Todennäköisyyslaskennan kurssi, 9. harjoitus (17.–21.11.2014)

1. *(Diskreetti uimarantaongelma.)* Pitkä uimaranta on muodoltaan reaaliakseli.

(a) Rannalla on kolme asiakasta paikoissa $a = 0$, $b = 2$ ja $c = 10$. Jäätelönmyyjä asettuu paikkaan x, jolloin asiakkaiden yhteenlaskettu matka jäätelöä ostamaan on $s(x) = |x - a| + |x - b| + |x - c|$. Mihin myyjän kannattaa asettua, jotta s minimoituu?

Vihje: Voit tarkastella s:ää erikseen tapauksissa (1) $x \le a$; (2) $a \le x \le b$; (3) $b \le x \le c$; (4) $c \le x$, jolloin pääset eroon itseisarvomerkeistä.

(b) Kaverisi ehdottaa, että myyjän kannattaa asettua asiakkaiden sijaintien keskiarvon kohdalle. Laske s siellä ja vertaa a-kohdan ratkaisuun.

(c) Rannalla onkin n asiakasta paikoissa $a_1 < a_2 < \ldots < a_n$. Mihin myyjän nyt kannattaa asettua, jotta minimoituu yhteenlaskettu matka $s(x) = \sum_{i=1}^{n} |x - a_i|$?

Vihje: mieti s:n suuntaa (kasvava vai vähenevä) esim. sellaisessa paikassa x, jonka oikealla puolella on enemmän asiakkaita kuin vasemmalla puolella. Voit selvyyden vuoksi tarkastella erikseen tapauksia “n on parillinen” ja “n on pariton”.

2. Olkoon satunnaismuuttujalla X jokin jatkuva jakauma, jolla on yksikäsitteinen mediaani m, siten että $P(X \le m) = \frac{1}{2}$. Asetamme piste-ennusteen c, ja yritämme minimoida *itseisvirheen* (ei siis neliövirheen) odotusarvon $s(c) = E(|X - c|)$.

(a) Osoita, että s minimoituu kun $c = m$. *Huom. Tehtävä ei ole aivan helppo. Katso ratkaisuvihjeitä tehtävien sivulta 3.*

(b) Olkoon erityisesti X:llä tiheys $f_X(x) = 2x$, kun $0 < x < 1$. Vertaa itseisvirheen odotusarvoa piste-ennusteilla $c = m$ ja $c = EX$.

3. Olkoot satunnaisvektorin $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ 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 $\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ kaavalla

$$\mathbf{Y} = \begin{bmatrix} 2014 \\ 3 \end{bmatrix} + \begin{bmatrix} -5 & 2 \\ 1 & -2 \end{bmatrix} \mathbf{X}$$

Laske $E\mathbf{X}$ ja $\operatorname{Cov}(\mathbf{X})$ sekä $E\mathbf{Y}$ ja $\operatorname{Cov}(\mathbf{Y})$ (käytä kaavaa (7.9)).

4. Satunnaismuuttujat X ja Y ovat riippumattomat ja kumpikin tasajakautuneet välillä $(0, 1)$. Tutkitaan satunnaisvektoria $\mathbf{V} = \begin{bmatrix} X \\ Y \end{bmatrix}$ ja sen lineaarimuunnosta $\mathbf{W} = \mathbf{A}\mathbf{V}$, missä

$$\mathbf{A} = \begin{bmatrix} 1 - a & a \\ a & 1 - a \end{bmatrix}.$$

(a) Laske kovarianssimatriisi $\operatorname{Cov}(\mathbf{V})$. (Voit käyttää monisteen sivulla 64 todettua tietoa, että riippumattomien muuttujien kovarianssi on nolla.)

(b) Olkoon $a = \frac{1}{4}$. Laske $\operatorname{Cov}(\mathbf{W})$. Ovatko $\mathbf{W}$:n komponentit riippumattomat?

5. Jatkoa edelliseen tehtävään.

(a) Olkoon edelleen $a = \frac{1}{4}$. Koska $\mathbf{V}$:llä on tasajakauma, ja kuvaus on lineaarikuvaus, on myös $\mathbf{W}$:llä tasajakauma eräässä alueessa. Tutki kuvallisesti mikä on kyseinen alue. *Vihje:* Lineaarikuvauksessa suorat kuvautuvat suoriksi. Mihin kuvautuvat yksikköneliön kulmapisteet, mihin sivut ja mihin sisäpisteet? Voit myös tutkia asiaa kokeellisesti pisteparvien avulla.

(b) Mitä tapahtuu, jos onkin $a = \frac{1}{2}$? Mihin yksikköneliö kuvautuu ja mikä on $\mathbf{W}$:n kovarianssimatriisi? Onko kuvaus edelleen bijektio? Mikä on $\mathbf{A}$:n determinantti tässä tapauksessa ja mitä se kertoo kuvajoukon pinta-alasta?

6. Olkoot X ja Y riippumattomia satunnaismuuttujia, joista kumpikin noudattaa eksponenttijakaumaa $\text{Exp}(\lambda)$, jossa $\lambda > 0$ on vakio. Määritellään satunnaismuuttujat U ja V kaavoilla

$$U = X + Y, \qquad V = X - Y.$$

(Silmäile esimerkkiä 7.6 ennen kuin lähdet laskemaan tätä laskua.)

a) Johda satunnaismuuttujien U ja V yhteistiheysfunktio.

b) Johda satunnaismuuttujien U ja V reunatiheysfunktiot. (U:n jakauma on tuttu luvusta 5; V:llä on ns. Laplacen jakauma eli kaksitahoinen eksponenttijakauma).

7. Satunnaismuuttujat T ja R ovat riippumattomat, $T \sim U(0, 2\pi)$ ja $R \sim U(0,1)$. Tutkitaan epälineaarista muunnosta (X, Y), missä $X = R\cos T$ ja $Y = R\sin T$.

Johda muuttujien X, Y yhteistiheysfunktio.

Lisätietoa. Esim. tehtävän 7 jakaumia voisi kokeellisesti havainnollistaa Matlabissa pisteparvien avulla seuraavasti. Kuva antaa karkean käsityksen vektorin (X, Y) yhteisjakaumasta, mutta täsmällisen ytf:n johtamiseen tarvitaan tiheyden muunnoskaava. Koodi löytyy kurssisivulta nimellä `pisteparvi.m`.

```
n = 2000;              % montako pistetta
T = rand(n,1)*2*pi;    % arvotaan yksikkovalilta, skaalataan valille (0,2pi)
R = rand(n,1);         % arvotaan yksikkovalilta

figure(1)
plot(T,R,'.');         % piirretaan alkuperaiset pisteet
axis equal             % sama mittakaava vaaka- ja pystysuunnassa
grid on                % viivaruudukko

X = R .* cos(T);       % kertolasku alkioittain operaattorilla .*
Y = R .* sin(T);       % samoin tassa

figure(2)
plot(X,Y,'.');         % piirretaan muunnetut pisteet
axis equal             % sama mittakaava vaaka- ja pystysuunnassa
grid on                % viivaruudukko
```

Ratkaisuvihjeitä tehtävään 2a.

Eräs ratkaisutapa on seuraava. Tutki aluksi, mitä tapahtuu jos piste-estimaatti asetetaan mediaania suuremmaksi, ts. $c = m + h$, missä $h > 0$. (Piirrä kuva!) Pyritään osoittamaan, että tällöin $s(c) \geq s(m)$.

Esitä erotus $s(c) - s(m)$ integraalina ja hajota se *kolmeen* osaan, nimittäin väleille $(-\infty, m)$, (m, c) ja (c, ∞). Tutki integrandia kullakin välillä erikseen ja osoita sille mahdollisimman suuri alaraja (vihje: se on eri väleillä joko $-h$ tai h kerrottuna X:n tiheysfunktiolla). Johda tästä itse integraaleja koskeva alaraja, ja hyödynnä tietoa, että $P(X \leq m) = \frac{1}{2}$ sen osoittamiseksi, että $s(c) \geq s(m)$.

Lopuksi päättele samaan tapaan tapaus $c < m$.