Todennäköisyyslaskennan kurssi, 9. harjoitus (28.–30.11.2012)

1. Olkoon $h(X)$ mielivaltainen sm:n X funktio, ja olkoon $m(X) = E(g(X,Y) \mid X)$. Todista, että

$$E[(g(X,Y) - h(X))^2] = E[(g(X,Y) - m(X))^2] + E[(m(X) - h(X))^2].$$

Opastus: kerro binomi auki kaavassa

$$E[(g(X,Y) - h(X))^2] = E\left\{[(g(X,Y) - m(X)) + (m(X) - h(X))]^2\right\},$$

ja käytä ehdollisen odotusarvon ominaisuuksia (lause 8.3 ja kaava (8.7)).

Tuloksessa jälkimmäinen termi on aina ei-negatiivinen ja nolla, jos $h = m$. Tämän takia sm:n $g(X,Y)$ paras ennuste sm:n X funktiona keskineliövirheen mielessä on ehdollinen odotusarvo $m(X) = E[g(X,Y) \mid X]$ (ts. kaava todistaa lauseen 8.2).

2. Todista lauseen 8.4 tulos: $\operatorname{var}(g(X,Y)) = E\operatorname{var}(g(X,Y) \mid X) + \operatorname{var} E(g(X,Y) \mid X)$.

3. Johda hierarkkisen esityksen

$$\begin{aligned} X \mid Y &\sim \mathrm{Poi}(\lambda Y) \\ Y &\sim \mathrm{Gam}(s,s), \end{aligned}$$

($\lambda, s > 0$ vakioita) avulla negatiivisen binomijakauman $\mathrm{NegBin}(r,p)$ odotusarvolle ja varianssille s. 66 annetut kaavat. Opastus: edellisissä harjoitukisissa osoitettiin, että tässä hierarkkisessa mallissa X:n reunajakauma on negatiivinen binomijakauma tietyillä parametreilla. Laske odotusarvo iteroituna odotusarvona ja varianssi lauseen 8.4 avulla sekä ota huomioon kuinka jakauman parametrit r ja p riippuvat vakioista λ ja s.

4. Oletetaan, että

$$\begin{aligned} Y \mid X &\sim N(m(X), (1-\rho^2)\sigma_Y^2), \\ X &\sim N(\mu_X, \sigma_X^2), \end{aligned}$$

jossa $\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y > 0$ ja $-1 < \rho < 1$ ovat vakioita, ja regressiofunktiolle $m(x)$ oletetaan lineaarinen muoto

$$m(x) = E(Y \mid X = x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

Osoita, että yhteistiheys $f_{X,Y}$ voidaan kirjoittaa esimerkissä 8.6 ilmoitettuun kaksiulotteisen normaalijakauman muotoon (ylin kaava s. 111).

Opastus: Matriisi $\mathbf{C}$ voidaan kääntää symbolisesti; kun $\mathbf{A}$ on 2×2-matriisi, eli

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \text{niin} \quad \mathbf{A}^{-1} = \frac{1}{\det\mathbf{A}}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \text{jossa} \quad \det\mathbf{A} = ad - bc.$$

5. Olkoon X sm, $\mathbf{Y} = (Y_1, Y_2)$ sv ja olkoon sv $\mathbf{Z} = (X, Y_1, Y_2)$ sellainen, että

$$\mathrm{Cov}(\mathbf{Z}) = \begin{bmatrix} 3 & 2 & 3 \\ 2 & 5 & 5 \\ 3 & 5 & 9 \end{bmatrix}.$$

Laske (ts. lue annetusta matriisista): a) $\operatorname{var}(X)$, b) $\mathrm{Cov}(\mathbf{Y})$, c) $\operatorname{cov}(X, \mathbf{Y})$, d) satunnaisvektorin (Y_2, X) kovarianssimatriisi eli $\mathrm{Cov}((Y_2, X))$.

6. (Choleskyn hajotelma.) Olkoon $\mathbf{\Sigma}$ kaksiulotteisen satunnaisvektorin $\mathbf{Z} = (X,Y)$ kovarianssimatriisi. Esitä se muodossa $\mathbf{A}\mathbf{A}^T$, jossa $\mathbf{A}$ on 2×2-alakolmiomatriisi, ts. etsi jokin muotoa

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho\,\sigma_X\sigma_Y \\ \rho\,\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix} = \mathbf{A}\mathbf{A}^T, \quad \text{jossa} \quad \mathbf{A} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}$$

oleva matriisin $\mathbf{\Sigma}$ esitys. (Opastus: Kirjoittamalla matriisitulon auki saat kolme riippumatonta yhtälöä. Yhtälöitä on yksi kutakin kovarianssimatriisin alkiota kohti, mutta matriisien symmetrisyys eliminoi yhtälöistä yhden. Etsi yhtälöille ratkaisu.)