HY, Todennäköisyysteoria I tentti, ratkaisut 11 maaliskuuta 2015

Valitse 4 tehtävää listasta 1,2,3,4,5,6 ja vastaa tehtävien kysymyksiin. Toki voit myös ratkaistaa kaikki tehtävät jos sinulle jää siihen tarpeeksi aikaa. Kokeen kesto on 3 tuntia.

1. (a) Esitä σ-algebran määritelmä.
   R. luentomoniste.

   (b) Esitä Borelin σ-algebran määritelmä.
   R. luentomoniste.

   (c) Esitä tulo σ-algebran määritelmä.
   R. luentomoniste.

   (d) Todista väite

$$\mathcal{B}(\mathbb{R}^d) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{d\text{-kertaa}}$$

jossa $\mathcal{B}(\mathbb{R}^d)$ on $\mathbb{R}^d$ avaruuden Borel σ-algebra ja oikealla puolella on d-kertainen tulo Borelin σ-algebroista.

R. (harjoitus 2.6)

$\mathbb{R}^d$ on separoituva metrinen avaruus jossa $\mathbb{Q}^d$ on numeroituva ja tiheä $\mathbb{R}^d$:ssa.

Määritelmän mukaan avaruuden $\Omega = \mathbb{R}^d$ Borelin σ algebra on avomien joukkojen virittämä σ-algebra.

Kun $t \in \mathbb{R}^d$, olkoon

$$(-\infty, t] = \{s \in \mathbb{R}^d : \ s_i \le t_i, i = 1, \dots d\}$$

Todistan ensin että luokka

$$\mathcal{I} = \big\{(-\infty, q), q \in \mathbb{Q}^d\big\}$$

on π-luokka joka virittää Borelin σ-algebran $\mathcal{B}(\mathbb{R}^d)$. Siitä seuraa

$$\mathcal{B}(\mathbb{R}^d)\sigma(\mathcal{I}) \subseteq \underbrace{\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{d\text{-kertaa}}$$

Jos $U \subseteq \mathbb{R}^d$ on avoin joukko, $\forall x \in U$ on olemassa pallo

$$B(q_x, r_x) := \big\{y : |y - q_x| < r_x\big\},$$

jolla $x \in B(q_x, r_x) \subseteq U$, $q_x \in \mathbb{Q}^d \cap U$, $r_x \in \mathbb{Q}^+$.
Koska $\mathbb{Q}^d$ on numeroituva, tästä seuraa että avoin joukko on numeroituva yhdiste avoimista palloista

$$U = \bigcup_{n \in \mathbb{N}} B(q_n, r_n)$$

jossa $q_n \in \mathbb{Q}^d \cap U$ ja $r \in \mathbb{Q}^+$.
Huomataan että voidaan myös korvata palloja d-ulotteisilla kuutioilla: Jos U on avoin $\mathbb{R}^d$:ssa, koska $\mathbb{Q}$ on tiheä $\mathbb{R}$:ssa, $\forall x \in U$ $\exists r, q \in \mathbb{Q}^d$ jolla $r < q$ (eli $r_i < q_i$ kun $i = 1, \dots, d$

$$x \in (r, q) := (r_1, q_1) \times (r_2, q_2) \times \cdots \times (r_d, q_d) \subseteq U.$$

Selvästi $\mathcal{B}(\mathbb{R}^d) \supseteq \sigma(\mathcal{I})$, koska $(-\infty, q) \in \mathcal{B}(\mathbb{R}^d)$.
Kun $\mathbf{q}, \mathbf{r} \in \mathbb{Q}^\mathbf{d}$, $(-\infty, q) \cap (-\infty, r) = (-\infty, r \wedge q)$ jossa $r \wedge q \in \mathbb{Q}^d$, $(r \wedge q)_i = \min(r_i, q_i)$, $i = 1, \dots d$, eli $\mathcal{I}$ on π-luokka.
Huomataan että avoin laatikko kulmapisteilla $r \leq q \in \mathbb{Q}^d$, (eli $r_i \leq q_i$ kun $i = 1, \dots, d$) on

$$L(r, q) := \{s \in \mathbb{R}^d : r < s < q\} \in \sigma(\mathcal{I}).$$

Esimerkiksi kun $d = 2$, laatikon indikaattori funktiolla on esitys

$$\begin{aligned} \mathbf{1}\{t \in L((r_1, r_1), (q_1, q_2)(\} &= \mathbf{1}\{t \in (-\infty, (q_1, q_2))\} + \mathbf{1}\{t \in (-\infty, (r_1, r_2))\} \\ &\quad -\mathbf{1}\{t \in (-\infty, (r_1, q_2))\} - \mathbf{1}\{t \in (-\infty, (q_1, r_2))\} \end{aligned}$$

joka yleistyy korkeammalle dimensioille.
Olkoon $U \subseteq \mathbb{R}^d$ avoin joukko. Koska U on avoin, kaikille $x \in U$ on olemassa avoin laatikko $L(r(x), q(x))$ jolla $x \in L(r(x), q(x)) \subseteq U$ kulmapisteillä $r(x) \leq q(x) \in \mathbb{Q}^d$. Tästä seuraa että

$$U = \bigcup_{x \in U} L(r(x), q(x)) \in \sigma(\mathcal{I})$$

koska yhdiste on numeroituva (vaikka avoin joukko U on ei-numeroituva): rationaalisten kulmapisteiden määrä on korkeintaan numeroituva.
Koska määritelän mukaan avoimet joukot virittävät Borelin σ-algebran , myös $\sigma(\mathcal{I}) \supseteq \mathcal{B}(\mathbb{R}^d)$.
Huomataan myös että kun $A_i \in \mathcal{B}(\mathbb{R})$, $i = 1, \dots, d$,

$$A_1 \times A_2 \times \cdots \times A_d = B_1 \cap B_2 \cap \cdots \cap B_d$$

jossa

$$B_i = (\underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{i-1\text{-kertaa}} \times A_i \times \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{d-i\text{-kertaa}}) \in \mathcal{B}(\mathbb{R}^d),$$

josta seuraa

$$\mathcal{B}(\mathbb{R}^d) \supseteq \underbrace{\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{d\text{-kertaa}}$$

2. n henkilöä käyvät teatterissa ja ennen esitystä jokainen jäättää hattunsa naulakkoon. Esityksen jälkeen jokainen saa takaisin naulakosta satunnaisesti poimitun hatun, ja kaikilla hattujen permutaatioilla on sama todennäköisyys.

   Olkoon X_k satunnaismuuttujat jolla

   $$X_k(\omega) = \begin{cases} 1 & \text{jos } k\text{-henkilö saa takaisin oman hattunsa} \\ 0 & \text{muuten} \end{cases}$$

   $$S_n(\omega) = \sum_{k=1}^{n} X_k(\omega) = \#\{ \text{ henkilöt jotka saavat takaisin oman hattunsa } \}$$

   (a) Osoita : kun $2 \leq k \leq n$, $z \in \{0,1\}$

   $$P_n(X_1 = z, X_2 = z, \dots X_k = z) \geq P_n(X_1 = z)P_n(X_2 = z) \dots P_n(X_k = z)$$

   R. Kun $z = 1$

   $$P_n(X_1 = 1, X_2 = 1, \dots, X_k = 1) = P_n(X_1 = 1)P_{n-1}(X_1 = 1, \dots, X_{k-1} = 1)$$
   $$= \cdots = P_n(X_1 = 1)P_{n-1}(X_1 = 1) \dots P_{n-k+1}(X_1 = 1) \geq P_n(X_1 = 1)^k$$

   koska kun $m \leq n$, $P_m(X_1 = 1) = m^{-1} \geq P_n(X_1 = 1) = n^{-1}$.
   Kun $z = 0$, $n = 2$

   $$P_2(X_1 = 0, X_2 = 0) = \frac{1}{2} \geq P_2(X_1 = 0)^2 = \frac{1}{4}$$

   Kun $z = 0$,

   $$P_n(X_1 = 0, X_2 = 0, \dots, X_k = 0)$$
   $$= P_n(X_1 = 0)P_n(X_2 = 0 | X_1 = 0) \cdots P_n(X_k = 0 | X_1 = 0, X_2 = 0 \dots, X_{k-1} = 0)$$

Kun $k=2$, epäyhtälö seuraa suoralla laskulla:

$P_n(X_2=0|X_1=0)=$
$P_n(X_2=0|X_1=0,B_{1,2})P_n(B_{1,2}|X_1=0)+P_n(X_2=0|X_1=0,B_{1,2}^c)P_n(B_{1,2}^c|X_1=0)$
$=1\times\dfrac{1}{n-1}+\dfrac{(n-2)}{(n-1)}\dfrac{(n-2)}{(n-1)}=\dfrac{1}{n-1}+\dfrac{(n-2)^2}{(n-1)^2}=1-\dfrac{(n-2)}{(n-1)^2}>P_n(X_2=0)=$
$\dfrac{n-1}{n}=1-\dfrac{1}{n}$

Yleisemmin, huomataan myös että epäyhtälöt

$P_n(X_1=0,X_2=0,\dots,X_k=0)>P_n(X_1=0,X_2=0,\dots X_{k-1}=0)P_n(X_k=0)$
$P_n(X_1=0,X_2=0,\dots,X_{k-1}=0|X_k=0)>P_n(X_1=0,X_2=0,\dots X_{k-1}=0)$
$P_n(X_k=0|X_1=0,X_2=0,\dots,X_{k-1}=0)>P_n(X_k=0)$

ovat yhtäpitäviä kun ovat voimassa.
Merkitään tapahtuma

$$B_{\ell,k}=\Big\{\ \ell\text{-henkilö on poiminut naulakosta } k\text{-henkilön hatun}\ \Big\}$$

Huomataan että $\forall \ell\neq\ell', k\neq k'$,

$$B_{\ell,k}\cap B_{\ell',k}=\emptyset,\quad B_{\ell,k}\cap B_{\ell,k'}=\emptyset,\quad X_k(\omega)=\mathbf{1}_{B_{k,k}}(\omega)$$

ja

$$\Omega=\Big(\bigcap_{\ell=1}^{k-1}B_{\ell,k}^c\Big)\cup\Big(\bigcup_{\ell=1}^{k-1}B_{\ell k}\Big)$$

$$\begin{aligned}
&P_n(X_k=1|X_1=0,\dots,X_{k-1}=0)\\
&=P_n\big(X_k=1\big|\bigcap_{\ell=1}^{k-1}B_{\ell,k}^c\big)P_n\big(\bigcap_{\ell=1}^{k-1}B_{\ell,k}^c\big|X_1=0,\dots,X_{k-1}=0\big)\\
&=\frac{1}{n-k+1}P_n\big(\bigcap_{\ell=1}^{k-1}B_{\ell,k}^c\big)\frac{P_n\big(\big|X_1=0,\dots,X_{k-1}=0\big|\bigcap_{\ell=1}^{k-1}B_{\ell,k}^c\big)}{P_n(X_1=0,\dots X_{k-1}=0)}\\
&=\frac{1}{n-k+1}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdots\frac{n-k+1}{n-k+2}\frac{P_{n-1}\big(X_1=0,\dots,X_{k-1}=0\big)}{P_n(X_1=0,\dots X_{k-1}=0)}\\
&=\frac{1}{n}\frac{P_{n-1}\big(X_1=0,\dots,X_{k-1}=0\big)}{P_n(X_1=0,\dots X_{k-1}=0)}\leq\frac{1}{n}=P_n(X_k=1)
\end{aligned}$$

jossa

$$P_{n-1}\big(X_1=0,\dots,X_{k-1}=0\big) \le P_n(X_1=0,\dots X_{k-1}=0)$$

ja on yhtä pitävä

$$P_n(X_k=0|X_1=0,\dots,X_{k-1}=0) \ge P_n(X_k=0) = \frac{n-1}{n}$$

(b) Ovatko satunnaismuuttujat $(X_k : 1 \le k \le n)$ P-riippumattomia? ovatko samoin jakautuneita?

R. Koska epäyhtälö on aito, satunnaismuuttujat eivät ole riippumattomia. Ovat kyllä samoin jakautuneita. Tilanne on symmetrinen koska oletetusti tasainen jakauma permutaatiorhymässä on permutaatio invariantti.

(c) Laske odotusarvot $E_{P_n}(X_k)$, $E_{P_n}(X_k X_\ell)$ kun $1 \le k \ne \ell \le n$. **R.**

$$\begin{aligned} E_{P_n}(X_1) &= E_{P_n}(X_k) = P_n(X_1=1) = \frac{1}{n}, \\ E_{P_n}(X_1X_2) &= P(X_1=1, X_2=1) = \frac{1}{n(n-1)} \end{aligned}$$

(d) Laske odotusarvot $E_{P_n}(S_n)$ eli keksimäärin montako ihmista saa takaisin juuri omaa hattunsa eikä toisen henkilön hatun? **R.**

$$E_{P_n}(S_n) = E_{P_n}\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n E_{P_n}(X_k) = nE_{P_n}(X_1) = nn^{-1} = 1$$

(e) Laske myös $E_{P_n}(S_n^2)$

$$E_{P_n}(S_n^2) = E_{P_n}\left(\left\{\sum_{k=1}^n X_k\right\}^2\right) =$$

$$\sum_{k=1}^n E_{P_n}(X_k) + 2\sum_{1\le k<\ell\le n} E_{P_n}(X_kX_\ell) = nE_{P_n}(X_1) + 2\binom{n}{2}E_{P_n}(X_1X_2)$$

$$= 1 + 2\frac{n(n-1)}{2}\frac{1}{n(n-1)} = 1+1=2$$

(f) Osoita

$$\lim_{n\to\infty} P_n(S_n=0) = \lim_{n\to\infty} P_n\big(X_k=0, \forall 1\le k\le n\big) =$$
$$\lim_{n\to\infty} P_n\big(\{\text{ ei kukaan saa omaa hattunsa takaisin }\}\big) \ge \exp(-1)$$

(Itse asiassa tämä epäyhtälö on aito yhtälö). **R.**

$$P_n(S_n = 0) \geq P_n(X_1 = 0)^n = \left(\frac{n-1}{n}\right)^n = \left(1 - \frac{1}{n}\right)^n \to \exp(-1) \quad \text{kun } n \to \infty$$

3. Olkoon satunnaismuuttujat $(U_k(\omega) : k \in \mathbb{N})$ P-riippumattomia ja tasaisesti jakautuneita välissä $[0,1]$, eli

$$P(U_1 \in (a_1, b_1], \ldots, U_n \in (a_n, b_n]) = \prod_{k=1}^{n} (b_k - a_k)$$

$\forall n \in \mathbb{N}$, $0 \leq a_k \leq b_k \leq 1$.

Olkoon

$$X_n(\omega) = \max\{U_1(\omega), \ldots, U_n(\omega)\}, \text{ ja } Y_n(\omega) = \min\{U_1(\omega), \ldots, U_n(\omega)\}$$

(a) Osoita etta $P(Y_n \leq t) = P(X_n \geq 1 - t)$ kun $t \in (0,1)$.
Vihje Osoita ensin että $V_n(\omega) = 1 - U_n(\omega)$ on myös tasaisesti jakautunut välissä $[0,1]$.

(b) Laske X_n ja Y_n kertymäfunktiot ja tiheysfunktiot.

(c) Laske $E(X_n^k)$ ja $E(Y_n^k)$ jossa $k \in \mathbb{N}$.

(d) Osoita

$$\lim_{n\to\infty} X_n(\omega) = 1, \quad \text{ja } \lim_{n\to\infty} Y_n(\omega) = 0, \quad P\text{-melkein varmasti .}$$

Vihje Muista Borel Cantellin lemmat.

R. (harjoitus 4.3) Riippumattomuuden nojalla,

$$P(X_n \leq t) = P(U_1 \leq t, U_2 \leq t, \ldots, U_n \leq t) = P(U_1 \leq t)P(U_2 \leq t) \ldots P(U_n \leq t)$$
$$= P(U_1 \leq t)^n = t^n$$

X_n jakauma on absoluuttisesti jatkuva koska kertymäfunktiolla on derivaatta

$$p_{X_n}(t) = \frac{d}{dt} P(X_n \leq t) \frac{d}{dt} t^n = n t^{n-1}$$

joka on jakuman tiheysfunktio Lebesgue mitan suhteen.

$$E(X_n^k) = n \int_0^1 t^{k+n-1} dt = \frac{n}{n+k}$$

Huomataan että $\forall \omega$, $X_n(\omega) \le X_{n+1}(\omega) \le 1$ ja siksi $\forall \omega$ on olemassa monotoninen raja $X_\infty(\omega) = \lim_{n\to\infty} X_n(\omega) \le 1$.
Koska $\forall t \in [0,1)$

$$\sum_{n=1}^{\infty} P(X_n \le t) = \sum_{n=1}^{\infty} t^n = \frac{t}{1-t} < \infty$$

siitä seuraa että $P(\limsup_n \{ X_n \le t\}) = 0 \ \forall t < 1$, ja siksi $\lim_{n\uparrow\infty} X_n(\omega) = 1$ P-melkein varmasti.
Olkoon $V(\omega) = 1 - U(\omega)$.

$$P(V \le t) = P(1-U \le t) = P(U \ge 1-t) = 1 - P(U < 1-t) = 1-(1-t) = t$$

$Y_n(\omega)$ ja $Y\prime(\omega)_n := \big(1 - X_n(\omega)\big)$ ovat samoin jakautuneita:

$$P(Y_n \le t) = P\big(\min\{U_1,\dots,U_n\} \le t\big) = P\big(\min\{U_1,\dots,U_n\} \le t\big) = P\big(1 - \max\{V_1,\dots$$

tai suoraan

$$P(Y_n \le t) = 1 - P(Y_n > t) = 1 - P(Y_1 > t)^n = 1 - (1-t)^n$$

Y_n tiheys funktio on

$$p_{Y_n}(t) = \frac{dP(Y_n \le t)}{dt} = n(1-t)^{n-1} = p_{X_n}(1-t)$$

$$E(Y_n^k) = n\int_0^1 (1-t)^{n-1} t^k dt = n\frac{\Gamma(n)\Gamma(k+1)}{\Gamma(n+k+1)} = \frac{n(n-1)!k!}{(n+k)!} = \frac{n!k!}{(n+k)!} = \binom{n+k}{n}^{-}$$

jossa laskuissa esintyy Beta integraali.
Koska $Y_n(\omega) = 1 - \max\{V_1(\omega), \dots V_n(\omega)\}$ jossa $V_i(\omega) = 1 - U_i(\omega)$ ovat riippumattomia ja tasaisesti jakautuneita välissä $[0,1]$, seuraa että P-melkein varmasti

$$\lim_{n\to\infty} Y_n(\omega) = 1 - \lim_{n\to\infty}(1 - Y_n(\omega)) = 1 - 1 = 0$$

4. Olkoon $X(\omega)$ Gaussinen satunnaismuuttuja, odotusarvolla $E(X) = 0$ ja varianssilla $E(X^2) = \sigma^2$, kertymäfunktiolla

$$F_X(t) = P(X \le t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{t} \exp\big(-\frac{x^2}{2\sigma^2}\big) dx$$

Olkoon $Y(\omega) = X(\omega)^2$. Laske Y:n jakauman kertymäfuktio $F_Y(t) = P(Y \le t)$ ja tiheysfunktio.

R. Kun $t \ge 0$,

$$\begin{aligned}
&F_Y(t) = P(Y \le t) = P(X^2 \le t) = P(-\sqrt{t} \le X \le \sqrt{t}) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\sqrt{t}}^{\sqrt{t}} \exp\bigl(-\frac{x^2}{2\sigma^2}\bigr)dx = \sqrt{\frac{2}{\pi\sigma^2}} \int_0^{\sqrt{t}} \exp\bigl(-\frac{x^2}{2\sigma^2}\bigr)dx \\
&= \sqrt{\frac{2}{\pi\sigma^2}} \int_0^{t} \exp\bigl(-\frac{y}{2\sigma^2}\bigr)\left|\frac{dx(y)}{dy}\right|dy = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^{t} \exp\bigl(-\frac{y}{2\sigma^2}\bigr)y^{-1/2}dy
\end{aligned}$$

muuttujan vaihdolla $y = x^2, x = \sqrt{y}$, ja Y:n jakauman tiheysfunktio on

$$p_Y(t) = \frac{dP(Y \le t)}{dt} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\bigl(-\frac{y}{2\sigma^2}\bigr)y^{-1/2}$$

5. Todennäköisyysavaruudella $(\Omega, \mathcal{F}, \mathbb{P})$, olkoon $(X_n(\omega) : n \in \mathbb{N})$ P-riippumattomia ja samoin jakautuneita satunnaismuuttujia, joilla

$$P(X_1 \ge 0) = 1,\ E_{\mathbb{P}}(X_1) < \infty,\ \text{ ja } E_{\mathbb{P}}\bigl(\frac{1}{X_1}\bigr) < \infty.$$

Olkoon $S_n(\omega) = X_1(\omega) + \cdots + X_n(\omega)$, $n \in \mathbb{N}$.

Laske

$$E_{\mathbb{P}}\left(\frac{S_m}{S_n}\right)$$

jossa $m \le n$.

Vihje

$$1 = \frac{X_1 + \cdots + X_n}{S_n}$$

R.

$$1 = E_P\bigl(\frac{S_n}{S_n}\bigr) = \sum_{k=1}^{n} E_P\bigl(\frac{X_k}{S_n}\bigr) = nE_P\bigl(\frac{X_1}{S_n}\bigr) \Longleftrightarrow E_P\bigl(\frac{X_1}{S_n}\bigr) = \frac{1}{n}$$

ja

$$1 = E_P\bigl(\frac{S_m}{S_n}\bigr) = \sum_{k=1}^{m} E_P\bigl(\frac{X_k}{S_n}\bigr) = mE_P\bigl(\frac{X_1}{S_n}\bigr) = \frac{m}{n}$$

6. Olkoon $X(\omega)$ satunnaismuuttuja todennäköisyysavaruudella $(\Omega, \mathcal{F}, \mathbb{P})$. Osoita että seuraavat väitteet ovat yhtäpitäviä

   (a) $E_{\mathbb{P}}(|X|) < \infty$

   (b) $\forall c > 0$, $\sum\limits_{n \in \mathbb{N}} \mathbb{P}\big(|X| > cn\big) < \infty$

   (c) $\exists c > 0$: $\sum\limits_{n \in \mathbb{N}} \mathbb{P}\big(|X| > cn\big) < \infty$

   R. Fubini lauseesta, kun $c > 0$

$$c^{-1} E_P(|X|) = E_P(c^{-1}|X|) = \int_0^\infty P(|X| > ct)dt$$

   seuraa yksinkertaisesti odotusarvon lineaarisuudesta jos on olemassa $c > 0$ jolla $\int\limits_0^\infty P(|X| > ct)dt < \infty$, sitä seuraa $\int\limits_0^\infty P(|X| > Kt)dt < \infty$ $\forall K > 0$.

   Väite seuraa koska $\forall t > 0$,

$$\begin{aligned} 0 &\leq \sum_{n=1}^\infty P(|X| > cn)\mathbf{1}((n-1)c \leq t < nc) \leq P(|X| > ct) \\ &\leq \sum_{n=0}^\infty P(|X| > cn)\mathbf{1}(nc \leq t < (n+1)c) \end{aligned}$$

   ja integroimaalla Lebesgue mitan suhteen saadaan

$$0 \leq \sum_{n=1}^\infty P(|X| > cn) \leq \int_0^\infty P(|X| > ct)dt \leq \sum_{n=0}^\infty P(|X| > cn)$$