Tilastollinen päättely, syksy 2009 - kevät 2010

Harjoitus 4

1. Olkoon $f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)$ tilastollinen malli, jonka parametri θ on yksiulotteinen. Olkoon $\phi = \phi(\theta)$ kääntäen yksikäsitteinen parametrimuunnos, jonka käänteismuunnos on $\theta = \theta(\phi)$. Tarkastellaan uudelleen parametroitua mallia

$$f^*_{\boldsymbol{Y}}(\boldsymbol{y};\phi) = f_{\boldsymbol{Y}}(\boldsymbol{y};\theta(\phi)).$$

Näytä, että sen havaittu informaatio ja Fisherin informaatio saadaan alkuperäisen mallin informaatioista kaavoilla

$$j^*(\hat{\phi};\boldsymbol{y}) = j(\hat{\theta};\boldsymbol{y})[\theta'(\hat{\phi})]^2 \quad \text{ja} \quad i^*(\phi) = i(\theta(\phi))[\theta'(\phi)]^2.$$

Mallin oletetaan täyttävän kaikki tarpeellliset säännöllisyysehdot ja parametrimuunnoksen oletetaan olevan riittävän monta kertaa derivoituva. [*Muista.* $l'(\hat{\theta},\boldsymbol{y}) = 0$ ja $E[l'(\theta,\boldsymbol{Y})] = 0$. Monisteen harjoitustehtävä 2.14]

2. Tarkastellaan toistokoemallia $Y_1,\ldots,Y_n \sim B(\theta) \perp\!\!\!\perp$. Totea, että suurimman uskottavuuden estimaattori $\hat{\theta} = (Y_1 + \cdots + Y_n)/n$ on harhaton: $E_\theta(\hat{\theta}) = \theta$, ja laske sen varianssi $\mathrm{var}(\hat{\theta})$. Palauta mieleen tämän mallin Fisherin informaatio $i(\theta)$. Mikä yhteys sillä on em. varianssiin?

3. Jatkoa edellisen harjoituksen tehtävälle 4. Olkoot $Y_1,\ldots,Y_n \perp\!\!\!\perp$ ja ja $Y_i \sim N(\beta x_i, \sigma_0^2)$, jossa $x_1,\ldots,x_n$ ovat tunnettuja lukuja ja $\sigma_0^2 > 0$ on tunnettu. Osoita, että parametrin β suurimman uskottavuuden estimaattori $\hat{\beta}$ on harhaton ja laske sen varianssi. Mikä yhteys $\hat{\beta}$:n varianssilla on tämän mallin Fisherin informaatioon $i(\beta)$?

4. Oletetaan, että havainnot $Y_1,\ldots,Y_n$ ovat riippumattomia ja noudattavat jakaumaa, jolla on odotusarvo μ ja varianssi σ^2. Tarkastellaan parametrin σ^2 estimointia muotoa cV olevilla estimaattoreilla, kun $c > 0$ on vakio ja $V = \sum_{i=1}^n (Y_i - \bar{Y})^2$. Näytä, että cV on harhaton, jos ja vain jos $c = 1/(n-1)$. [*Ehdotus.* Käytä harjoituksen 2 tehtävän 1 hajotelmaa $\sum_{i=1}^n (y_i - a)^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - a)^2$ valinnalla $a = \mu$. Monisteen harjoitustehtävä 3.3]

5. Hernekasvit luokitellaan niiden tuottamien herneiden mukaan yhtäältä muodon puolesta (pyöreä tai särmikäs) ja toisaalta värin puolesta (keltainen tai vihreä). Näin hernekasvit jakautuvat neljään luokkaan: PK, PV, SK ja SV. Niiden esiintymistodennäköisyydet ovat genetiikan mukaan vastaavasti $\alpha\beta$, $\alpha(1-\beta)$, $(1-\alpha)\beta$ ja $(1-\alpha)(1-\beta)$, jossa $0 < \alpha < 1$ ja $0 < \beta < 1$.

   Parametrien α ja β estimoimiseksi kasvatettiin sata (toisistaan riippumatonta) kasvia, jolloin havaittiin, että ne jakautuivat em. luokkiin seuraavasti:

Luokka:	PK	PV	SK	SV
Havaintoja:	52	21	17	10

   a) Ilmoita näitä havaintoja vastaava uskottavuusfunktio ja log-uskottavuusfunktio sekä suurimman uskottavuuden estimaatti $(\hat{\alpha}, \hat{\beta})$.

   b) Ovatko tarkasteltavan mallin parametrit α ja β ortogonaaliset? [Vrt. monisteen kohta 2.4.6]