9.3.21

תרגול 1 - ① [illegible]

$X_1, \dots, X_n \overset{iid}{\sim} F$

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1\{X_i \le x\}$$

שאלה ①

(א) הראו שתחת P_r $\quad n\hat{F}_n(x) \sim Bin(n, F(x))$

$n\hat{F}_n(x) = \sum_{i=1}^{n} 1\{X_i \le x\}$ סכום של n מ"מ iid המתפלגים $Ber(\underbrace{P_r(X_1 \le x)}_{\text{הסיכוי לקבל את התכונה } \{X_i \le x\}})$

⇐ ולכן התוצאה נובעת.

(ב) הוכיחו כי לכל $x_1 < \dots < x_k$ וקטור ההפרשים

$n(\hat{F}_n(x_1), \hat{F}_n(x_2) - \hat{F}_n(x_1), \dots, \hat{F}_n(x_k) - \hat{F}_n(x_{k-1}), 1 - \hat{F}_n(x_k)) \sim multi$ בעל התפלגות מולטינומית ולמצוא את הפרמטרים.

תזכורת: אפיון התפלגות מולטינומית

יהיו $\psi_1, \dots, \psi_n$ הם מ"מ בת"ל שמקבלים ערכים $\{a_1, \dots, a_k\}$

כך ש $P(\psi_1 = a_1) = p_i$

נגדיר: $y_j = \sum_{i=1}^{n} 1\{\psi_i = a_j\}$ $\quad j=1,\dots,k$ אז לווקטור $(y_1, \dots, y_k)$ יש התפלגות

$Mult(n, p) \quad p = (p_1, \dots, p_k)$

במקרה שלנו, נסמן $x_j < x_{j+1}$ $\quad j = 0, \dots, k$

ונגדיר $x_0 = -\infty \Rightarrow P(X_1 \le x_0) = 0$

$x_{k+1} = \infty \Rightarrow P(X_1 \le x_{k+1}) = 1$

$$n(\hat{F}_n(x_{j+1}) - \hat{F}_n(x_j)) = n \cdot \frac{1}{n} \sum_{i=1}^{n} (1\{X_i \le x_{j+1}\} - 1\{X_i \le x_j\}) =$$

$$= \underbrace{\sum_{i=1}^{n} 1\{x_j \le X_i \le x_{j+1}\}}_{\sim Ber(\underbrace{F(x_{j+1}) - F(x_j)}_{\text{הסיכוי לקבל את התכונה}})} = y_i$$

$p_i = F(x_{j+1}) - F(x_j)$

ולכן הווקטור מתפלג $Mult(n, p)$

שאלה 2

הראו שלכל התפלגות F $\mathbb{E}_F \int_{\mathbb{R}} [\sqrt{n}(\hat{F}_n(x)-F(x))^2]\, dF(x) \le \frac{1}{4}$

וכאשר F רציפה הביטוי שווה ל- $\frac{1}{6}$. הסיקו כי $\hat{F}_n$ עקיבה.

פתרון

$\hat{F}_n(x) = \frac{1}{n}\sum_{j=1}^{n} 1\{X_j \le x\}$

$$\mathbb{E}_F \int_{\mathbb{R}} [\sqrt{n}(\hat{F}_n(x)-F(x))]^2 dF(x) = \int_{\mathbb{R}} \underbrace{\mathbb{E}_F[n(\hat{F}_n(x)-F(x))]^2}_{var(\hat{F}_n(x))} dF(x) =$$

$$= \int_{\mathbb{R}} n \frac{var(1\{X_j \le x\})}{n} dF(x) \underset{!}{=} \int_{\mathbb{R}} F(x)(1-F(x))\, dF(x) \le$$

שונות של מקדם.

$$\le \int_{\mathbb{R}} \frac{1}{4} dF(x) = \frac{1}{4}$$

מקרה שבו F רציפה

$$\int_{\mathbb{R}} (F(x)(1-F(x))\, dF(x) = \int_{\mathbb{R}} F(x)(1-F(x)) f(x)\, dx =$$

$dF(x) = \underbrace{f(x)dx}_{\text{צפיפות}}$

החלפת משתנים.

החלף משתנים: $u = F(x)$
$du = dF(x) = f(x)dx$

$$= \int_0^1 u(1-u)\, du = \frac{u^2}{2} - \frac{u^3}{3} \Big|_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

ומכאן:

$$\int_{\mathbb{R}} (\hat{F}_n(x)-F(x))^2 dF(x) = \frac{1}{6n} \le \frac{1}{4n} \xrightarrow[n\to\infty]{} 0 \Rightarrow \text{עקיבה.}$$

שאלה 3

(a) הראו כי MISE מקיים: $\sup_{F \in \mathcal{F}} \mathbb{E}_F \|\sqrt{n}(\hat{F}_n - F)\|_2^2 = \infty$

$F\in\mathcal{F}$: Poiss מכל התפלגות דיסקרטית

תזכורת: MISE

$$MISE = \underbrace{\mathbb{E}_F \int_{\mathbb{R}} (\hat{F}_n(x) - F(x))^2 dx}_{\mathbb{E}_F \|\hat{F}_n - F\|_2^2}$$

<u>פתרון</u>

$$\mathbb{E}_F \|\sqrt{n}(\hat{F}_n - F)\|_2^2 = \mathbb{E}_F \int_{\mathbb{R}} [\sqrt{n}(\hat{F}_n(x) - F(x))]^2 dx =$$

$$\underset{\text{בהרצאה ②}}{=} \int_{\mathbb{R}} F(x)(1-F(x))\,dx$$

נמצא פונק' שעבורה הינו אינסופי עבור F אינסופי.

נבחר לדוג' את התפל' קושי (Caushy) סטנדרטי:

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2} \quad ; \quad F(x) = \int_{-\infty}^{x} \frac{1}{\pi}\frac{1}{1+t^2}\,dt = \frac{1}{\pi}\arctan(x) + \frac{1}{2}$$

נראה שלא פונ' זו אינטגרבילית:

עבור $x \le -1$ מתקיים כי $F(x) < \frac{1}{2}$ (כיוון ש $F(0) = \frac{1}{2}$) ולכן $1 - F(x) > \frac{1}{2}$

$$(1-F(x))F(x) \ge \frac{1}{2}F(x) = \frac{1}{2}\int_{-\infty}^{x} \frac{1}{\pi}\frac{1}{1+t^2}\,dt \ge \frac{1}{2\pi}\int_{-\infty}^{x} \frac{1}{t^2}\,dt =$$

$$= \frac{1}{2\pi}\left(-\frac{1}{t}\Big|_{t=-\infty}^{x}\right) = -\frac{1}{2\pi}\frac{1}{x}$$

ולכן:

$$\int_{-\infty}^{-1} (1-F(x))F(x)\,dx \ge \int_{-\infty}^{-1} -\frac{1}{2\pi}\frac{1}{x}\,dx = \int_{1}^{\infty} \frac{1}{2\pi}\frac{1}{x}\,dx \longrightarrow \infty$$

② הראו כי MISE מקיים $\sup_{F \in \mathcal{F}_M} \mathbb{E}_F \|\sqrt{n}(\hat{F}_n(x) - F(x))\| \le M$

כאשר $\mathcal{F}_M = \{F \in \mathcal{F} : \int_{\mathbb{R}} |x|\,dF(x) \le M\}$ <u>רמז</u>: עבור מ"מ $\xi \ge 0$

$$\mathbb{E}\xi = \int_0^\infty 1 - F_\xi(x)\,dx$$

<u>פתרון</u>:

$$\mathbb{E}_F \|\sqrt{n}(\hat{F}_n(x) - F(x))\|_2^2 \underset{\text{לפי קודם}}{=} \int_{\mathbb{R}} F(x)(1-F(x))\,dx = \int_{-\infty}^{0} F(x)(1-F(x))\,dx +$$

$$+ \int_0^\infty F(x)(1-F(x))\,dx \le \int_{-\infty}^{0} F(x)\,dx + \int_0^\infty F(x)\,dx \quad (*)$$

כי $F(x)$ בין 0 ל-1 ולכן $F(x)(1-F(x)) \le F(x)$

המטרה היא להראות כי שווה ל $\mathbb{E}_F|x|$. <u>תזכורת</u>: $|x| = x^+ + x^-$

כאשר $\begin{cases} x^+ = (x \vee 0) \\ x^- = -(x \wedge 0) \end{cases}$ אי שליליים

עבור X^+, $x \ge 0$ (עבור $x < 0$ $F_{X^+} = 0$)

$$F_{X^+}(x) = P(X^+ \le x) = P(X^+ \le x, x \le 0) + P(X^+ \le x, x > 0) = F(0) + F(x) - F(0) = F(x)$$

$$= \underbrace{P(X^+ \le x \mid x \le 0)}_{=1}\underbrace{P(x \le 0)}_{F(0)} + \underbrace{P(X^+ \le x \mid x > 0)}_{\frac{P(0 \le x \le x^+)}{P(x > 0)} = F(x) - F(0)} \cancel{P(x > 0)}$$

לכן: $EX^+ \underset{\text{רמז}}{=} \int_0^\infty (1-F_{X^+}(x))\,dx = \int_0^\infty (1-F(x))\,dx$

באותו אופן, עבור X^-, כש $X\geq 0$ ונקבל ש:

$$\boxed{F_{X^-}(x) = 1-F(-x)}$$

ולכן: $EX^- = \int_0^\infty 1-F_{X^-}(x)\,dx = \int_0^\infty F(-x)\,dx = \int_{-\infty}^0 F(x)\,dx$

<u>ומכאן לאי שוויון:</u> $(*) = EX^- + EX^+ = E|X| \leq M \quad \forall F\in\mathcal{F}_M$

הסקנו: ש $\hat{F}_n(x)$ עקבית עם $\mathcal{F}_M$ מבחינת נורמת L^2

$$E_F\|\hat{F}_n - F\|_2^2 \leq \sup_{F\in\mathcal{F}_M} E_F\|\hat{F}_n - F\|_2^2 \leq \frac{M}{n} \xrightarrow[n\to\infty]{} 0$$

<u>שאלה 7</u>

יהיו $X_1,\dots,X_n \overset{iid}{\sim} F$ כאשר F רציפה.

אם נשתמש ב $\hat{F}_n$ (פונ' ההתפ' האמפירית) כדי לאמוד את F נקבל שהסיכון המינימלי שווה ל $\frac{1}{6}$. נניח כי מוגבלת ש $F\in(F_\lambda)_{\lambda\in\mathbb{R}^+}$ כך ש $F_\lambda\sim exp(\lambda)$.

תשבו את הגבול כאשר $n\to\infty$ של הסיכון עבור $F_{\hat{\lambda}_n}$ כאשר $\hat{\lambda}_n$ הוא $MLE(\lambda)$

<u>פתרון</u>

$$\hat{\lambda}_n = \frac{1}{\bar{X}_n}$$

$$E_\lambda \int_0^\infty [n\cdot(F_{\hat{\lambda}}(x)-F(x))]^2\,dF(x) = n E_\lambda \int_0^\infty \left(e^{-\hat{\lambda}_n x} - e^{-\lambda x}\right)\lambda e^{-\lambda x}\,dx =$$

$F_\lambda(x) = 1-e^{-\lambda x}$

רציפה ולכן $dF(x)=f(x)dx$

יש כאן דילוג על הרבה מעברים

$$= n E_\lambda \int_0^\infty \left(e^{-x(\lambda+2\hat{\lambda}_n)} - 2e^{-x(2\lambda+\hat{\lambda}_n)} + e^{-3\lambda x}\right)dx =$$

דילוג על מעברים

$$= n\lambda E_\lambda\left(\frac{1}{\lambda+2\hat{\lambda}_n} - \frac{2}{2\lambda+\hat{\lambda}_n} + \frac{1}{3\lambda}\right) = E_\lambda \frac{2}{3}\underbrace{\frac{(\sqrt{n}(\bar{X}_n - 1/\lambda)\lambda)^2}{(\lambda\bar{X}_n+2)(2\lambda\bar{X}_n+1)}}_{Y_n}$$

דילוג על מעבר

<u>מהלך:</u> (*) נרצה להראות שניתן להחליף גבול ותוחלת.

(*) לצורך כך נוכיח ש Y_n מ"מ אינטגרבילי במ"ש.

<u>(*) תזכורת de la Vallee Pousin:</u>

Y_n אינטגרבילים במ"ש אם"ם קיימת פונ' φ כך ש $\lim_{x\to\infty}\frac{\varphi(x)}{x}=\infty$

כך ש $\sup E\varphi(|Y_n|)<\infty$

במקרה שלנו נבחר $\varphi(x)=x^2$ ונראה כי $\le C\in\mathbb{R}$, y_n , 8 (ובגלל שהמקדם חסום

[illegible]

[illegible]

$$y_n \xrightarrow{d} \frac{2}{2?} Z^2 \quad , Z\sim N(0,1)$$