Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos
MATA280 Stokastiikan perusteet

Laskuharjoitus 3 – Ratkaisut
18–19.11.2013
L. Leskelä / M. Kuronen

3.1 Havaitaan ensin, että

$$\mathbb{E}X = \sum_{i=1}^{3} i\, P_X(i) = \sum_{i=1}^{3} i\, \frac{1}{3} = 2$$

ja tietysti myös $\mathbb{E}Y = 2$.

(a) Lineaarisuuden perusteella $\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y = 4$.

(b) Lineaarisuuden perusteella $\mathbb{E}(X-Y) = \mathbb{E}X - \mathbb{E}Y = 0$.

(c) Koska X ja Y ovat riippumattomat, $\mathbb{E}(XY) = \mathbb{E}X\, \mathbb{E}Y = 4$.

(d) Satunnaismuuttujan $1/Y$ odotusarvo on

$$\mathbb{E}(1/Y) = \sum_{i=1}^{3} (1/i)\, P_Y(i) = \sum_{i=1}^{3} \frac{1}{i}\, \frac{1}{3} = \frac{11}{18}.$$

Nyt $X \perp\!\!\!\perp Y \implies X \perp\!\!\!\perp 1/Y$, joten $\mathbb{E}X/Y = \mathbb{E}X\, \mathbb{E}(1/Y) = \frac{11}{9}$.

(e) Merkitään $f(s,t) = s^t$. Tällöin

$$\mathbb{E}X^Y = \mathbb{E}f(X,Y) = \sum_{s=1}^{3} \sum_{t=1}^{3} s^t P_{(X,Y)}(s,t),$$

missä $P_{(X,Y)}$ on satunnaisvektorin (X,Y) yhteisjakauma. Koska $X \perp\!\!\!\perp Y$, pätee

$$P_{(X,Y)}(s,t) = P_X(s) P_Y(t) = \frac{1}{9}$$

kaikilla $(s,t) \in \{1,2,3\}^2$. Näin ollen

$$\mathbb{E}X^Y = \frac{1}{9} \sum_{s=1}^{3} \sum_{t=1}^{3} s^t = \frac{56}{9}.$$

(f) Ensin lasketaan

$$\mathbb{E}\sin(\pi/X) = \frac{\sin(\pi/1) + \sin(\pi/2) + \sin(\pi/3)}{3} = \frac{1}{3}\left(1 + \frac{\sqrt{3}}{2}\right)$$

ja

$$\mathbb{E}\cos(\pi/Y) = \frac{\cos(\pi/1) + \cos(\pi/2) + \cos(\pi/3)}{3} = -\frac{1}{6}.$$

Nyt $X \perp\!\!\!\perp Y \implies \sin(\pi/X) \perp\!\!\!\perp \cos(\pi/Y)$ ja näin ollen

$$\mathbb{E}\sin(\pi/X)\cos(\pi/Y) = \mathbb{E}\sin(\pi/X)\, \mathbb{E}\cos(\pi/Y) = -\frac{1}{18}\left(1 + \frac{\sqrt{3}}{2}\right).$$

3.2 (a) Jos X ja Y ovat riippumattomia satunnaislukuja, joilla on odotusarvot, niin aina pätee $\mathbb{E}XY = \mathbb{E}X\,\mathbb{E}Y$.

(b) Olkoon X jokin satunnaisluku, jolla on odotusarvo $\mathbb{E}X$ ja varianssi $0 < \mathrm{Var}(X) < \infty$. Määritellään $Y = X$. Tällöin varianssin laskukaavaa $\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$ käyttämällä nähdään, että

$$\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathrm{Var}(X).$$

Koska oletuksen mukaan $\mathrm{Var}(X) > 0$, seuraa tästä, että $\mathbb{E}XY > \mathbb{E}X\,\mathbb{E}Y$.

3.3 (a) Määritellään satunnaisvektori $\psi_k : \Omega \to \{0,1\}^4$ kaavalla $\psi_k = (\theta_{4(k-1)+1}, \dots, \theta_{4k})$, missä $k = 1, 2, \dots$ Todistetaan ensiksi, että $\psi_1 \perp\!\!\!\perp \psi_2$. Havaitaan ensiksi, että

$$(\psi_1, \psi_2) = (\theta_1, \theta_2, \dots, \theta_8).$$

Koska kyseisen vektorin komponentit ovat riippumattomia, noudattaa (ψ_1, ψ_2) tasajakaumaa joukossa $\{0,1\}^8$. Näin ollen

$$\mathbb{P}(\psi_1 = s,\, \psi_2 = t) = \left(\frac{1}{2}\right)^8$$

kaikilla $s, t \in \{0,1\}^4$. Koska toisaalta

$$\mathbb{P}(\psi_1 = s) = \left(\frac{1}{2}\right)^4 \quad \text{ja} \quad \mathbb{P}(\psi_2 = t) = \left(\frac{1}{2}\right)^4,$$

voidaan päätellä että

$$\mathbb{P}(\psi_1 = s,\, \psi_2 = t) = \mathbb{P}(\psi_1 = s)\mathbb{P}(\psi_2 = t)$$

kaikilla s ja t. Siispä $\psi_1 \perp\!\!\!\perp \psi_2$. Samaan tapaan voidaan todistaa, että satunnaismuuttujat $\psi_1, \dots, \psi_n$ ovat riippumattomat kaikilla n. Näin ollen myös satunnaismuuttujat $X_1 = \phi(\psi_1), \dots, X_n = \phi(\psi_n)$ ovat riippumattomat kaikilla n. Väite seuraa tästä.

(b) Olkoon $\phi : \{0,1\}^4 \to \{1, \dots, 16\}$ jokin bijektio. Koska $(\theta_1, \dots, \theta_4)$ on tasajakautunut joukossa $\{0,1\}^4$ ja koska ϕ on bijektio, niin myös $X_1 = \phi(\theta_1, \dots, \theta_4)$ on tasajakautunut joukossa $\phi(\{0,1\}^4) = \{1, \dots, 16\}$. Sama pätee myös satunnaismuuttujille X_i, $i \geq 2$.

(c) Määritellään $\phi(s_1, \dots, s_4) = s_1 \cdots s_4$. Tällöin

$$\mathbb{P}(\phi(\theta_1, \dots, \theta_4) = 1) \;=\; \mathbb{P}(\theta_1 = 1, \dots, \theta_4 = 1) \;=\; \frac{1}{16}.$$

Koska ϕ:n arvojoukko on $\{0,1\}$, nähdään että

$$\mathbb{P}(\phi(\theta_1, \dots, \theta_4) = 0) \;=\; 1 - \mathbb{P}(\phi(\theta_1, \dots, \theta_4) = 1) \;=\; \frac{15}{16}.$$

Siispä X_1 on Bernoulli-jakautunut parametrilla $1/16$. Sama pätee myös satunnaismuuttujille X_i, $i \geq 2$.

(d) Vastaus on ei. Näin ei voi käydä, sillä mielivaltaiselle kuvaukselle $\phi : \{0,1\}^4 \to \{1,2,3\}$ pätee

$$\mathbb{P}(\phi(\theta_1,\ldots,\theta_4)=1) = \sum_{s\in\phi^{-1}\{1\}} P_{(\theta_1,\ldots,\theta_4)}(s) = \sum_{s\in\phi^{-1}\{1\}} \frac{1}{16} = \frac{|\phi^{-1}\{1\}|}{16},$$

missä $\phi^{-1}\{1\} = \{s : \phi(s) = 1\}$. Ylläolevan yhtälön oikea puoli ei selvästikään voi koskaan saada arvoa $\frac{1}{3}$.

3.4 (a) Geometrisen sarjan summakaavaa käyttäen saadaan

$$\mathbb{P}(X \geq k) = \sum_{j=k}^{\infty} \mu(j) = \sum_{j=k}^{\infty} (1-p)^{j-1} p = p\frac{(1-p)^{k-1}}{1-(1-p)} = (1-p)^{k-1}.$$

Vastaavasti $\mathbb{P}(Y \geq k) = (1-q)^{k-1}$.

(b) Havaitaan ensiksi, että $Z \geq k \iff X \geq k$ ja $Y \geq k$. Koska lisäksi $X \perp\!\!\!\perp Y$, niin

$$\begin{aligned}
\mathbb{P}(Z \geq k) &= \mathbb{P}(X \geq k, Y \geq k) \\
&= \mathbb{P}(X \geq k)\mathbb{P}(Y \geq k) \\
&= (1-p)^{k-1}(1-q)^{k-1} \\
&= ((1-p)(1-q))^{k-1} \\
&= (1-(p+q-pq))^{k-1}.
\end{aligned}$$

Siispä $\mathbb{P}(Z \geq k) = (1-r)^{k-1}$, missä $r = p+q-pq$.

(c) Koska

$$\begin{aligned}
\mathbb{P}(Z = k) &= \mathbb{P}(Z \geq k) - \mathbb{P}(Z \geq k+1) \\
&= (1-r)^{k-1} - (1-r)^k \\
&= (1-r)^{k-1}(1-(1-r)) \\
&= (1-r)^{k-1} r,
\end{aligned}$$

niin Z noudattaa geometrista jakaumaa.

3.5 (a) Olkoot $n = 10^8$ ja $\xi_1, \ldots, \xi_n$ apinan lyömät kirjaimet. Olkoon

$$A_i = \{\xi_i = K,\ \xi_{i+1} = I,\ \xi_{i+2} = V,\ \xi_{i+3} = I\}$$

tapahtuma, että sana kivi esiintyy tekstissä i:nnestä kirjaimesta alkaen. Koska kirjaimet ovat riippumattomat, pätee

$$\mathbb{P}(A_i) = \left(\frac{1}{50}\right)^4.$$

Sanan 'KIVI' esiintymiskertojen lukumäärä tekstissä on

$$\sum_{i=1}^{n-3} 1_{A_i}$$

ja sen odotusarvo on

$$\mathbb{E}\sum_{i=1}^{n-3} 1_{A_i} = \sum_{i=1}^{n-3} \mathbb{P}(A_i) = (n-3)\frac{1}{50^4} \approx 16.$$

(b) Olkoot $\xi_1, \ldots \xi_n$ apinan lyömät kirjaimet ja olkoot $z_1, \ldots, z_k$ *Seitsemän veljestä* -teoksen kirjaimet, missä $k = 635864$. Olkoon

$$A_i = \{\xi_i = z_1,\, \xi_{i+1} = z_2, \ldots,\, \xi_{i+k-1} = z_k\}$$

tapahtuma, että *Seitsemän veljestä* esiintyy apinan tekstissä i:nnestä kirjaimesta alkaen. Tällöin *Seitsemän veljestä* esiintymien lukumäärä apinan tuottamassa n:n merkin tekstissä on

$$\sum_{i=1}^{n-(k-1)} 1_{A_i}$$

ja sen odotusarvo on

$$\mathbb{E} \sum_{i=1}^{n-(k-1)} 1_{A_i} = \sum_{i=1}^{n-(k-1)} \mathbb{P}(A_i) = (n-k+1)\left(\frac{1}{50}\right)^k.$$

Jotta esiintymisiä olisi keskimäärin yksi, täytyy apinan kirjoittaa teksti, jonka pituus on $n = 50^k + k - 1 \geq 10^{600000}$. Vertailun vuoksi havaittavissa olevassa maailmankaikkeudessa arvioidaan olevan noin 10^{80} atomia.