|h(x,\theta)| \le k(x)$ 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 $\mu(\theta) = \mathsf{E}[h(X_1,\theta)]$. Lisäksi funktio μ on jatkuva Θ:ssa.

4. (Hieman erikoisempi su-estimaattorin asymptoottinen jakauma.) Tarkastellaan mallia $Y_1, \ldots, Y_n \sim \mathrm{Tas}(0,\theta) \perp\!\!\!\perp$, jossa $\theta > 0$. Palauta mieleen, että θ:n su-estimaattori on $\hat{\theta}_n = Y_{(n)} = \max(Y_1, \ldots, Y_n)$. 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 lauseen 3.2 tulos 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 $x \in \mathbb{R}$.