האוניברסיטה העברית בירושלים
המחלקה לסטטיסטיקה
הסתברות ותהליכים מקריים
מורה הקורס : עופר קלע

מידה מסומנת היא פונקציה

$$\nu : \mathcal{F} \to \mathbb{R} \cup \{\infty\}$$

או

$$\nu : \mathcal{F} \to \mathbb{R} \cup \{-\infty\}$$

המקיימת

$$\nu(\phi) = 0$$

ולכל $A_1, A_2, \ldots \in \mathcal{F}$ זרים בזוגות מתקיים כי

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu\left(A_n\right)$$

הפרוק של Haan :
קיים $B \in \mathcal{F}$ כך שלכל $A \in \mathcal{F}$ מתקיים כי

$$\begin{aligned} \nu(A \cap B) &\geq 0 \\ \nu(A \cap B^c) &\leq 0 \end{aligned}$$

נשים לב כי
1. $(A \cap B) \cup (A \cap B^C) = \Omega$
2. $(A \cap B) \cap (A \cap B^C) = \emptyset$

שמקיים את תכונות של הקבוצה B מתקיים כי לכל $E \subset (B \triangle C) \cup (B^C \triangle C^C)$ $v(E) = 0$.
כלומר אם קיימת עוד קבוצה שחוצה את המרחב Ω שאני חי בו .
אז היא נבדלת מ B (בגלל זה אני מסתכל על $B \triangle C$) עד כדי מידה 0 .

לכל C אחר המקיים זאת מתקיים כי המידה של כל תת קבוצה מדידה של $(B\Delta C) \cup (B^c \Delta C^c)$ היא אפס.

שתי מידות μ_1, μ_2 נקראות סינגולריות אחת ביחס לשניה אם קיים $B \in \mathcal{F}$ כך ש-$\mu_1(B) = 0$ ו-$\mu_2(B^c) = 0$. אם נסתכל על

$$\begin{aligned} \nu^+(A) &= \nu(A \cap B) \\ \nu^-(A) &= -\nu(A \cap B^c) \end{aligned}$$

אז $\nu^\pm$ הן מידות סינגולריות ומתקיים כי $\nu = \nu^+ - \nu^-$. לא יתכן כי $\nu(B) = \infty$ וגם $\nu(B^c) = -\infty$ כי אז

$$\nu(\Omega) = \nu(B) + \nu(B^c)$$

והסכום מצד ימין אינו מוגדר. לכן אחד מהם חייב להיות סופי ולכן לפחות אחת מהמידות $\nu^\pm$ היא סופית. מהפרוק של Haan נובע כי ההצגה של ν כהפרש של שתי מידות סינגולריות הוא יחיד. דהיינו אם ν_1, ν_2 הן מידות סינגולריות ומתקיים כי $\nu = \nu_1 - \nu_2$ אז $\nu_1 = \nu^+$ ו-$\nu_2 = \nu^-$.
פרוק זה נקרה הפרוק של Jordan.
נסמן $|\nu(A)| = \nu^+(A) + \nu^-(A)$. נניח כי μ היא מידה ו-ν היא מידה מסומנת. נאמר כי ν רציפה בהחלט ביחס ל-μ ונסמן $\nu \ll \mu$ אם לכל $A \in \mathcal{F}$ עבורו $\mu(A) = 0$ מתקיים כי $\nu(A) = 0$.
משפט Radon-Nikodym

נניח כי μ היא מידה σ-סופית וכי ν היא מידה מסומנת σ-סופית (בפרט $\nu^{\pm}$ הן σ-סופיות). אם מתקיים כי $\mu \ll \nu$ אז קיימת פונקציה מדידה g כך ש-

$$\nu(A) = \int_A g(x)d\mu(x)$$

לכל פונקציה מדידה אחרת h המקיימת זאת מתקיים כי $g = h$ כמעט בכל מקום ביחס ל-μ, דהיינו לקבוצה $\{x|h(x) \neq g(x)\}$ יש מידה אפס.

יישום אחד של משפט זה הוא כדלהלן : נניח כי X הוא משתנה מקרי כך שמידת ההתפלגות שלו $P^X(B) = P(X \in B)$ רציפה בהחלט ביחס למידת לבג. דהיינו לכל קבוצת בורל B בעלת מידת לבג אפס מתקיים כי $P(X \in B) = 0$, אז קיימת פונקציית בורל f כך שלכל קבוצת בורל B מתקיים

כאן עופר מראה שפונק צפיפות f_X של משתנה מקרי איננה שלילית !

$$P^X(B) = \int_B f(x)m(dx)$$

נסמן ב-$B_n = \{x|f(x) \leq -1/n\}$. מכיוון ש-$f$ היא פונקציית בורל, B_n היא קבוצת בורל ולכן

$$0 \leq P^X(B_n) = \int_{B_n} f(x)m(dx) \leq -\frac{1}{n}\int_{B_n} m(dx) = -\frac{1}{n}m(B_n) \leq 0$$

מכאן נובע כי $m(B_n) = 0$ לכל $n \geq 1$. כמו כן, $B_n \subset B_{n+1}$ ומתקיים כי $\bigcup_{n=1}^{\infty} B_n = B \equiv \{x|f(x) < 0\}$. לכן

$$m(B) = \lim_{n\to\infty} m(B_n) = 0$$

לכן, אם נסמן $f_X(x) = f^+(x)$ נקבל כי f_X היא פונקצית בורל אי שלילית אשר נבדלת מ-f על קבוצת בורל שמידת לבג שלה היא אפס ולכן מתקיים כי לכל קבוצת בורל B

$$E1_B(X) = P(X \in B) = \int_B f_X(x)m(dx) = \int_{\mathbb{R}} 1_B(x)f_X(x)m(dx)$$

מכאן גם שעבור פונקציות בורל פשוטות מתקיים כי

$$Eg(X) = \int_{\mathbb{R}} g(x)f_X(x)dx$$

ולכן גם עבור גבולות לא יורדים ואי שליליים שלהן. כלומר הנוסחה הזאת מתקיימת לכל פונקצית בורל אי שלילית ולבסוף היא גם מתקיים לכל פונקציית בורל עבורה או $Eg^+(X) < \infty$ או $Eg^-(X) < \infty$. f_X נקראת "צפיפות" והיא אינה יחידה, אך נבדלת מכל צפיפות אחרת על קבוצת בורל שמידת לבג שלה היא אפס. דהיינו, שתי צפיפויות של X שוות אחת לשניה כמעט בכל מקום (כ.ב.מ.) ביחס למידת לבג.

באותו אופן, ניתן נאמר כי עבור זוג המשתנים המקריים (X,Y) יש צפיפות אם $P^{(X,Y)}$ רציפה בהחלט ביחס למידת לבג $m \otimes m$ על $\left(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2)\right)$ כאשר $\mathcal{B}(\mathbb{R}^2) \equiv \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. אז באותו אופן, קיימת פונקצית בורל אי שלילית $f_{XY}(x,y)$ המקיימת לכל $C \in \mathcal{B}(\mathbb{R}^2)$ כי

$$P((X,Y) \in C) = \int_C f_{XY}(x,y)dm \otimes m(x,y)$$

כאן עופר מפרמל את נוסחת ההסתברות השלמה לצפיפות

ממשפט טונלי-פוביני אנו יודעים כי $f_{XY}(x,\cdot)$ היא פונצקית בורל לכל x קבוע, כי $f_{XY}(\cdot,y)$ היא פונצקית בורל לכל y קבוע, כי $f_X(x)=\int_{\mathbb{R}} f_{XY}(x,y)dm$ היא פונצקית בורל (בהכרח אי שלילית) ובאותו אופן גם $f_Y(y)=\int_{\mathbb{R}} f_{XY}(x,y)dm(x)$. כמו כן, מתקיים כי

$$\begin{aligned} P(X\in B) &= P((X,Y)\in B\times\mathbb{R}) = \int_{B\times R} f_{XY}(x,y)dm\otimes m(dx,dy) = \\ &= \int_B \left(\int_{\mathbb{R}} f_{XY}(x,y)dm(y)\right)dm(x) = \int_B f_X(x)dm(x) \end{aligned}$$

ובאותו אופן רק שמשתמשים בשוויון השני במשפט טונלי-פוביני:

$$P(Y\in B) = \int_B f_Y(y)dm(y)$$

טענה: נניח כי X,Y הם משתנים מקריים שלכל אחד מהם יש התפלגות בעלת צפיפות f_X ו-f_Y בהתאמה (דהיינו, רציפה בהחלט ביחס למידת לבג). אז X,Y הם בלתי תלויים אם ורק אם ההתפלגות של (X,Y) היא בעלת צפיפות השווה ל-$f_X(x)f_Y(y)$ כ.ב.מ. ביחס למידת לבג.
הוכחה: אם ל-(X,Y) צפיפות משותפת $f_{XY}(x,y)=f_X(x)f_Y(y)$ אז ממשפט טונלי פונביני נובע כי לכל זוג קבוצות בורל A,B

$$\begin{aligned} P(X\in A,Y\in B) &= \int_{A\times B} f_X(x)f_Y(y)dm\otimes m(x,y) \\ &= \int_B\left(\int_A f_X(x)dm(x)\right)f_Y(y)dm(y) \\ &= \int_A P(X\in A)f_Y(y)dm(y) = P(X\in A)\int_B f_Y(y)dm(y) \\ &= P(X\in A)P(Y\in B) \end{aligned}$$

ולכן X,Y בלתי תלויים. הכיוון ההפוך: נניח כי X,Y בלתי תלויים. אז לכל זוג קבוצות בורל A,B מתקיים מהמשוואה האחרונה כי

$$P(X\in A,Y\in B) = P(X\in A)P(Y\in B) = \int_{A\times B} f_X(x)f_Y(y)dm\otimes m(x,y)$$

ממה שהראינו בכיתה, ניתן להרחיב את $P(X\in A,Y\in B)$ למידה יחידה על $(\mathbb{R}^2,\mathcal{B}(\mathbb{R}^2))$ המקיימת לכל $C\in\mathcal{B}(\mathbb{R}^2)$ כי

$$P((X,Y)\in C) = \int_C f_X(x)f_Y(y)dm\otimes m(x,y)$$

מכאן שלמידה זו יש צפיפות והצפיפות היא $f_X(x)f_Y(y)$ כמעט בכל מקום ביחס למידת לבג.
לבסוף, מכיוון ש-$f_X(x)$ היא פונצקית בורל (חד מימדית) ו-$f_{XY}(x,y)$ היא פונצקית בורל (דו מימדית) אז אם עבור $f(x,y)$ פונקצית בורל אי שלילית כלשהי המקיימת $\int_{\mathbb{R}} f(x,y)dy=1$ נסמן:

$$f_{Y|X=x}(y) = \begin{cases} \frac{f_{XY}(x,y)}{f_X(x)} & f_X(x)\neq 0 \\ f(x,y) & f_X(x)=0 \end{cases}$$

אז זו גם כן צפיפות הנקראת הצפיפות המותנה של Y בהנתן X וזו פונקציית בורל של שני המשתנים x, y. שימו לב כי לכל x, y כך ש-$f_X(x) > 0$ מתקיים כי

$$f_{XY}(x,y) = f_X(x) f_{Y|X=x}(y)$$

כמו כן

$$P(X \in \{x | f_X(x) = 0\}) = \int_{\{x|f_X(x)=0\}} f_X(x) dm(x) = 0$$

שימו לב כי

$$(x,y) | f_{XY}(x,y) \neq f_X(x) f_{Y|X=x}(x) \subset \{(x,y) | f_X(x) = 0\} = \{x | f_X(x) = 0\} \times \mathbb{R}$$

מידת לבג (הדו-מימדית) של הקבוצה שמופיעה בצד ימין היא אפס ומכאן ש-$f_{XY}(x,y) = f_X(x) f_{Y|X=x}(y)$ כמעט בכל מקום ביחס ל-$m \otimes m$.ואם נסמן עבור פונקציית בורל g (של שני משתנים)

$$E(g(X,Y) | X = x) = \int_B g(x,y) f_{Y|X=x}(y) dm(y)$$

אז ממשפט טונלי-פוביני נובע כי זו פונקציית בורל של x המקיימת

$$\begin{aligned}
\int_{\mathbb{R}} E(g(X,Y)|X=x) f_X(x) m(dx) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x,y) f_{Y|X=x}(y) dm(y) \right) f_X(x) dm(x) \\
&= \int_{\mathbb{R}^2} g(x,y) f_{Y|X=x}(y) f_X(x) dm \otimes m(x,y) \\
&= \int_{\mathbb{R}^w} g(x,y) f_{XY}(x,y) dm \otimes m(x,y) \\
&= Eg(X,Y)
\end{aligned}$$

כאשר השוויון השלישי נובע מכך ש-$f_{Y|X=x}(y) f_X(x) = f_{XY}(x,y)$ כמעט בכל מקום ביחס ל-$m \otimes m$. בפרט, אם ניקח $g(X,Y) = Y$ נקבל כי

$$\int_{\mathbb{R}} E(Y|X=x) f_X(x) m(dx) = EY$$

יותר מאוחר נראה איך זה מתקשר לתוחלת מותנה כפי שנגדיר אותה עכשיו.
בהנתן מרחב הסתברות $(\Omega, \mathcal{F}, P)$ ומשתנה מקרי X בעל תוחלת מוגדרת, דהיינו, עם $EX^+ < \infty$ או $EX^- < \infty$, נסמן

$$\nu(A) = EX1_A$$

אז ν היא מידה מסומנת (מדוע?) המקיימת כי $P \ll \nu$. בפרט אם $\mathcal{F}_0 \subset \mathcal{F}$ אז אם נסתכל על P ו-ν כמידות על $(\Omega, \mathcal{F}_0)$ אז גם מעל מרחב זה מתקיים כי $P \ll \nu$ ולכן קיים משתנה מקרי שנסמנו ב-$E[X|\mathcal{F}_0]$ על $(\Omega, \mathcal{F}_0)$ המקיים לכל $A \in \mathcal{F}_0$ כי

$$EX1_A = \nu(A) = \int_A E[X|\mathcal{F}_0](\omega) dP(\omega) = EE[X|\mathcal{F}_0]1_A$$

כלומר, קיים משתנה מקרי $E[X|\mathcal{F}_0]$ המדיד ביחס ל-$\mathcal{F}_0$ המקיים לכל $A \in \mathcal{F}_0$ כי

$$EE[X|\mathcal{F}_0]1_A = EX1_A$$

כל שני משתנים מקריים המקיימים זאת שווים אחד לשני בהסתברות אחת.
תוצאות:

1. $EE[X|\mathcal{F}_0] = EX$. הוכחה: פשוט מציבים בהגדרה $A = \Omega$.

2. אם $X \in \mathcal{F}_0$ אז $E[X|\mathcal{F}_0] = X$ בהסתברות אחת. הוכחה: מכיוון ש-$X \in \mathcal{F}_0$ ו-$EX1_A = EX1_A$ לכל $A \in \mathcal{F}_0$ אז X מקיים את התכונות הנדרשות.

3. X בלתי תלוי ב-$\mathcal{F}_0$ אז $E[X|\mathcal{F}_0] = EX$ בהסתברות אחת. הוכחה: EX הוא קבוע ולכן כמשתנה מקרי הוא בכל σ-שדה ובפרט ב-$\mathcal{F}_0$. ברור כי $E(EX)1_A = EX \cdot P(A)$. מצד שני

$$EE[X|\mathcal{F}_0]1_A = EX1_A = EXE1_A = EXP(A)$$

כאשר השוויון השני נובע מאי התלות בין X ו-1_A.

4. אם $X \geq 0$ בהסתברות אחת אז $E[X|\mathcal{F}_0] \geq 0$ בהסתברות אחת ואם בנוסף $EX < \infty$ אז גם $E[X|\mathcal{F}_0] < \infty$ בהסתברות אחת. הוכחה: נסמן $A_n = \left\{\omega | E[X|\mathcal{F}_0](\omega) \leq -\frac{1}{n}\right\}$. מכיוון ש-$E[X|\mathcal{F}_0]$ מדיד ביחס ל-$\mathcal{F}_0$ אז A_n היא קבוצה ב-$\mathcal{F}_0$ מדידה ולכן

$$EE[X|\mathcal{F}_0]1_{A_n} = EX1_{A_n}$$

מכיוון ש-$X \geq 0$ בהסתברות אחת ו-$1_{A_n} \geq 0$ נובע כי $EX1_{A_n} \geq 0$. עכשיו,

$$0 \leq EX1_{A_n} = EE[X|\mathcal{F}_0]1_{A_n} \leq -\frac{1}{n}E1_{A_n} = -\frac{1}{n}P(A_n) \leq 0$$

ולכן $P(A_n) = 0$. מכיוון ש-$A_n \subset A_{n+1}$ לכל n ומתקיים כי $\bigcup_{n=1}^{\infty} A_n = \{\omega | E[X|\mathcal{F}_0](\omega) < 0\}$ נובע מרציפות ההסתברות כי

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = P(E[X|\mathcal{F}_0] < 0) = \lim_{n\to\infty} P(A_n) = 0$$

ולכן $E[X|\mathcal{F}_0] \geq 0$ בהסתברות אחת. עכשיו, אם בנוסף $EX < \infty$ אז מכיוון ש-$E[X|\mathcal{F}_0] \geq 0$ בהסתברות אחת, נובע מאי שוויון מרקוב כי

$$P(E[X|\mathcal{F}_0] \geq n) \leq \frac{EE[X|\mathcal{F}_0]}{n} = \frac{EX}{n}$$

לכל $n \geq 1$ ומכיוון ש-$EX < \infty$ נובע שאם נשאיף $n \to \infty$ נקבל כי

$$P(E[X|\mathcal{F}_0] = \infty) = 0$$

ומכאן ש-$E[X|\mathcal{F}_0] < \infty$ בהסתברות אחת.

5. לכל קבוע c, $E[cX|\mathcal{F}_0] = cE[X|\mathcal{F}_0]$ בהסתברות אחת. הוכחה:

$$E\left(E[cX|\mathcal{F}_0]1_A\right) = E(cX1_A) = cEX1_A = cEE[X|\mathcal{F}_0]1_A = E(cE[X|\mathcal{F}_0])1_A$$

6. אם $EX^+, EY^+ < \infty$ (ואז גם $E(X+Y)^+ < \infty$) או $EX^-, EY^- < \infty$ (ואז גם $E(X+Y)^- < \infty$) אז $E[X+Y|\mathcal{F}_0] = E[X|\mathcal{F}_0] + E[Y|\mathcal{F}_0]$ בהסתברות אחת.
הוכחה: לכל $A \in \mathcal{F}_0$

$$\begin{aligned} EE[X+Y|\mathcal{F}_0]1_A &= E(X+Y)1_A = E(X1_A + Y1_A) \\ &= EX1_A + EY1_A = EE[X|\mathcal{F}_0]1_A + EE[Y|\mathcal{F}_0]1_A \\ &= E\left(E[X|\mathcal{F}_0]1_A + E[Y|\mathcal{F}_0]1_A\right) \\ &= E\left(E[X|\mathcal{F}_0] + E[Y|\mathcal{F}_0]\right)1_A \end{aligned}$$

7. אם $X \le Y$ בהסתברות אחת ומתקיים כי $EX^- < \infty$ (בפרט אם $X \ge 0$ בהסתברות אחת) או $EY^+ < \infty$ (בפרט אם $Y \le 0$ בהסתברות אחת), אז $E[X|\mathcal{F}_0] \le E[Y|\mathcal{F}_0]$
הוכחה: אם $EX^- < \infty$ אז מכיוון ש-$-Y \le -X$ בהסתברות אחת, אז גם $EY^- < \infty$. באותו אופן, אם $EY^+ < \infty$ אז גם $EX^+ < \infty$. במקרה הראשון, מתקיים כי $E(Y-X)^- = 0$ ולכן מסעיף 4 ו-6 נובע כי

$$E[X|\mathcal{F}_0] \le E[X|\mathcal{F}_0] + E[Y-X|\mathcal{F}_0] = E[Y|\mathcal{F}_0]$$

כאשר $EY^+ < \infty$ אז $E(-Y)^- < \infty$ ומכיוון ש-$E(Y-X)^- = 0$ אז בהסתברות אחת

$$E[-Y|\mathcal{F}_0] \le E[-Y|\mathcal{F}_0] + E(Y-X|\mathcal{F}_0] = E[-X|\mathcal{F}_0]$$

עכשיו ניעזר בסעיף 5 וסיימנו.

8. אם $0 \le X_n \le X_{n+1}$ בהסתברות אחת לכל $n \ge 1$ אז

$$\lim_{n\to\infty} E[X_n|\mathcal{F}_0] = E\left[\lim_{n\to\infty} X_n|\mathcal{F}_0\right]$$

בהסתברות אחת. הוכחה: נסמן $X = \lim_{n\to\infty} X_n$ (כפי שראינו בעבר, הגבול קיים בהסתברות אחת). מכיוון שלפי סעיף 7 נובע כי $E(X_n|\mathcal{F}_0)$ היא סדרה אי שלילית ולא יורדת בהסתברות אחת אז גם לה יש גבול בהסתברות אחת שנסמנו ב-Y. ברור כי $Y \in \mathcal{F}_0$ מכיוון ש-Y הוא גבול של משתנים מקריים ב-$\mathcal{F}_0$. לכל $A \in \mathcal{F}_0$. מתקיים כי

$$EY1_A = E\lim_{n\to\infty} E(X_n|\mathcal{F}_0)1_A = \lim_{n\to\infty} EE(X_n|\mathcal{F}_0)1_A = \lim_{n\to\infty} EX_n1_A = EX1_A$$

כאשר השוויון השני והאחרון נובעים ממשפט ההתכנסות המונוטונית הרגיל. מכיוון ש-$Y \in \mathcal{F}_0$ נובע מהגדרת התוחלת המותנה כי $Y = E(X|\mathcal{F}_0)$ בהסתברות אחת כנדרש.

9. אם $X_n \ge 0$ בהסתברות אחת אז

$$E\left[\liminf_{n\to\infty} X_n|\mathcal{F}_0\right] \le \liminf_{n\to\infty} E[X_n|\mathcal{F}_0]$$

בהסתברות אחת. הוכחה: נובע באותו אופן כמו במקרה של הלמה של Fatou הרגילה.

10. אם $|X_n| \le Y$ בהסתברות אחת, $EY < \infty$ ו-$X_n \to X$ בהסתברות אחת, אז

$$\lim_{n\to\infty} E[X_n|\mathcal{F}_0] = E[X|\mathcal{F}_0]$$

בהסתברות אחת. הוכחה: נובע באותו אופן כמו משפט ההתכנסות הנשלטת הרגילה.

11. אם $E|X|, E|Y|, E|XY| < \infty$ ומתקיים כי $Y \in \mathcal{F}_0$ אז

$$E[XY|\mathcal{F}_0] = YE[X|\mathcal{F}_0]$$

הוכחה: : עבור $Y = 1_B$ כאשר $B \in \mathcal{F}_0$ זה נובע מכיוון ש-

$$EE[X1_B|\mathcal{F}_0]1_A = EX1_B1_A = EX1_{A\cap B} = EE[X|\mathcal{F}_0]1_{A\cap B} = E\left(1_B E[X|\mathcal{F}_0]\right)1_A$$

לכן זה נכון עבור Y פשוט (לצורך זה מספיק ש-$EX^+ < \infty$ או $EX^- < \infty$). מכאן נובע באמצעות סעיף 8 כי אם $X \geq 0$ אז התוצאה נכונה כאשר $Y \geq 0$. נשים לב כי במקרה הכללי גם $Y^+, Y^- \in \mathcal{F}_0$ ונקבל כי

$$\begin{aligned}
E[X^+Y^+|\mathcal{F}_0] &= Y^+E[X^+|\mathcal{F}_0] \\
E[X^+Y^-|\mathcal{F}_0] &= Y^-E[X^+|\mathcal{F}_0] \\
E[X^-Y^+|\mathcal{F}_0] &= Y^+E[X^-|\mathcal{F}_0] \\
E[X^-Y^-|\mathcal{F}_0] &= Y^-E[X^-|\mathcal{F}_0]
\end{aligned}$$

התוצאה הסופית מתקבלת באמצעות סעיפים קודמים אם נשים לב כי

$$XY = (X^+ - X^-)(Y^+ - Y^-) = X^+Y^+ - X^+Y^- - X^-Y^+ + X^-Y^-$$

12. אם $\mathcal{F}_0 \subset \mathcal{F}_1$ אז בהסתברות אחת מתקיים כי

$$E[E[X|\mathcal{F}_0]|\mathcal{F}_1] = E[E[X|\mathcal{F}_1]|\mathcal{F}_0] = E[X|\mathcal{F}_0]$$

הוכחה: מכיוון ש-$E[X|\mathcal{F}_0]$ מדיד ביחס ל-$\mathcal{F}_0$ ולכן גם ביחס ל-$\mathcal{F}_1$ נובע מסעיף 2 כי

$$E[E[X|\mathcal{F}_0]|\mathcal{F}_1] = E[X|\mathcal{F}_0]$$

עכשיו, נניח כי $A \in \mathcal{F}_0$. אז

$$EE[E[X|\mathcal{F}_1]|\mathcal{F}_0]1_A = E[E[X|\mathcal{F}_1]1_A$$

אך מכיוון ש-$A \in \mathcal{F}_1$ (כי $\mathcal{F}_0 \subset \mathcal{F}_1$) מתקיים גם כי

$$E[E[X|\mathcal{F}_1]1_A = EX1_A$$

אם כן, $E[E[X|\mathcal{F}_1]|\mathcal{F}_0]$ מדיד ביחס ל-$\mathcal{F}_0$ ומקיים כי

$$EE[E[X|\mathcal{F}_1]|\mathcal{F}_0]1_A = EX1_A$$

לכל $A \in \mathcal{F}_0$ ולכן הוא שווה בהסתברות אחת ל-$E[X|\mathcal{F}_0]$.

13. נניח כי $A_1, \ldots, A_n \in \mathcal{F}$ זרים בזוגות עם $P\left(\bigcup_{i=1}^n A_i\right) = 1$ ו-$P(A_i) > 0$ לכל $i = 1, \ldots, n$. נסמן $\mathcal{F}_0 = \sigma(A_1, \ldots, A_n)$. אז בהסתברות אחת

$$E[X|\mathcal{F}_0] = \sum_{i=1}^n \frac{EX1_{A_i}}{P(A_i)}1_{A_i}$$

אם במקום $P\left(\bigcup_{i=1}^{n} A_i\right) = 1$ מניחים כי $\bigcup_{i=1}^{n} A_i = \Omega$ אז השוויון הוא לכל ω ולא רק בהסתברות אחת. בפרט, התוחלת המותנה כאן היא משתנה מקרי המקבל את הערך $EX1_{A_i}/P(A_i)$ בהסתברות $P(A_i)$.למשל, אם ניקח זוג משתנים (U,V) מקריים בעלי התפלגות בדידה ונגדיר

$$\begin{aligned} X &= 1_{\{V=v_j\}} \\ A_i &= \{U=u_i\} \end{aligned}$$

ונסמן

$$P(V=v_j|\sigma(U)) = E[X|\sigma(U)]$$

אז נקבל כי

$$P[V=v_j|\sigma(U)] = \sum_{i=1}^{n} \frac{P(U=u_i, V=v_j)}{P(U=u_i)} 1_{\{U=u_i\}} = \sum_{i=1}^{n} P(V=v_j|U=u_i) 1_{\{U=u_i\}}$$

14. נניח כי (X,Y) בעל צפיפות ועבור פונצקית בורל אי שלילית שרירותית מסויימת $f(x,y)$ המקיימת $\int_{\mathbb{R}} f(x,y)dy = 1$ נסמן

$$f_{Y|X=x}(y) = \begin{cases} \frac{f_{XY}(x,y)}{f_X(x)} & f_X(x) \neq 0 \\ f(x,y) & f_X(x) = 0 \end{cases}$$

וכן

$$h(x) = E[g(X,Y)|X=x] \equiv \int_{\mathbb{R}} g(x,y) f_{Y|X=x}(y)dy$$

אז בהסתברות אחת

$$E[g(X,Y)|\sigma(X)] = h(X)$$

הוכחה: תרגיל.