**Tilastollisen päättelyn jatkokurssi - yleistentti 15.11.2012**

**1**. Havaintoja vastaavat satunnaismuuttujat ovat $Y_1, ..., Y_n \sim P(\mu) \;\underline{\perp\!\!\!\perp}$. Tarkastellaan otoskeskiarvoa $\bar{Y}_n = n^{-1} \sum_{i=1}^{n} Y_i$ odotusarvon $\mu$ estimaattorina.

(i) Mikä on otoskeskiarvon $\bar{Y}_n$ asymptoottinen jakauma eli mihin muuttuja $\sqrt{n}\left(\bar{Y}_n - \mu\right)$ konvergoi jakaumaltaan?

(ii) Johda muunnoksen $\sqrt{\bar{Y}_n}$ asymptoottinen jakauma eli mihin muuttuja $\sqrt{n}\left(\sqrt{\bar{Y}_n} - \sqrt{\mu}\right)$ konvergoi jakaumaltaan?

**2**. (i) Selvitä yksityiskohtaisesti millainen on tilastollisen mallin tuloesitys.

(ii) Tarkastellaan autoregressiivista aikasarjamallia

$$Y_i = \phi Y_{i-1} + \varepsilon_i, \quad i = 1, ..., n,$$

jossa $Y_0 = y_0$ on tunnettu vakio, $\varepsilon_1, ..., \varepsilon_n \sim \mathsf{N}\left(0, \sigma^2\right) \;\underline{\perp\!\!\!\perp}$ ja $\left(\phi, \sigma^2\right) \in \mathbb{R} \times (0, \infty)$. Johda tilastollinen malli tässä tapauksessa.

**3**. Olkoon $L(\theta; \boldsymbol{y})$ aineistoon $\boldsymbol{y} = (y_1, ..., y_n)$ liittyvä parametrin $\theta \in \Theta \subseteq \mathbb{R}^d$ uskottavuusfunktio. Tarkastellaan lineaarista nollahypoteesia $\mathsf{H}_0 : A\theta_0 = c$, jossa $\theta_0$ on parametrin $\theta$ todellinen arvo, matriisi $A$ $(q \times d)$ ja vektori $c$ $(q \times 1)$ ovat tunnettuja, ja $A$:n aste on $q$.

Johda Waldin testi $\mathsf{H}_0$:lle vaihtoehtoa $A\theta_0 \neq c$ vastaan selvittäen huolellisesti perustellen testisuureen asymptoottinen jakauma, kun nollahypoteesi on voimassa ja tavanomaiset suurimman uskottavuuden estimoinnin asymptoottiset tulokset pätevät.

**4**. Oletetaan, että havaintoja vastaavat satunnaismuuttujat $Y_1, ..., Y_n$ ovat riippumattomat ja

$$Y_i \sim \mathsf{N}\left(\frac{x_i}{1 + \beta x_i}, 1\right),$$

jossa $x_i$, $i = 1, ..., n$, ovat kiinteän selittävän muuttujan arvoja $\beta \in \mathbb{R}$ on tuntematon parametri.

(i) Johda parametrin $\beta$ pistemääräfunktio, havaittu informaatio ja Fisherin informaatio.

(ii) Johda Raon pistemäärätesti hypoteesille $\mathsf{H}_0 : \beta = 0$ vaihtoehtoa $\beta \neq 0$ vastaan, kun testin tavanomaisen asymptoottisen jakauman vaatimat oletukset ovat voimassa.

**Perustele vastauksesi!**

KÄÄNTÖPUOLELLA MM. TEHTÄVISSÄ ESIINTYVIEN JAKAUMIEN TIHEYSFUNKTIOIDEN JA PISTETODENNÄKÖISYYSFUNKTIOIDEN LAUSEKKEET.

**MUISTIN TUEKSI**

**Lause 1.4.** Oletetaan, että sv–jonoille $X_1, X_2, ...$ $(k \times 1)$ ja $Y_1, Y_2, ...$ $(l \times 1)$ pätee $X_n \xrightarrow{d} Z$ ja $Y_n \xrightarrow{p} c$, jossa $c$ on vakiovektori. Tällöin $(X_n, Y_n) \xrightarrow{d} (Z, c)$.

**Pistemäärän ja SU–estimaattorin tavanomaiset asymptoottiset ominaisuudet**. Olkoon $s(\theta; \boldsymbol{y})$ ja $\mathcal{J}(\theta; \boldsymbol{y})$ aineistoon $\boldsymbol{y} = (y_1, ..., y_n)$ liittyvä pistemäärä ja havaittu informaatiomatriisi ja $\theta_0$ parametrin $\theta \in \Theta \subseteq \mathbb{R}^d$ todellinen arvo. "Sopivien säännöllisyysehtojen" voimassa ollessa on olemassa positiivisesti definiitti matriisi $\bar{\mathcal{I}}(\theta_0)$ siten, että pistemäärälle ja SU–estimaattorille $\hat{\theta}$ pätee

$$\frac{1}{\sqrt{n}} s(\theta_0; \mathbf{Y}) \xrightarrow{d} \mathsf{N}\left(0, \bar{\mathcal{I}}(\theta_0)\right)$$

ja

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathsf{N}\left(0, \bar{\mathcal{I}}(\theta_0)^{-1}\right).$$

Lisäksi, $n^{-1}\mathcal{J}(\hat{\theta}; \mathbf{Y}) \xrightarrow{p} \bar{\mathcal{I}}(\theta_0)$.

**Multinormaalijakauman ominaisuuksia**. (i) Jos $\boldsymbol{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ja $\boldsymbol{Y} = \boldsymbol{BX} + \boldsymbol{\nu}$, jossa $\boldsymbol{B}$ on ei-satunnainen astetta $q$ oleva $q \times k$ matriisi ja $\boldsymbol{\nu}$ on ei-satunnainen $q \times 1$ vektori, niin $\boldsymbol{Y} \sim \mathsf{N}_q(\boldsymbol{B\mu} + \boldsymbol{\nu}, \boldsymbol{B\Sigma B}')$.

(ii) Jos $\boldsymbol{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ja $\boldsymbol{\Sigma}$ on positiivisesti definiitti, niin $(\boldsymbol{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2_k$.

**Tehtävissä esiintyvistä jakaumista**

- Jos satunnaismuuttuja $Y \sim P(\mu)$ (Poisson-jakauma), niin $Y$:n pistetodennäköisyysfunktio on $f(y; \mu) = \frac{1}{y!}\mu^y e^{-\mu}$, $y = 0, 1, ...$, $\mu > 0$. Lisäksi $\mathsf{E}(Y) = \mu$ ja $\mathsf{Var}(Y) = \mu$.
- Jos satunnaismuuttuja $Y \sim \mathsf{N}\left(\mu, \sigma^2\right)$ (normaalijakauma), niin $Y$:n tiheysfunktio on $f(y) = \left(2\pi\sigma^2\right)^{-1/2} e^{-(y-\mu)^2/2\sigma^2}$.