HY Todennäkösyysteoria III, harjoitustehtävät 5, kevät 2014, Ratkaisut 12.2.2014

1. Olkoon $\{X_i : i \in \mathbb{N}\}$ i.i.d. satunnausmuuttujat kertymäfunktiolla $F(t) = P(X_1 \leq t)$.

   Otoksen $(X_1, \ldots, X_n)$ Empiirinen jakauma on satunnaisprosessi

   $$F_n(t, \omega) := \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(X_i(\omega) \leq t) \qquad n \in \mathbb{N},\ t \in \mathbb{R}$$

   (a) Osoita: $E_P(F_n(t)) = F(t)$.

   R: suoraan odotusarvon lineaarisuudesta.

   (b) : osoita $\forall t \in \mathbb{R}$,

   $$F_n(t, \omega) \to F(t)$$

   P m.v. kun $n \to \infty$.

   R: Koska $E_P(\mathbf{1}(X_1 \leq t)) = F(t)$ väite seuraa Kolmogorovin suurten lukujen laista.

   (c) Todista Glivenko-Cantelli lemma

   $$\sup_{t \in \mathbb{R}} |F_n(t, \omega) - F(t)| \to 0$$

   P m.v. kun $n \to \infty$.

   Vihje osoita ensin että

   $$\sup_{t \in \mathbb{R}} |F_n(t, \omega) - F(t)| = \sup_{q \in \mathbb{Q}} |F_n(t, \omega) - F(t)|$$

   F_n ja F ovat oikelta jatkuvia.

   Huomataan että kun $a \leq t \leq b$,

   $$F_n(a) - F(a) + F(a) - F(b) \leq F_n(t) - F(t) \leq F_n(b) - F(b) + F(b) - F(a)$$

   josta seuraa

   $$|F_n(t) - F(t)| \leq |F_n(b) - F(b)| \wedge |F_n(a) - F(a)| + |F(b) - F(a)|$$

R. Koska $F_n(t)$ ja $F(t)$ ovat oikealta jatkuvia,

$$\sup_{t\in\mathbb{R}} |F_n(t,\omega) - F(t)| = \sup_{q\in\mathbb{Q}} |F_n(q,\omega) - F(q)|$$

joka on $\mathcal{F}$-mitallinen, eli se on satunnaismuuttuja.

Kun $\varepsilon > 0$ löytyy pisteitä

$$-\infty = t_0 < t_1 < \cdots < t_m < t_{m+1} = +\infty,\ m \in \mathbb{N}$$

jolla

$$F(t_{i+1}) - F(t_i) \le \varepsilon \text{ ja } F(t_{i+1}-) - F(t_i-) < \varepsilon, \quad \forall i = 0, 1, \ldots, m.$$

Koska $\forall t \in \mathbb{R}$ $F_n(t,\omega) \to F(t)$ P-melkein varmasti, se pätee myös tasaisesti äärellisissä joukoissa, eli P-melkein varmasti, $\forall\varepsilon$ $\exists n_\varepsilon(\omega)$ jolla $\forall n > n_\varepsilon(\omega)$,

$$\sup_{1\le i\le m} |F_n(t_i,\omega) - F(t_i)| < \varepsilon$$

Koska $t \mapsto F_n(t,\omega)$ on ei-vaheneva

$\forall\omega \in \Omega,\ i \in 0,\ldots,n, t_i < t < t_{i+1}$ P-melkein varmasti $\forall n > n_\varepsilon(\omega)$

$$-2\varepsilon \le F_n(t_i,\omega) - F(t_i) - \varepsilon \le F_n(t,\omega) - F(t) \le F_n(t_{i+1},\omega) - F(t_{i+1}) + \varepsilon \le 2\varepsilon$$

ja väite seuraa.

(d) Laske kovarianssi

$$\mathrm{Cov}(F_n(t), F_m(s)) = E_P\big\{\big(F_n(t,\omega) - F(t)\big)\big(F_m(s,\omega) - F(s)\big)\big\}$$

for $n, m \in \mathbb{N}$, $t, s \in \mathbb{R}$. **R .**

$$E_P\big(\mathbf{1}(X_i \le t)\mathbf{1}(X_i \le s)\big) = E_P\big(\mathbf{1}(X_i \le t \wedge s)\big) = F(t \wedge s)$$

ja

$$\mathrm{Cov}_P(\mathbf{1}(X_i \le t), \mathbf{1}(X_i \le s)) = F(t \wedge s) - F(t)F(s)$$

Seuraa kun $s \le t$

$$\mathrm{Cov}(F_n(t), F_m(s)) = \frac{n \wedge m}{nm} F(s)(1 - F(t))$$

(e) Osoita että kun $s, t \in R$

$$\left(\sqrt{n}(F_n(s) - F(s)), \sqrt{n}(F_n(t) - F(t))\right) \xrightarrow{d} (G(s), G(t))$$

jossa satunnais-pari $(G(s), G(t))$ on gaussinen. Mitkä ovat $E(G(t))$, $E(G(s)G(t))$?

R.

Koska $\mathbf{1}(X_i \leq t) - F(t)$:lla on odotusarvo 0 ja varianssi $F(t)(1 - F(t))$, ja (X_i) on i.i.d. jono seuraa keskeisen rajaarvon lauseesta että $\sqrt{n}(F_n(t) - F(t)) \xrightarrow{d} X(t)$ (jakauman mielessä, ei missään nimessä stokastisesti) jossa $X(t)$ on gaussinen jolla $E_P(X(t)) = 0$ ja $E_P(X(t)^2) = F(t)(1 - F(t))$.

Koska satunnaisparilla $\left(\mathbf{1}(X_i \leq t) - F(t), \mathbf{1}(X_i \leq t) - F(t)\right)$ on nolla odotusarvo kovarianssi $F(s \wedge t) - F(s)F(t)$, seuraa moniulotteisen keskeisen raja-arvo lauseesta että

$$n^{-1/2} \sum_{i=1}^{n} \left(\mathbf{1}(X_i \leq t) - F(t), \mathbf{1}(X_i \leq t) - F(t)\right) \xrightarrow{d} (G(s), G(t))$$

jossa satunnaispari $(G(s), G(t))$ on Gaussinen, nolla odotusarvolla ja kovarianssi matriisilla

$$\begin{pmatrix} F(s)(1 - F(s)) & F(t \wedge s)(1 - F(t \vee s)) \\ F(t \wedge s)(1 - F(t \vee s)) & F(s)(1 - F(s)) \end{pmatrix}$$

Huomataan myös että Gaussisella prosessilla $(G(s) : s \in \mathbb{R})$ on esitys $G(s) = X(F(s))$ jossa $(X(u) : u \in [0, 1])$ on keskitetty Gaussinen prosessi kovarianssille $E_P(X(u)X(v)) = (u \wedge v)(1 - u \vee v)$. $X(u)$ kutsutaan Brownin sillaksi. Sen jakauma on sama kuin Brownin liikkeen prosessin $(W(u) : u \in [0, 1])$ ehdollinen jakaumaa ehdolla tapahtumaa $\{W(1) = 0\}$.

2. Olkoon

$$e_k(t) = \frac{1}{2\pi} \exp(ikt), \quad k \in \mathbb{Z}, \quad t \in (-\pi, \pi]$$

(a) Osoita että $\{e_n : n \in Z\}$ ovat ortonormaalisia $L^2((-\pi,\pi],dt)$ avaruudessa eli.

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(ikt)dt = \mathbf{1}(k=0) \quad \forall k \in \mathbb{Z}$$

R

$$\int \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(i(n+k)t)\exp(-in)dt = \int \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(ikt)dt$$
$$= \begin{cases} \frac{1}{ik}\Big\{\exp(ik2\pi)-\exp(ik2\pi)\Big\} = 0 & k \neq 0 \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} 1dt = 1 & k = 0 \end{cases}$$

(b) Olkoon $X(\omega)$ satunnaismuuttuja jolla $P(X \in \mathbb{Z}) = 1$, karakteristella funktiolla $\varphi_X(t)$.

$$P(X=k) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi_X(t)\exp(-ikt)dt \quad \forall k \in \mathbb{Z}$$

R

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\varphi_X(t)\exp(-ikt)dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Bigg(\sum_{z\in\mathbb{Z}}\exp(int)P(X=n)\Bigg)\exp(-ikt)dt = \sum_{z\in\mathbb{Z}} P(X$$

Summan ja integroinnin jäerjestyksen vaihto on sallittu koska Fubinin lauseen oletus on voimassa $|\exp(i(z-k)t)| \leq 1$ jossa

$$\sum_{z\in\mathbb{Z}} P(X=z)\frac{1}{2\pi}\int_{-\pi}^{\pi} 1dt = 1 < \infty$$

(c) Olkoon $S_n(\omega) = X_1(\omega) + \cdots + X_n(\omega)$ jossa $X_1,\ldots,X_n$ ovat P-riippumattomia kopioita satunnaimuuttujasta $X(\omega) \in \mathbb{Z}$. Osoita

$$P(S_n=k) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\big(\varphi_X(t)\big)^k\exp(-ikt)dt \quad \forall k \in \mathbb{Z}$$

Koska $P(X_1 \in \mathbb{Z}) = 1$, seuraa että $P(S_n \in \mathbb{Z}) = 1$ ja tehtävän edellisen pykälä pätee S_n:lle jolla on karateristinen funktio $\varphi_{S_n}(t) = \varphi_{S_n}(t) = \big(\varphi_{S_n}(t)\big)^n$.