Todennäköisyyslaskennan kurssi, 8. harjoitus (18.–22.11.2013)

1. Tutki, ovatko satunnaismuuttujat X ja Y riippumattomia, kun niiden yhteistiheysfunktio on

a) $f(x,y) = 4\,x\,y$, kun $0 < x < 1$ ja $0 < y < 1$ (ja nolla muuten);

b) $f(x,y) = 8\,x\,y$, kun $0 < y < x < 1$ (ja nolla muuten).

2. Olkoot X, Y ja Z satunnaismuuttujia sekä a, b ja c vakioita. Esitä lineaarikombinaation $U = aX+bY+cZ$ varianssi $\operatorname{var} U$ suureiden $\operatorname{var} X$, $\operatorname{var} Y$, $\operatorname{var} Z$, $\operatorname{cov}(X,Y)$, $\operatorname{cov}(X,Z)$ ja $\operatorname{cov}(Y,Z)$ avulla (joiden oletamme olevan äärellisiä). Käytä hyväksi kovarianssioperaattorin $\operatorname{cov}(\cdot,\cdot)$ bilineaarisuutta (esim. esimerkin 7.6 tapaan).

3. Oletetaan, että käytettävissä on aineisto $(x_1,y_1),\dots,(x_n,y_n)$, jossa skalaari x_i on selittävän muuttujan arvo ja skalaari y_i on selitettävän muuttujan arvo mitattuna i:nnessä otosyksikössä. Tällöin muuttujan y arvoja usein selitetään muuttujan x arvon avulla muotoa

$$y = a + b\,x$$

olevan *regressiosuoran* avulla, ts. regressiosuora antaa selittävän muuttujan arvoa x_i vastaavan ennusteen (tai sovitteen) $a + b\,x_i$. Regressiosuoran kertoimet lasketaan aineistosta kaavoilla

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{j=1}^n (x_j - \bar{x})^2}, \qquad a = \bar{y} - b\,\bar{x}.$$

Tässä $\bar{x}$ ja $\bar{y}$ ovat otoskeskiarvot, eli $\bar{x} = \frac{1}{n}\sum x_i$ ja $\bar{y} = \frac{1}{n}\sum y_i$. (Tietenkin oletamme, että nimittäjä $\sum_j (x_j - \bar{x})^2$ ei ole nolla.)

Olkoot X ja Y diskreettejä satunnaismuuttujia siten, että todennäköisyydellä $1/n$ on $(X,Y) = (x_i,y_i)$. Tarkista, että tällöin saamme regressiosuoran yhtälön muodostamalla jakson 7.6 mukaisen parhaan lineaarisen ennusteen. (Regressiosuora motivoidaan tavallisesti aivan toisenlaisella tavalla kuin tässä tehtävässä.)

4. Laske satunnaisvektorin (X,Y) odotusarvovektori ja kovarianssimatriisi, kun vektorilla on tiheysfunktio

$$f(x,y) = 8\,x\,y, \quad \text{kun } 0 < y < x < 1.$$

5. Todista, että (ks. kaava (7.7)),

$$E(\mathbf{AZB} + \mathbf{C}) = \mathbf{A}(E\mathbf{Z})\mathbf{B} + \mathbf{C},$$

jossa $\mathbf{Z}$ on satunnaismatriisi ja $\mathbf{A}$, $\mathbf{B}$ ja $\mathbf{C}$ ovat vakiomatriiseja, joiden dimensiot ovat sellaiset, että lauseke $\mathbf{AZB} + \mathbf{C}$ on määritelty.

Opastus: Näytä, että kaikilla (i,j)

$$[E(\mathbf{AZB} + \mathbf{C})]_{ij} = [\mathbf{A}(E\mathbf{Z})\mathbf{B} + \mathbf{C}]_{ij}\,,$$

jossa alaindekseillä ij merkitään matriisiarvoisen lausekkeen kohdassa (i,j) olevaa alkiota. (Sovella matriisikertolaskun määritelmää, odotusarvon lineaarisuutta sekä satunnaismatriisin odotusarvon määritelmää.)

6. Olkoot satunnaisvektorin $\mathbf{X} = (X_1,X_2)$ komponentit riippumattomia satunnaismuuttujia, joille

$$EX_1 = 1, \quad EX_2 = 2, \quad \operatorname{var} X_1 = 1, \quad \operatorname{var} X_2 = 3.$$

Määritellään satunnaisvektori $\mathbf{Y} = (Y_1,Y_2)$ kaavalla

$$\mathbf{Y} = \begin{bmatrix} 2013 \\ 3 \end{bmatrix} + \begin{bmatrix} -10 & 2 \\ 1 & -2 \end{bmatrix} \mathbf{X}$$

Laske $E\mathbf{X}$ ja $\operatorname{Cov}(\mathbf{X})$ sekä $E\mathbf{Y}$ ja $\operatorname{Cov}(\mathbf{Y})$ (käytä kaavaa (7.9)). Huomaa, että tämä on sama lasku kuin 5. harjoitusten tehtävä 3.