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

$\mathcal{F}_1,\mathcal{F}_2$ נקראים בלתי תלויים אם לכל $A_1\in\mathcal{F}_1$ ו-$A_2\in\mathcal{F}_2$ מתקיים כי $P(A_1\cap A_2)=P(A_1)P(A_2)$.

נניח כי X,Y הם משתנים מקריים. אז התנאים הבאים שקולים. אם אחד מהם מתקיים (ואז גם כולם) אז X,Y נקראים בלתי תלויים.

1. $\sigma(X)$ ו-$\sigma(Y)$ בלתי תלויים.
2. לכל זוג קבוצות בורל A,B מתקיים כי $P(X\in A,Y\in B)=P(X\in A)P(Y\in B)$
3. לכל s,t מתקיים כי $\mathbb{P}(X\le s,Y\le t)=\mathbb{P}(X\le s)\cdot\mathbb{P}(Y\le t)$
4. לכל f,g פונקציות בורל אי שליליות או שעבורן מתקיים כי $E|f(X)|<\infty$ ו-$E|g(Y)|<\infty$ מתקיים כי $Ef(X)g(Y)=Ef(X)Eg(Y)$.
5. לכל f,g רציפות וחסומות מתקיים כי $Ef(X)g(Y)=Ef(X)Eg(Y)$.
6. לכל f,g בורל וזוג קבוצות בורל מתקיים כי

$$P(f(X)\in A,g(Y)\in B)=P(f(X)\in A)P(g(Y)\in B)$$

(דהיינו, $f(X)$ ו-$g(Y)$ בלתי תלויים).

הוכחה :

$2\Leftrightarrow 1$: מכיוון ש-$\sigma(X)=\left\{X^{-1}(B)|B\in\mathcal{B}(\bar{\mathbb{R}})\right\}$ אז 1 ו-2 הן בעצם אותו הדבר כי כל מאורע ב-$\sigma(X)$ הוא מהצורה $\{X\in A\}$ עבור קבוצת בורל A כלשהי (ואותו הדבר לגבי Y).

$3\Rightarrow 2$: אוסף הקבוצות $\mathcal{C}$ מהצורה $[-\infty,s]$ סגור תחת חיתוכים סופיים. כמו כן האוסף

$$(a)\quad \{A|P(X\in A,Y\le t)=P(X\in A)P(Y\le t)\}$$

הרעיון הוא להראות תחילה שהקבוצה (a) היא בורל ואז לכל $A\in Borel(\bar{\mathbb{R}})$ נקבל כי (b) היא בורל גם כן בעצם קיבענו פעם אחת את Y ופעם אחרת את X

מכיל את $\bar{\mathbb{R}}$, את $\mathcal{C}$, סגור תחת הפרשים וגבולות לא יורדים (בדקו) ולכן אוסף זה מכיל את המחלקה המונוטונית הקטנה ביותר שמכילה את $\mathcal{C}$ ולפי משפט המחלקה המונוטונית המחלקה המונוטונית הקטנה ביותר היא בהכרח $\sigma(\mathcal{C})=\mathcal{B}\left(\bar{\mathbb{R}}\right)$. לכן, לכל קבוצת בורל מתקיים כי

$$(b)\quad P(X\in A,Y\le t)=P(X\in A)P(Y\le t)$$

בהנתן קבוצת בורל A, גם האוסף

$$\{B|P(X\in A,Y\in B)=P(X\in A)P(Y\in B)\}$$

מכיל את Ω, את $\mathcal{C}$, סגור תחת הפרשים וגבולות לא יורדים ולכן גם אוסף זה הוא $\mathcal{B}\left(\bar{\mathbb{R}}\right)$. מכאן שתנאי 2 מתקיים לכל זוג קבוצות בורל A,B.

$2\Rightarrow 4$: ניקח תחילה $f=1_A$ ו-$g=1_B$ ונקבל כי

$$\begin{aligned}Ef(X)g(Y)&=E1_A(X)1_B(Y)=E1_{A\times B}(X,Y)=P(X\in A,Y\in B)\\&=P(X\in A)P(Y\in B)=E1_A(X)E1_B(Y)=Ef(X)Eg(Y)\end{aligned}$$

מהלינאריות של התוחלת זה מתקיים עבור פונקציות (בורל) פשוטות ולכן גם עבור גבולות לא יורדים של פונקציות פשוטות ואי שליליות. מכאן שזה מתקיים כאשר f, g פונקציות בורל אי שליליות. אם הן לא בהכרח אי שליליות אז

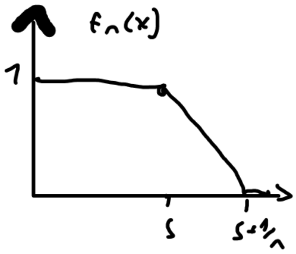
$$Ef^{\pm}(X)g^{\pm}(Y) = Ef^{\pm}(X)Eg^{\pm}(Y)$$

יש כאן ארבעה מקרים. מכיוון ש-$E|f(X)|, E|g(Y)|$ סופיים אז גם $Ef^{\pm}(X), Eg^{\pm}(Y)$ סופיים ומהשוויון האחרון גם $Ef^{\pm}(X)g^{\pm}(Y)$ סופיים (שימו לב כי גם $E|f(X)g(Y)| < \infty$ כי $E|f(X)g(X)| = E|f(X)|E|g(Y)|$). עכשיו

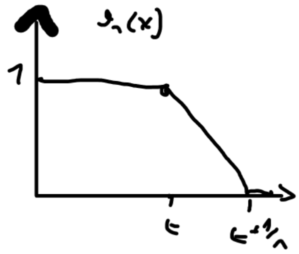
$$\begin{aligned}
&Ef(X)g(Y) = E(f^+(X) - f^-(X))(g^+(Y) - g^-(Y)) \\
&= E(f^+(X)g^+(Y) - f^-(X)g^+(Y) - f^+(X)g^-(Y) + f^-(X)g^-(Y)) \\
&= Ef^+(X)g^+(Y) - Ef^-(X)g^+(Y) - Ef^+(X)g^-(Y) + Ef^-(X)g^-(Y) \\
&= Ef^+(X)Eg^+(Y) - Ef^-(X)Eg^+(Y) - Ef^+(X)Eg^-(Y) + Ef^-(X)Eg^-(Y) \\
&= (Ef^+(X) - Ef^-(X))(Eg^+(Y) - Eg^-(Y)) \\
&= Ef(X)Eg(Y)
\end{aligned}$$

$4 \Rightarrow 5$: פונקציות רציפות הן בהכרח בורל ומכיוון שהן חסומות אז $E|f(X)|$ ו-$E|g(Y)|$ סופיות.

$5 \Rightarrow 3$: ניקח



$$\begin{aligned}
f_n(x) &= \begin{cases} 1 & x \leq s \\ 1 - n(x-s) & s < x < s + 1/n \\ 0 & x \geq s + 1/n \end{cases} \\
g_n(y) &= \begin{cases} 1 & y \leq t \\ 1 - n(y-t) & t < y < t + 1/n \\ 0 & y \geq t + 1/n \end{cases}
\end{aligned}$$



אז f_n, g_n רציפות וחסומות ולכן

$$Ef_n(X)g_n(Y) = Ef_n(X)Eg_n(Y)$$

עכשיו

$$\lim_{n\to\infty} f_n(x) = 1_{[-\infty,s]}(x), \ \lim_{n\to\infty} g_n(y) = 1_{[-\infty,t]}(y)$$

ולכן ממשפט ההתכנסות החסומה מתקיים כי

$$\begin{aligned}
P(X \leq s, Y \leq t) &= E1_{[-\infty,s]}(X)1_{[-\infty,t]}(Y) = E1_{[-\infty,s]}(X)E1_{[-\infty,t]}(Y) \\
&= P(X \leq s)P(Y \leq t)
\end{aligned}$$

$2 \Leftrightarrow 6$: אם ב-6 ניקח $f(x) = x$ ו-$g(y) = y$ נקבל את 2. כמו כן, אם נניח את 2 אז $f(X) \in A$ ו-$g(Y) \in B$ אם ורק אם $X \in f^{-1}(A)$ ו-$Y \in g^{-1}(B)$. מכיוון ש-f, g פונקציות בורל אז גם $f^{-1}(A)$ ו- $g^{-1}(B)$ הן קבוצות בורל ולכן התוצאה ב-6 נובעת מ-2.
המסקנה המיידית ממה שמופיע לעיל הוא שאם $E|X|, E|Y|$ סופיים אז גם $E|XY|$ סופי ומתקיים כי

$$EXY = EXEY$$

דהיינו, כאשר המשתנים המקריים בלתי תלויים ובעלי תוחלת סופית, השונות המשותפת מוגדרת ושווה לאפס. שימו לב כי בדרך כל, יש להניח כי EX^2, EY^2 סופיים כדי שגם EXY תהיה מוגדרת וסופית (אי שוויות קושי שוורץ), אך כאשר המשתנים המקריים בלתי תלויים זו הנחה מיותרת. כמובן אפשר באינדוקציה להכליל את מה שכתבנו עבור שני משתנים לאוסף סופי כלשהו של משתנים. כזכור, אוסף אינסופי של משתנים מקריים הם בלתי תלויים אם ורק אם בכל תת אוסף סופי המשתנים בלתי תלויים. אם האוסף הוא בן מניה, אז לא צריך להסתכל על כל תת אוסף. בפרט, התנאים הבאים שקולים (תרגיל) ושניהם מהווים הגדרה של $X_1, X_2, \ldots$ הם בלתי תלויים.

1. $X_{n_1}, X_{n_2}, \ldots, X_{n_m}$ הם בלתי תלויים לכל $m \geq 2$ ולכל $n_1 < \ldots < n_m$.

2. $X_1, \ldots, X_n$ הם בלתי תלויים לכל $n \geq 1$.

עבור אוסף בן מנייה **בלבד של** משתנים מקריים

שימו לב כי מרציפות ההסתברות זה שקול לכך שלכל $B_1, B_2, \ldots$ קבוצות בורל מתקיים כי

$$P\left(\bigcap_{i=1}^{\infty}\{X_i \in B_i\}\right) = \prod_{i=1}^{\infty} P\left(X_i = B_i\right)$$

מהמקרה הסופי מקבלים את השוויון האחרון על ידי לקיחת גבול בשני האגפים ואילו מהשוויון האחרון מקבלים את המקרה הסופי על ידי כך שניקח $B_i = \bar{\mathbb{R}}$ לכל $i \notin \{1, \ldots, n\}$ (או $i \notin \{n_1, \ldots, n_m\}$).
יהיה לנו עוד משהו לומר על משתנים מקריים בלתי תלויים כשנדבר על צפיפויות, אך בשלב זה נעבור לנושא הבא. נניח כי $(\Omega_1, \mathcal{F}_1)$ ו-$(\Omega_2, \mathcal{F}_2)$ הם מרחבים מדידים. אז, עבור $A_i \subset \Omega_i$, $i = 1, 2$ ניזכר כי "מכפלה קרטזית" מוגדרת באופן הבא:

$$A_1 \times A_2 = \{(\omega_1, \omega_2) | \omega_1 \in A_1, \omega_2 \in A_2\}$$

שימו לב כי מההגדרה נובע כי המכפלה הקרטזית של קבוצה כלשהי עם הקבוצה הריקה היא הקבוצה הריקה. דהיינו, $A_1 \times \phi = \phi \times A_2 = \phi \times \phi = \phi$.
נגדיר

$$\mathcal{F}_1 \times \mathcal{F}_2 = \{A_1 \times A_2 | A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$$

(זו אמנם אינה מכפלה קרטזית אך במקרה זה נהוג להשתמש בסימון זה בכל זאת). בדרך כלל $\mathcal{F}_1 \times \mathcal{F}_2$ אינו σ-שדה (המשלים של מכפלה קרטזית לא חייב להיות מכפלה קרטזית ואיחוד של מכפלות קרטזיות לא חייב להיות מכפלה קרטזית) אבל הוא כן סגור תחת חיתוכים סופיים. בפרט מתקיים כי

$$(A_1 \times A_2) \cap (B_1 \times B_2) = (A_1 \cap B_1) \times (A_2 \cap B_2)$$

קיים כמובן σ-שדה קטן ביותר שמכיל את אוסף הקבוצות הזה. נסמן אותו ב-

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma\left(\mathcal{F}_1 \times \mathcal{F}_2\right)$$

טענה:
נניח כי $C \in \mathcal{F}_1 \otimes \mathcal{F}_2$. אז

$$\begin{aligned} C_{\cdot\omega_2} &= \{\omega_1 | (\omega_1, \omega_2) \in C\} \in \mathcal{F}_1 \\ C_{\omega_1\cdot} &= \{\omega_2 | (\omega_1, \omega_2) \in C\} \in \mathcal{F}_2 \end{aligned}$$

הראשון לכל $\omega_2 \in \Omega_2$ והשני לכל $\omega_1 \in \Omega_1$.
הוכחה:
נסמן

$$\mathcal{H} = \{C | C \in \mathcal{F}_1 \otimes \mathcal{F}_2,\ C_{\cdot\omega_2} \in \mathcal{F}_1,\ \forall \omega_2\}$$

ברור כי $\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{H}$ כי

$$(A_1 \times A_2)_{\omega_2 \cdot} = \begin{cases} A_1 & \omega_2 \in A_2 \\ \phi & \omega_1 \notin A_2 \end{cases}$$

ולכן $\mathcal{H}$ אינו ריק. עכשיו

$$\begin{aligned}(C^c)_{\cdot\omega_2} &= \{\omega_1 | (\omega_1, \omega_2) \in C^c\} = \{\omega_1 | (\omega_1, \omega_2) \notin C\} \\ &= \{\omega_1 | (\omega_1, \omega_2) \in C\}^c = C^c_{\cdot\omega_2}\end{aligned}$$

ולכן אם $C \in \mathcal{H}$ אז גם $C^c \in \mathcal{H}$. כמו כן,

$$\left(\bigcup_{n=1}^{\infty} C_n\right)_{\cdot\omega_2} = \left\{\omega_1 | (\omega_1, \omega_2) \in \bigcup_{n=1}^{\infty} C_n\right\} = \bigcup_{n=1}^{\infty} \{\omega_1 | (\omega_1, \omega_2) \in C_n\} = \bigcup_{n=1}^{\infty} (C_n)_{\cdot\omega_2}$$

ולכן אם $C_1, C_2, \ldots \in \mathcal{H}$ אז גם $\bigcup_{n=1}^{\infty} C_n \in \mathcal{H}$ ומכאן ש-$\mathcal{H}$הוא σ-שדה המכיל את $\mathcal{F}_1 \times \mathcal{F}_2$ ומוכל ב-$\mathcal{F}_1 \otimes \mathcal{F}_2$. לכן $\mathcal{H} = \mathcal{F}_1 \otimes \mathcal{F}_2$.
טענה:
נניח כי $f : \Omega_1 \times \Omega_2 \to \bar{\mathbb{R}}$ היא פונקציה $\mathcal{F}_1 \otimes \mathcal{F}_2$-מדידה. דהיינו, f היא משתנה מקרי על $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$. אז לכל $\omega_2 \in \Omega_2$ קבוע, הפוקציה $g(\omega_1) = f(\omega_1, \omega_2)$ (כפונקציה של ω_1) היא $\mathcal{F}_1$-מדידה, דהיינו, משתנה מקרי על $(\Omega_1, \mathcal{F}_1)$. באותו אופן, הפונצקיה $h(\omega_2) = f(\omega_1, \omega_2)$ היא $\mathcal{F}_2$-מדידה לכל $\omega_1 \in \Omega_1$ נתון.
הוכחה:
עבור אינדיקטורים זה נובע מהטענה הקודמת. לכן זה נובע גם עבור פונצקיות פשוטות ולכן גם עבור גבולות לא יורדים של פונצקיות פשוטות אי שליליות ומדידות. לבסוף כל פונצקיה מדידה היא הפרש של שתי פונצקיות מדידות אי שליליות.
עכשיו, נניח כי $(\Omega_i, \mathcal{F}_i, P_i)$, $i = 1, 2$ הם מרחבי הסתברות.
טענה:
לכל $C \in \mathcal{F}_1 \otimes \mathcal{F}_2$ מתקיים כי

$$P_2(C_{\omega_1 \cdot})$$

היא פונצקיה $\mathcal{F}_1$-מדידה.
הוכחה:
נסמן

$$\mathcal{H} = \{C | C \in \mathcal{F}_1 \otimes \mathcal{F}_2,\ P_2(C_{\omega_1 \cdot}) \in \mathcal{F}_1\}$$

אם $D \subset C$ שניהם ב-$\mathcal{H}$ אז לכל ω_1 מתקיים כי $C_{\omega_1 \cdot} \subset D_{\omega_1 \cdot}$ וכ-$(D \setminus C)_{\omega_1 \cdot} = D_{\omega_1 \cdot} \setminus C_{\omega_1 \cdot}$
ולכן

$$P_2\left((D \setminus C)_{\omega_1 \cdot}\right) = P_2\left((D)_{\omega_1 \cdot}\right) - P_2\left((C)_{\omega_1 \cdot}\right)$$

צד ימין היא פונקציה מדידה כי היא הפרש של פונקציות מדידות ולכן גם צד שמאל. מכאן ש- $D \setminus C \in \mathcal{H}$, כלומר $\mathcal{H}$ סגורה תחת הפרשים. כמו כן אם $C^1 \subset C^2 \subset \ldots$ הן ב-$\mathcal{H}$ אז גם $C^1_{\omega_1\cdot} \subset C^1_{\omega_1\cdot} \cdots$ הן קבוצות ב-$\mathcal{F}_2$ ומכיוון ש-

$$\left(\bigcup_{n=1}^{\infty} C^n\right)_{\cdot\omega_2} = \bigcup_{n=1}^{\infty} C^n_{\omega_1\cdot}$$

נובע כי

$$P_2\left(\left(\bigcup_{n=1}^{\infty} C^n\right)_{\cdot\omega_1}\right) = P_2\left(\bigcup_{n=1}^{\infty} C^n_{\omega_1\cdot}\right) = \lim_{n\to\infty} P_2\left(C^n_{\omega_1\cdot}\right)$$

צד ימין הוא פונקציה מדידה כי הוא גבול של פונקציות מדידות, לכן גם צד שמאל היא פונקציה מדידה ולכן $\bigcup_{n=1}^{\infty} C^n \in \mathcal{H}$. כלומר $\mathcal{H}$ סגורה תחת גבולות לא יורדים. אם כן $\mathcal{H}$ היא מחלקה מונוטונית המכילה את $\mathcal{F}_1 \times \mathcal{F}_2$ ו-$\mathcal{F}_1 \times \mathcal{F}_2$ מכילה את $\Omega_1 \times \Omega_2$ וסגורה תחת חיתוכים סופיים. לכן $\mathcal{H} = \mathcal{F}_1 \otimes \mathcal{F}_2$ (כי היא מוכלת וגם מכילה את צד ימין).

טענה: לכל $C \in \mathcal{F}_1 \otimes \mathcal{F}_2$ נסמן:

$$P_1 \otimes P_2(C) = \int_{\Omega_1} P_2(C_{\omega_1\cdot})dP_1(\omega_1)$$

(סימון לתוחלת של המשתנה המקרי $X(\omega_1) = P_2(C_{\omega_1\cdot})$ על מרחב ההסתברות $(\Omega_1, \mathcal{F}_1, P_1)$).
אז $P_1 \otimes P_2$ היא ההסתברות היחידה על $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ המקיימת לכל $A_1 \times A_2 \in \mathcal{F}_1 \times \mathcal{F}_2$ כי

$$P_1 \otimes P_2(C) = P_1(A_1)P_2(A_2)$$

הוכחה: היחידות נובעת ממשהו שהראינו בעבר. דהיינו, שניתן להרחיב מידה מאוסף הסגור תחת חיתוכים סופיים ל-σ-שדה הקטן ביותר שמכיל אותו באופן יחיד. נשאר להראות כי $P_1 \otimes P_2$ היא באמת הסתברות ומקיימת את הנדרש. תחילה, נזכור כי לכל A_1, A_2 כפי שמופיעים בטענה,

$$(A_1 \times A_2)_{\omega_1\cdot} = \begin{cases} A_2 & \omega_1 \in A_1 \\ \phi & \omega_1 \notin A_1 \end{cases}$$

ולכן

$$P_2\left((A_1 \times A_2)_{\omega_1\cdot}\right) = P_2(A_2)1_{A_1}(\omega_1)$$

מכאן ש-

$$\begin{aligned} P_1 \otimes P_2(A_1 \times A_2) &= \int_{\Omega_1} P_2((A_1 \times A_2)_{\omega_1\cdot})dP_1(\omega_1) = \int_{\Omega_1} P(A_2)1_{A_1}(\omega_1)dP_1(\omega_1) \\ &= P(A_2)\int_{\Omega_1} 1_{A_1}(\omega_1)dP_1(\omega_1) = P_2(A_2)P_1(A_1) \end{aligned}$$

ובפרט

$$P_1 \otimes P_2(\Omega_1 \times \Omega_2) = P_1(\Omega_1)P_2(\Omega_2) = 1 \cdot 1 = 1$$

כמו כן אם $C^1, C^2, \ldots \in \mathcal{F}_1 \otimes \mathcal{F}_2$ זרים בזוגות אז $C^1_{\omega_1\cdot}, C^2_{\omega_2\cdot}, \ldots \in \mathcal{F}_2$ זרים בזוגות. לכן

$$P_2\left(\bigcup_{n=1}^{\infty} C^n_{\omega_1\cdot}\right) = \sum_{n=1}^{\infty} P_2\left(C^n_{\omega_1\cdot}\right)$$

וממשפט ההתכנסות המונוטונית (גרסה שניה) נובע כי

$$\begin{aligned} P_1 \otimes P_2\left(\bigcup_{n=1}^{\infty} C^n\right) &= \int_{\Omega_1} P_2\left(\bigcup_{n=1}^{\infty} C^n_{\omega_1\cdot}\right) dP_1(\omega_1) = \int_{\Omega_1} \sum_{n=1}^{\infty} P_2\left(C^n_{\omega_1\cdot}\right) dP_1(\omega_1) \\ &= \sum_{n=1}^{\infty} \int_{\Omega_1} P_2\left(C^n_{\omega_1\cdot}\right) dP_1(\omega_1) = \sum_{n=1}^{\infty} P_1 \otimes P_2\left(C^n\right) \end{aligned}$$

ולכן $P_1 \otimes P_2$ היא אמנם הסתברות.

משפט Tonelli-Fubinni להסתברויות
נניח כי f היא $\mathcal{F}_1 \otimes \mathcal{F}_2$-מדידה ואי שלילית. אז

$$\int_{\Omega_2} f(\omega_1, \omega_2) dP_2(\omega_2)$$

היא $\mathcal{F}_1$-מדידה ו-

$$\int_{\Omega_1} f(\omega_1, \omega_2) dP_1(\omega_1)$$

היא $\mathcal{F}_2$-מדידה ומתקיים כי

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2) dP_1 \otimes P_2(\omega_1, \omega_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2) dP_1(\omega_1)\right) dP_2(\omega_2) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) dP_2(\omega_2)\right) dP_1(\omega_1) \end{aligned}$$

אם f אינה בהכרח אי שלילית אך עבור f^+ או f^- אחד מהאינטגרלים (ואז כולם המופיעים במשוואה האחרונה הוא סופי אז השוויון האחרון מתקיים עבור f.

הוכחה:
עבור אינדיקטורים זה נובע מהטענה הקודמת. לכן זה נכון עבור פונקציות פשוטות ומדידות וגבולות לא יורדים של פונקציות פשוטות ומדידות. דהיינו, עבור פונקציות מדידות ואי שליליות. אם f אינה בהכרח אי שלילית, אז השוויון מתקיים עבור f^+ וגם עבור f^-. במקרה זה

$$\int_{\Omega_2} f(\omega_1, \omega_2) dP_2(\omega_2) = \int_{\Omega_2} f^+(\omega_1, \omega_2) dP_2(\omega_2) - \int_{\Omega_2} f^-(\omega_1, \omega_2) dP_2(\omega_2)$$

ואז צד ימין היא פונקציה מדידה כי הוא הפרש של פונקציות מדידות. באותו אופן עבור האינטגרל ביחס ל-P_1. השוויון הסופי מתקבל כאשר נכתוב אותו בנפרד עבור f^+ ו-f^- ואז נחסיר.
בהינתן מרחב מדיד $(\Omega, \mathcal{F})$ הפונקציה $\mu : \mathcal{F} \to [0, \infty]$ נקראת "מידה" אם היא מקיימת

1. $\mu(\phi) = 0$,

2. אם $A_1, A_2, \ldots \in \mathcal{F}$ זרים בזוגות אז $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu\left(A_n\right)$.

מכאן שהסתברות P היא קודם כל מידה אך בנוסף מקיימת כי $P(\Omega) = 1$. כשהגדרנו הסתברות לא הנחנו כי $P(\phi) = 0$ אך זאת מכיוון שבמקרה זה ההנחה הזו היא מיותרת מכיוון שהיא בהכרח מתקיימת.

מידה ניקראת σ-סופית אם קיימים $A_1, A_2, \ldots$ זרים בזוגות ואיחודם הוא Ω כך ש-$\mu\left(A_n\right) < \infty$. ברור כי אם המידה היא סופית, דהיינו, $\mu(\Omega) < \infty$ (בפרט אם מדובר במידת הסתברות) אז אפשר לקחת $A_1 = \Omega$ ו-$A_n = \phi$ לכל $n \geq 2$ ולהסיק כי מידה סופית היא גם σ-סופית.

אם μ היא מידה σ-סופית אך לא סופית, $A_n \in \mathcal{F}$ זרות בזוגות ואיחודם הוא Ω ו-$\mu\left(A_n\right) < \infty$ אז יש אוסף אינסופי של אינדכסים n עבורם $\mu\left(A_n\right) > 0$. זאת מכיוון שאם זה לא היה נכון אז $\mu(\Omega) = \sum_{n=1}^{\infty} \mu\left(A_n\right)$ היה מספר סופי בניגוד להנחה ש-$\mu(\Omega) = \infty$. עכשיו נסמן

$$N = \{n \mid \mu\left(A_n\right) = 0\}$$

אז

$$\mu\left(\bigcup_{n \in N} A_n\right) = \sum_{n \in N} \mu\left(A_n\right) = \sum_{n \in N} 0 = 0$$

ולכן אם $k \notin N$ הוא אינדכס כלשהו (יש אינסוף כאילה) ונחליף את A_k ב-

$$A_k \cup \bigcup_{n \in N} A_n$$

אז

$$\mu\left(A_k \cup \bigcup_{n \in N} A_n\right) = \mu\left(A_k\right) + \mu\left(\bigcup_{n \in N} A_n\right) = \mu\left(A_k\right) + 0 = \mu\left(A_k\right)$$

נשארנו איפה עם אינסוף מאורעות זרים בזוגות שאיחודם הוא Ω שהמידה של כל אחד מהם היא חיובית. לכן, כאשר μ היא σ-סופית ולא סופית אז קיימים מאורעות $A_n \in \mathcal{F}$ זרים בזוגות שאיחודם הוא Ω כך ש-$0 < \mu\left(A_n\right) < \infty$ לכל $n \geq 1$. עבור מאורעות כאילה נסמן

$$\lambda_n = \mu\left(A_n\right) \qquad P_n(A) = \frac{\mu\left(A \cap A_n\right)}{\mu\left(A_n\right)}$$

ונקבל כי P_n הן מידות הסתברות על $(\Omega, \mathcal{F})$ ומתקיים לכל $A \in \mathcal{F}$ כי

$$\mu(A) = \sum_{n=1}^{\infty} \lambda_n P_n(A)$$

מכאן שכל מידה σ-סופית היא קומבינציה חיובית של מידות הסתברות שניצבות אחת לשניה במובן שקיימים $A_n \in \mathcal{F}$ זרים בזוגות ואיחודם הוא Ω כך ש-$P_k\left(A_n\right) = 0$ לכל $n \neq k$.

את משפט טונלי-פוביני ניתן להכליל למידות σ-סופיות μ_1 ו-μ_2 והתוצאה היא לגמרי זהה פרט לכך ש-$\mu_1 \otimes \mu_2$ היא לא בהכרח מידת הסתברות, אבל היא כן מידה σ-סופית. ניתן לראות זאת על ידי זה שניקח $A_n \in \mathcal{F}_1$ זרים בזוגות כך שאיחודם הוא Ω_1 ו-$B_n \in \mathcal{F}_2$ זרים בזוגות כך שאיחודם הוא Ω_2 ואז לשים לב כי האיחוד של $A_k \times B_n$ על כל $k, n \geq 1$ הוא $\Omega_1 \times \Omega_2$ ומתקיים כי

$$\mu_1 \otimes \mu_2\left(A_k \times B_n\right) = \mu_1\left(A_k\right) \mu_2\left(B_n\right) < \infty$$

נניח כי $\mu_1(A_k), \mu_2(B_n) > 0$ ואז ממשפט טונלי-פוביני עבור הסתברויות נקבל כי

$$\int_{B_n} f(\omega_1, \omega_2) d\mu_2(\omega_2)$$

היא $\mathcal{F}_1$-מדידה ו-

$$\int_{A_k} f(\omega_1, \omega_2) d\mu_1(\omega_1)$$

היא $\mathcal{F}_2$-מדידה ומתקיים כי

$$\begin{aligned}\int_{A_k \times B_n} f(\omega_1, \omega_2) \frac{d\mu_1 \otimes \mu_2(\omega_1, \omega_2)}{\mu_1 \otimes \mu_2 (A_k \times B_n)} &= \int_{A_k} \left(\int_{B_n} f(\omega_1, \omega_2) \frac{d\mu_1(\omega_1)}{\mu_1 (A_k)} \right) \frac{d\mu_2(\omega_1)}{\mu_2 (B_n)} \\ &= \int_{B_n} \left(\int_{A_k} f(\omega_1, \omega_2) \frac{d\mu_2(\omega_1)}{\mu_2 (B_n)} \right) d\frac{d\mu_1(\omega_1)}{\mu_1 (A_k)}\end{aligned}$$

על ידי סיכום או הכפלה ב-$\mu_1 (A_k) \mu_2 (B_n)$ וסיכום ושימוש במשפט ההתכנסות המונוטונית עבור מידות σ-סופיות נקבל:

משפט Tonelli-Fubini כללי:
בהנתן מרחבי מידה σ-סופיים $(\Omega_i, \mathcal{F}_i, \mu_i)$ עבור $i = 1,2$ קיימת מידה יחידה על $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ שנסמנה ב-$\mu_1 \otimes \mu_2$ המקיימת לכל $A_i \in \mathcal{F}_i$ עבור $i = 1,2$ כי

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

מידה זו היא σ-סופית ומתקיים לכל f אי שלילית ו-$\mathcal{F}_1 \otimes \mathcal{F}_2$-מדידה כי

$$\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2)$$

היא $\mathcal{F}_1$-מדידה ו-

$$\int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1)$$

היא $\mathcal{F}_2$-מדידה ומתקיים כי

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2) d\mu_1 \otimes \mu_2(\omega_1, \omega_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right) d\mu_1(\omega_1)\end{aligned}$$

אם f אינה בהכרח אי שלילית אך עבור f^+ או f^- אחד מהאינטגרלים (ואז כולם) המופיעים במשוואה האחרונה הוא סופי אז התוצאה השוויון האחרון מתקיים עבור f.
לפני שנראה דוגמה, צריך להגדיר למה אנו מתכוונים במידת לבג על $\mathbb{R}$. אם כן, m_1 נקראת מידת לבג על $((0,1], \mathcal{B}((0,1]))$ אם היא קודם כל מידה, דהיינו, מקיימת כי $m_1(\emptyset) = 0$ ועבור

$B_1, B_2, \ldots,$ קבוצות בורל המוכלות ב-$(0,1]$ וזרות בזוגות מתקיים כי $m_1\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} m_1\left(B_i\right)$. כמו כן צריך להתקיים כי לכל $0 \le a < b \le 1$ מתקיים כי

$$m_1((a,b]) = b - a$$

אם קיימת מידה כזו האם היא יחידה? שימו לב כי $\mathcal{C} = \{(a,b] | 1 \le a < b \le 1\}$ הוא אוסף שסגור תחת חיתוכים סופיים וממה שראינו בתרגיל אם יש מידה על $\sigma(\mathcal{C})$ המוגדרת באופן הרצוי על $\mathcal{C}$ אז היא בהכרח יחידה. דהיינו, לא יכולות להיות שתי מידות שונות שיש להם את אותה תכונה. שימו לב כי $m_1(0,1]) = 1 - 0 = 1$ ולכן אם מידה כזו קיימת, אז היא מידת הסתברות. השאלה שנשארה היא מדוע מידה כזו קיימת. אנו לא נוכיח את זה, אך נתן להראות כי אם לכל תת קבוצה A של $(0,1]$

$$m_1^*(A) = \inf\left\{ \sum_{i=1}^{n} (b_i - a_i) \middle| A \subset \bigcup_{i=1}^{n} (a_i, b_i],\ 0 \le a_i < b_i \le 1,\ \forall i = 1, \ldots, n \right\}$$

אז אם נגדיר $m_1(B) = m_1^*(B)$ לכל קבוצת בורל (שימו לב כי m_1^* מוגדרת גם עבור קבוצות שאינן קבוצות בורל) אז נקבל באמצעות סדרה של תוצאות (אותן לומדים בקורס "תורת המידה") כי m_1 היא מידה על $\mathcal{B}((0,1])$. אנו נפסח על החלק הזה. לא שהוא כל כך קשה אך הוא דורש זמן ויש לנו עוד דברים ללמוד בקורס הזה. בקורס תורת המידה מראים קיום של מידת לבג על $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (ויותר) אך מכיוון שאנו מתעסקים בהסתברות כאן, יותר קל לי להסביר מה זו מידת לבג באופן שהגדרתי כאן.

ובכן, באותו אופן שקיימת מידת לבג m_1 על $((0,1], \mathcal{B}((0,1]))$ אז קיימת גם מידת לבג m_i על $((i-1,1], \mathcal{B}((i-1,i]))$. נגדיר איפה מידת לבג באופן כללי כך:

$$m(B) = \sum_{i=-\infty}^{\infty} m_i\left(B \cap (i-1,i]\right)$$

קל מאוד לבדוק כי זו המידה היחידה המקיימת לכל $-\infty < a < b < \infty$ כי $m((a,b]) = b - a$. בפרט, מרציפות המידה (באופן דומה לרציפות ההסתברות) נובע כי

$$m(\{b\}) = \lim_{a \uparrow b} m((a,b]) = \lim_{a \uparrow b} (b - a) = 0$$

ובפרט מידת לבג של כל קבוצה בת מניה היא אפס. כמו כן נובע מכך גם כי

$$m((a,b]) = m([a,b]) = m([a,b)) = m((a,b)) = b - a$$

לבסוף אם לכל $\omega \in \mathbb{R}$ נגדיר $U(\omega) = \omega$ (ואז $U^{-1}(A) = A$ לכל A) אז U הוא משתנה מקרי על $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (פונקציה לינארית, לכן רציפה ולכן מדידה). נגדיר הסתברות על $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ באופן הבא. לכל קבוצת בורל B נגדיר

$$P(B) = m(B \cap (0,1])$$

מכיוון ש-m היא מידה, נובע כי גם P היא מידה המקיימת

$$P(\mathbb{R}) = m(\mathbb{R} \cap (0,1]) = m((0,1]) = 1$$

עכשיו אם ניקח B קבוצת בורל כלשהי נקבל כי

$$P(U\in B)=P(U^{-1}(B))=P(B)=m(B\cap(0,1])$$

בפרט לכל $x<0$ נקבל

$$P(U\le x)=m((-\infty,x]\cap(0,1])=m(\emptyset)=0$$

לכל $x\ge 1$ נקבל

$$P(U\le x)=m((-\infty,x]\cap(0,1])=m((0,1])=1$$

ולכל $0\le x<1$ נקבל

$$P(U\le x)=m((-\infty,x]\cap(0,1])=m((0,x])=x$$

ולכן

$$F_U(x)=P(U\le x)=\begin{cases}0 & x<0\\ x & 0\le x<1\\ 1 & x\ge 1\end{cases}$$

התפלגות זו מוכרת לכם? אם כן, זה אומר שקיים מרחב הסתברות ומשתנה מקרי עליו המתפלג אחיד. לכן ממה שראינו בשבוע שעבר, לכל פונקצית התפלגות קיים מרחב הסתברות ומשתנה מקרי עליו שזו פונקצית ההסתברות שלו. בפרט זה גם מראה (לפי מה שהראיתם בתרגיל) כי לכל פונקצית התפלגות F קיימת מידת הסתברות יחידה Q על $\left(\bar{\mathbb{R}},\mathcal{B}\left(\bar{\mathbb{R}}\right)\right)$ המקיימת

$$Q([-\infty,x])=F(x)$$

לכל $x\in\mathbb{R}$ (מספיק להסתכל על x סופי כי את המקרה האינסופי מקבלים בגבול). מכאן שההגדרה של האינטגרל ביחס ל-F להיות התוחלת ביחס ל-Q מוצדקת.

דוגמה:

נניח שאחד המרחבים הוא מרחב ההסתברות $(\Omega,\mathcal{F},P)$ והשני הוא $(\mathbb{R},\mathcal{B}(\mathbb{R}),m)$ כאשר m היא מידת לבג. מידת לבג היא σ-סופית. נניח כי X הוא משתנה מקרי סופי על $(\Omega,\mathcal{F})$ ונסתכל על הפונצקיה

$$g(\omega,t)=\begin{cases}1 & 0\le t<X^+(\omega)\\ 0 & \text{otherwise}\end{cases}$$

זו פונצקיה אי שלילית ואם היא הייתה $\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$-מדידה אז

$$\begin{aligned}\int_{\Omega\times\mathbb{R}}g(\omega,t)dP\otimes m(\omega,t) &= \int_\Omega\left(\int_\mathbb{R}g(\omega,t)dm(t)\right)dP(\omega)\\ &= \int_\mathbb{R}\left(\int_\Omega g(\omega,t)dP(\omega)\right)dm(t)\end{aligned}$$

נשים לב כי לכל $\omega\in\Omega$

$$\int_\mathbb{R}g(\omega,t)dm(t)=\int_\mathbb{R}1_{[0,X^+(\omega))}(t)dm(t)=m([0,X^+(\omega))=X^+(\omega)$$

וזו פונקציה מדידה על $(\Omega, \mathcal{F})$. כמו כן

$$\int_\Omega \left(\int_\mathbb{R} g(\omega,t)dm(t) \right) dP(\omega) = \int_\Omega X^+(\omega)dP(\omega) = EX^+$$

עכשיו לכל $t \geq 0$ מתקיים

$$\begin{aligned} \int_\Omega g(\omega,t)dP(\omega) &= \int_\Omega 1_{(t,\infty)}(X^+(\omega))dP(\omega) = P(X^+ \in (t,\infty)) = P(X^+ > t) \\ &= P(X > t) = 1 - F_X(t) \end{aligned}$$

אם $t < 0$ אז $g(\omega, t) = 0$ לכל $\omega \in \Omega$ ואז האינגרל שחישבנו כרגע הוא אפס. דהיינו, באופן כללי קבלנו כי

$$\int_\Omega g(\omega,t)dP(\omega) = 1 - F_X(t)$$

$$\int_\mathbb{R} \left(\int_\Omega g(\omega,t)dP(\omega) \right) dm(t) = \int_{[0,\infty)} (1 - F_X(t))dm(t)$$

מכיוון ש-F_X מונוטונית אז אוסף נקודות אי הרציפות שלה הוא לכל היותר בן מניה ולכן במקרה זה ידוע כי הפונקציה היא אינטגרבילית רימן על כל קטע סופי ומתקיים כי

$$\int_{[0,\infty)} (1 - F_X(t))dm(t) = \int_0^\infty (1 - F_X(t))dt$$

כאשר האינטגרל מצד שמאל הוא אינטגרל לבג וזה שבצד ימין הוא אינטגרל רימן. אם כן, אם התנאים מתקיימים אז קבלנו כי

$$EX^+ = \int_0^\infty (1 - F_X(t)dt = \int_0^\infty P(X > t)dt$$

מכיוון שהפונקציה $P(X = t)$ שווה לאפס פרט לאוסף שהוא לכל היותר בן מניה של נקודות t אז $\int_0^\infty P(X = t)dt = 0$ ולכן גם

$$EX^+ = \int_0^\infty P(X \geq t)dt$$

אם ניקח את $-X$ במקום X אז נקבל כי

$$EX^- = E(-X)^+ = \int_0^\infty P(-X \geq t)dt = \int_0^\infty P(X \leq -t)dt = \int_0^\infty F_X(-t)dt$$

ואם נבצע את החלפת המשתנים $s = -t$ נקבל כי

$$EX^- = -\int_0^{-\infty} F_X(s)ds = \int_{-\infty}^0 F_X(s)ds$$

מכאן שאם $EX^+ = \int_0^\infty (1-F_X(t))dt < \infty$ או $EX^- = \int_{-\infty}^0 F_X(t)dt$ אז

$$\begin{aligned} EX &= -\int_{-\infty}^0 F_X(t)dt + \int_0^\infty (1-F_X(t))dt \\ E|X| &= \int_{-\infty}^0 F_X(t)dt + \int_0^\infty (1-F_X(t))dt \end{aligned}$$

כדי לראות שזה מותר נשאר להראות כי g היא $\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$-מדידה. אך מכיוון ש-$g$ היא אינדיקטור מספיק להראות כי הקבוצה הבאה מדידה:

$$\left\{(\omega,t)|0\leq t< X^+(\omega)\right\}$$

נראה את השוויון הבא

$$\left\{(\omega,t)|0\leq t< X^+(\omega)\right\} = \bigcup_{q\in\mathbb{Q}}\{\omega|X(\omega)>q\}\times\{t|0\leq t<q\} = \bigcup_{q\in\mathbb{Q}}\{X>q\}\times[0,q)$$

כאשר $[0,q)$ היא הקבוצה הריקה עבור $q\leq 0$. אם $\omega\in\{X>q\}$ וגם $t\in[0,q)$ אז גם מתקיים כי $0\leq t<X(\omega)$ ולכן גם $0\leq t<X^+(\omega)$ (כי $X(\omega)>0$). לכן (ω,t) שייך לקבוצה שמשמאל. מצד שני, אם (ω,t) שייך לקבוצה משמאל אז מכיוון ש-$0\leq t<X^+(\omega)$ אז קיים מספר רציונלי חיובי המקיים $t<q<X^+(\omega)$ ולכן $\omega\in\{X^+>q\}$ וגם $t\in[0,q)$. עכשיו, מכיוון ש-X משתנה מקרי אז

$$\{X>q\}\in\mathcal{F}$$

וברור כי

$$[0,q)\in\mathcal{B}(\mathbb{R})$$

לכן

$$\{X>q\}\times[0,q)\in\mathcal{F}\times\mathcal{B}(\mathbb{R})\subset\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$$

מכיוון ש-$\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$ הוא σ-שדה ו-$\mathbb{Q}$ היא קבוצה בת מנה אז גם

$$\bigcup_{q\in\mathbb{Q}}\{X>q\}\times[0,q)\in\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$$

ולכן

$$\left\{(\omega,t)|0\leq t< X^+(\omega)\right\}\in\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$$

ולכן g באמת $\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$-מדידה.
אפשר להכליל את הרעיון הזה באופן הבא: נתחיל במקרה ש-X הוא משתנה מקרי אי שלילי ו-h היא פונקצית בורל אי שלילית על $[0,\infty)$ ונסמן $H(x)=\int_{(0,x]}h(t)dm(t)$ עבור $x\geq 0$. בפרט $H(\cdot)$ היא פונקציה לא יורדת ורציפה המקיימת $H(0)=0$.

$$g(\omega,t)=h(t)1_{\{(\omega,t)|X(\omega)>t\}}$$

היא פונקציה $\mathcal{F}\otimes\mathcal{B}(\mathbb{R})$-מדידה (כמכפלה של שתי פונקציות מדידות). אם נחזור על הרעיון הקודם נקבל כי

$$EH(X) = E\int_0^X h(t)dm(t) = \int_{(0,\infty)} h(t)(1-F_X(t))dm(t)$$

אם h אינה בהכרח אי שלילית אז ניקח $h = h^+ - h^-$ ונקבל כי אם $\int_{(0,\infty)} h^+(t)(1-F_X(t))dm(t) < \infty$ או $\int_{(0,\infty)} h^-(t)(1-F_X(t))dm(t) < \infty$ אז

$$EH(X) = \int_{(0,\infty)} h(t)(1-F_X(t))dm(t)$$

ובפרט אם בנוסף h אינטגרבילית רימן על קטעים סופיים (רציפה כמעט בכל מקום ביחס למידת לבג וחסומה על קטעים סופיים) אז

$$EH(X) = \int_0^\infty h(t)(1-F_X(t))dt$$

למשל, עבור $h(t) = nt^{n-1}$ נקבל כי $H(x) = x^n$ עבור $x \geq 0$ ולכן

$$EX^n = n\int_0^\infty t^{n-1}(1-F_X(t))dt$$

אם X אינו בהכרח אי שלילי ונסמן $H(x) = \int_0^x h(t)dt = -\int_x^0 h(t)dt = -\int_0^{-x} h(-t)dt$ עבור $x \leq 0$ אז

$$\begin{aligned}
EH(X)1_{\{X\geq 0\}} &= EH(X^+) = \int_0^\infty h(t)(1-F_{X^+}(t))dt = \int_0^\infty h(t)(1-F_X(t))dt \\
EH(X)1_{\{X<0\}} &= H(-X^-) = -\int_0^\infty h(-t)(1-F_{X^-}(t)dt = -\int_0^\infty h(-t)P(X^- > t)dt \\
&= -\int_0^\infty h(-t)P(X^- \geq t)dt = -\int_0^\infty h(-t)P(-X \geq t)dt \\
&= -\int_0^\infty h(-t)P(X \leq -t)dt = -\int_{-\infty}^0 h(t)F_X(t)dt
\end{aligned}$$

ולכן

$$EH(X) = -\int_{-\infty}^0 h(t)F_X(t)dt + \int_0^\infty h(t)(1-F_X(t))dt$$

בתנאי ש-

$$\int_0^\infty h^+(t)(1-F_X(t)dt, \int_{-\infty}^0 h^-(t)F_X(t)dt < \infty$$

או

$$\int_0^\infty h^-(t)(1-F_X(t))dt, \int_{-\infty}^0 h^+(t)(1-F_X(t))dt < \infty$$

בפרט

$$EX^n = -n\int_{-\infty}^{0} t^{n-1}F_X(t)dt + n\int_0^{\infty} t^{n-1}(1-F_X(t))dt$$

למשל

$$EX^2 = -2\int_{-\infty}^{0} tF_X(t)dt + 2\int_0^{\infty} t(1-F_X(t))dt = 2\left(\int_{-\infty}^{0} |t|F_X(t)dt + \int_0^{\infty} |t|(1-F_X(t))dt\right)$$