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

תוחלת ומשפטי גבול

עבור $A \subset \Omega$ נסמן

$$1_A = 1_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

נשים לב לתכונות הבאות (פרט ל־6,7 הראיתם אותן בתרגיל בית):

1. $1_{A^c} = 1 - 1_A$.
2. לכל אוסף (סופי או בן מניה) של קבוצות $1_{\cap_{j\in J} A_j} = \prod_{j\in J} 1_{A_i} = \min_{j\in J} 1_{A_i}$.
3. לכל אוסף (סופי או בן מניה) של קבוצות $1_{\cup_{j\in J} A_j} = 1 - \prod_{j\in J}(1 - 1_{A_i}) = \max_{j\in J} 1_{A_j}$
4. $\limsup_{n\to\infty} 1_{A_n} = 1_{\limsup A_n}$ ו־ $\liminf_{n\to\infty} 1_{A_n} = 1_{\liminf A_n}$, לכן, הגבול $\lim_{n\to\infty} 1_{A_n}$ (לכל $\omega \in \Omega$) קיים אם ורק אם $\limsup A_n = \liminf A_n (= \lim A_n)$ ומתקיים כי $\lim_{n\to\infty} 1_{A_n} = 1_{\lim A_n}$
הוכחה: נראה רק את השוויון הראשון: מסעיף 2 נובע כי

$$\sup_{m\geq n} 1_{A_m} = 1_{\cup_{m\geq n} A_m}$$

עכשיו

$$\limsup_{n\to\infty} 1_{A_n} = \inf_{n\geq 1} \sup_{m\geq n} 1_{A_m} = \inf_{n\geq 1} 1_{\cup_{m\geq n} A_m} = 1_{\cap_{n\geq 1} \cup_{m\geq n} A_m} = 1_{\limsup A_n}$$

כאשר השוויון השלישי נובע מסעיף 1.

5. לכל אוסף של קבוצות זרות (סופי או אינסופי) $1_{\cup_{j\in J} A_j} = \sum_{j\in J} 1_A$.
6. לכל A_1, A_2, $1_{A_1\cup A_2} = 1_{A_1} + 1_{A_2} - 1_{A_1\cap A_2}$.
7. עבור $A_1, \ldots, A_n$ נסמן

$$S_k = \sum_{i_1<\ldots<i_k} 1_{A_{i_1}\cap\ldots\cap A_{i_k}}$$

אז

$$1_{A_1\cup\ldots\cup A_n} = \sum_{k=1}^{n} (-1)^{k-1} S_k$$

ההוכחה זהה להוכחה לגבי הסתברות של איחוד של מאורעות.

בהנתן מרחב מדיד $(\Omega,\mathcal{F})$, מתי נוכל לומר כי 1_A הוא משתנה מקרי? אם כן, ניקח $B\in\mathcal{B}(\mathbb{R})$. אז

$$1_A^{-1}(B)=\{\omega|1_A(\omega)\in B\}=\begin{cases}\phi & 0\notin B,1\notin B\\ A & 0\notin B,1\in B\\ A^c & 0\in B,1\notin B\\ \Omega & 0\in B,1\in B\end{cases}$$

מכיוון שזה צריך להתקיים לכל קבוצת בורל נקבל כי 1_A הוא משתנה מקרי אם ורק אם $A\in\mathcal{F}$, דהיינו, אם A הוא מאורע. אחרת הוא אינו משתנה מקרי.

עבור $A\in\mathcal{F}$ אז 1_A הוא משתנה מקרי המקבל את הערך 1 בסיכוי $P(A)$ ואת הערך 0 בסיכוי $1-P(A)$. עבור משתנה מקרי כזה אנו נגדיר את התוחלת על ידי

$$\boxed{E1_A=1\cdot P(A)+0\cdot(1-P(A))=P(A)}$$

כלומר התוחלת של אינדיקטור שווה לסיכוי של המאורע שהוא מציין.

עכשיו נניח כי X הוא משתנה מקרי המקבל n ערכים שונים אחד מהשני שנסמנם ב־$a_1,\dots,a_n$. נסמן

$$A_i\quad=\quad X^{-1}(\{a_i\})=\{\omega|X(\omega)=a_i\}$$

אז בהכרח $A_1,\dots,A_n\in\mathcal{F}$ (כי X הוא משתנה מקרי) ומתקיים כי $A_i\cap A_j=\emptyset$ לכל $i\neq j$ וגם $\cup_{i=1}^n A_i=\Omega$. ניתן לכתוב

$$X=\sum_{i=1}^n a_i 1_{A_i}$$

דהיינו, $X(\omega)=a_i$ לכל $\omega\in A_i$ ולכן X מקבל את הערכים $a_1,\dots,a_n$ בהסתברויות $P(A_1),\dots,P(A_n)$. נגדיר את התוחלת של משתנה מקרי כזה כפי שהגדרנו בקורס ראשון בהסתברות:

$$\boxed{EX=\sum_{i=1}^n a_iP(A_i)}$$

מה קורה אם $\sum_{i=1}^n a_i1_{A_i}$ כאשר $a_1,\dots,a_n$ אינם שונים (אך $A_1,\dots,A_n$ עדיין זרים בזוגות)? מקבלים $n<m$ ערכים שונים $b_1,\dots,b_m$ ואם נסמן $K_k=\{i|a_i=b_k\}$ ו־$B_k=\cup_{i\in K_k}A_i$ אז

$$X=\sum_{k=1}^m b_k1_{B_k}$$

ולכן מכיוון ש־$b_1,\dots,b_m$ ו־$B_1,\dots,B_m$ זרים בזוגות ואיחודם הוא Ω (מדוע?) נקבל מההגדרה הקודמת כי

$$\begin{aligned}EX&=\sum_{k=1}^m b_kP(B_k)=\sum_{k=1}^m b_k\sum_{i\in K_k}P(A_i)=\sum_{k=1}^m\sum_{i\in K_k}b_kP(A_i)\\&=\sum_{k=1}^m\sum_{i\in K_k}a_iP(A_i)=\boxed{\sum_{i=1}^n a_iP(A_i)}\end{aligned}$$

דהיינו, נוסחת התוחלת היא אותה נוסחה. עכשיו, נניח כי

$$X=\sum_{i=1}^n a_i 1_{A_i},\ Y=\sum_{j=1}^m b_j 1_{B_j}$$

כאשר $A_1,\ldots,A_n$ זרים בזוגות וגם $B_1,\ldots,B_m$ זרים בזוגות ומתקיים כי $\cup_{i=1}^n A_i=\cup_{j=1}^m B_j=\Omega$. נסמן $C_{ij}=A_i\cap B_j$ ו־$c_{ij}=a_i+b_j$. אז $\{C_{ij}|1\le i\le n,\ 1\le j\le m\}$ זרות בזוגות (בידקו) ומתקיים כי

$$\bigcup_{i=1}^n A_i=\bigcup_{i=1}^n\bigcup_{j=1}^m C_{ij}=\bigcup_{j=1}^m\bigcup_{i=1}^n C_{ij}=\bigcup_{j=1}^m B_j=\Omega$$

ולכן

$$\begin{aligned}E(X+Y)&=\sum_{i=1}^n\sum_{j=1}^m c_{ij}1_{C_{ij}}=\sum_{i=1}^n\sum_{j=1}^m c_{ij}P(C_{ij})=\sum_{i=1}^n\sum_{j=1}^m(a_i+b_j)P(C_{ij})\\&=\sum_{i=1}^n a_i\boxed{\sum_{j=1}^m P(A_i\cap B_j)}+\sum_{j=1}^m b_j\boxed{\sum_{i=1}^n P(A_i\cap B_j)}\\&=\sum_{i=1}^n a_i\boxed{P(A_i)}+\sum_{j=1}^m b_j\boxed{P(B_j)}=EX+EY\end{aligned}$$

אפשר עכשיו להסיק באינדוקציה בדיוק באותו אופן כי אם $X_1,\ldots,X_n$ הם משתנים מקריים המקיימים

$$X_i=\sum_{j=1}^{k_i}a_{ij}1_{A_{ij}}$$

ולכל $i=1,\ldots,n$ מתקיים כי $A_{i1},\ldots,A_{ik_i}$ זרים בזוגות עם $\cup_{j=1}^{k_i}A_{ij}=\Omega$ אז

$$E\sum_{i=1}^n X_i=\sum_{i=1}^n EX_i$$

הרעיון הוא לבחור

$$\begin{aligned}X&=\sum_{i=1}^{n-1}X_i=\sum_{i_1=1}^{k_1}\cdots\sum_{i_{n-1}=1}^{k_{n-1}}\left(a_{i_1}+\ldots+a_{i_{n-1}}\right)1_{A_{i_1}\cap\ldots\cap A_{i_{n-1}}}\\Y&=X_n\end{aligned}$$

חיתוכי המאורעות שמשתתפים ב־X הן קבוצות זרות בזוגות שאיחודם הוא Ω ולכן ממה שכבר הראינו נובע כי

$$E\sum_{i=1}^n X_i=E(X+Y)=EX+EY=E\sum_{i=1}^{n-1}X_i+EX_n$$

ולכן אם נניח באינדוקציה ש־$E\sum_{i=1}^{n-1}X_i = \sum_{i=1}^{n-1}EX_i$ נקבל כי גם $E\sum_{i=1}^{n}X_i = \sum_{i=1}^{n}EX_i$ כנדרש.

נסתכל עכשיו על משתנה מקרי מהצורה $X = a1_A$ עבור $a \neq 0$. משתנה מקרי זה מקבל את הערכים $a, 0$ בהסתברויות $P(A), 1-P(A)$ בהתאמה. לכן התוחלת היא

מוטיבציה להגדרה שהפעם אין דרישה לזרים בזוגות

$$EX = aP(A) + 0(1-P(A)) = aP(A) = aE1_A$$

בפרט אם $A = \Omega$ ואז $X(\omega) = a$ לכל $\omega \in \Omega$ ואז $EX = aP(\Omega) = a$. אם $a = 0$ אז $X(\omega) = 0$ לכל $\omega \in \Omega$ ואז $EX = 0 = 0 \cdot P(A)$. כלומר הראינו עכשיו כי אם $X = a1_A$ אז

$$EX = aP(A)$$

לכל בחירה של a, A.

עכשיו ניקח $X_i = a_i1_{A_i} = a_i1_{A_i} + 0 \cdot 1_{A_i^c}$ ונשים לב כי A_i, A_i^c זרות ואיחודן הוא Ω. ממה שהראינו לעיל נובע כי

$$E\sum_{i=1}^{n}a_i1_{A_i} = E\sum_{i=1}^{n}X_i = \sum_{i=1}^{n}EX_i = \sum_{i=1}^{n}Ea_i1_{A_i} = \sum_{i=1}^{n}a_iP(A_i)$$

הגדרה: משתנה מקרי X נקרא פשוט אם קיים $n \geq 1$ (סופי), $A_1, \dots, A_n \in \mathcal{F}$ (לא בהכרח זרים בזוגות ואיחודם אינו בהכרח Ω) ו־$a_1, \dots, a_n \in \mathbb{R}$ (סופיים) כך ש־

$$X = \sum_{i=1}^{n}a_i1_{A_i}$$

משתנה מקרי כזה בהכרח מקבל אוסף סופי של ערכים אפשריים (לא בהכרח n) ומכאן שהתוחלת שלו מוגדרת היטב. ממה שהראינו, בהכרח מתקיים כי

$$EX = \sum_{i=1}^{n}a_iP(A_i)$$

שימו לב כי לכל משתנה מקרי כזה יכולות להיות מספר הצגות שונות מהצורה $X = \sum_{i=1}^{n}a_i1_{A_i}$ אך התוחלת הוא מספר שאינו תלוי בהצגה הספציפית שבחרנו. לכל ההצגות מקבלים את אותה תוחלת (כפי שתיווכחו בתרגיל בית). כמו כן, קל לבדוק כי לכל קבוע c גם cX הוא משתנה מקרי פשוט המקיים

$$E(cX) = \sum_{i=1}^{n}ca_iP(A_i) = c\sum_{i=1}^{n}a_iP(A_i) = cEX$$

וכי אם X, Y פשוטים אז גם $X+Y$ פשוט ואז

$$E(X+Y) = E\left(\sum_{i=1}^{n}a_i1_{A_i} + \sum_{j=1}^{m}b_j1_{B_j}\right) = \sum_{i=1}^{n}a_iP(A_i) + \sum_{j=1}^{m}b_jP(B_j) = EX+EY$$

הרבה יותר קל להוכיח את זה ככה, אך היינו צריכים גם להוכיח את זה כפי שהראינו קודם כדי להסיק את נוסחת התוחלת לכל משתנה מקרי פשוט.

16. סימונים וטענות בסיסיות : שבוע 3

(a) $a \wedge b = Min(a,b)$

(b) $a \vee b = Max(a,b)$

(c) $a^+ = Max(a,0)$ וגם $a^- = -Min(a,0)$ אז נקבל כי $\begin{cases}(-a)^- = a^+ \\ (-a)^+ = a^-\end{cases}$

(d) $a = a^+ - a^-$

(e) $|a| = a^+ + a^-$

(f) $(a+b)^+ \le a^+ + b^+$

(g) $(a+b)^- = a^- + b^+$

ובאינדוקציה, אם $X_1, \dots, X_n$ פשוטים, אז לכל $c_1, \dots, c_n, d$ מתקיים כי

$$E\left(\sum_{i=1}^{n} c_i X_i + d\right) = \sum_{i=1}^{n} c_i E X_i + d$$

לבסוף, אם X, Y הם משתנים מקריים פשוטים המקיימים $X(\omega) \le Y(\omega)$ לכל $\omega \in \Omega$ אז קל לבדוק כי $EX \le EY$. זאת על ידי כך שנפרק את המרחב למאורעות זרים בזוגות $A_1, \dots, A_n$ כך ש-$X(\omega) = a_i \le b_i = Y(\omega)$ לכל $\omega \in A_i$ (כיצד?) ואז נקבל

$$EX = \sum_{i=1}^{n} a_i P(A_i) \le \sum_{i=1}^{n} b_i P(A_i)$$

נניח עכשיו כי $X(\omega) \ge 0$ לכל $\omega \in \Omega$. ונסתכל על

$$X_n(\omega) = \frac{\lfloor 2^n X \rfloor}{2^n} \wedge n$$

עכשיו, מכיוון ש-$x \le \lfloor x \rfloor$, נובע כי $2 \lfloor x \rfloor \le 2x$ ומכיוון שצד שמאל הוא מספר שלם אז גם $2 \lfloor x \rfloor \le \lfloor 2x \rfloor$ ומכאן ש-

$$\lfloor x \rfloor \le \frac{\lfloor 2x \rfloor}{2}$$

אם נציב במקום x $2^n x$ ונחלק בשני האגפים ב-2^n נקבל כי

$$\frac{\lfloor 2^n x \rfloor}{2^n} \le \frac{\lfloor 2^{n+1} x \rfloor}{2^{n+1}} \le \frac{2^{n+1} x}{2^{n+1}} = x$$

עכשיו, נשים לב כי

$$\frac{\lfloor 2^n x \rfloor}{2^n} \wedge n \le \frac{\lfloor 2^{n+1} x \rfloor}{2^{n+1}} \wedge n \le \frac{\lfloor 2^{n+1} x \rfloor}{2^{n+1}} \wedge (n+1) \le x \wedge (n+1) \le x$$

כמו כן, מכיוון ש-$x - \lfloor x \rfloor \le 1$ לכל x אז נובע גם כי

$$x - \frac{\lfloor 2^n x \rfloor}{2^n} = \frac{2^n x - \lfloor 2^n x \rfloor}{2^n} \le \frac{1}{2^n}$$

ומכאן שלכל $x \le n$ גם

$$x - \frac{\lfloor 2^n x \rfloor}{2^n} \wedge n \le \frac{1}{2^n}$$

בפרט, לכל $0 \le x < \infty$ מתקיים כי

$$\lim_{n \to \infty} \frac{\lfloor 2^n x \rfloor}{2^n} \wedge n = x$$

זאת מכיוון שעבור n מספיק גדול מתקיים כי $x \le n$. עכשיו, אם $x = \infty$ אז

$$\lim_{n \to \infty} \frac{\lfloor 2^n x \rfloor}{2^n} \wedge n = \lim_{n \to \infty} n = \infty$$

מכל מה שכתבנו עכשיו נובע כי לכל $\omega \in \Omega$ מתקיים כי

$$X_n(\omega) \leq X_{n+1}(\omega) \leq X(\omega)$$

וכי

$$\lim_{n\to\infty} X_n(\omega) = X(\omega)$$

מכיוון שהערכים האפשריים של $X_n(\omega)$ הם

$$\{k/2^n \,|\, k = 0, \ldots, n2^n\}$$

נובע כי בהכרח X_n הוא משתנה מקרי אי שלילי ופשוט (מקבל אוסף סופי של ערכים). מכאן שכל משתנה מקרי אי שלילי (לא בהכרח סופי) הוא גבול לא יורד של סדרת משתנים פשוטים ואי שליליים.

נסמן את אוסף המשתנים הפשוטים והאי שליליים ב־$\mathcal{S}_+$.

משפט:

נניח כי X_n היא סדרה לא יורדת כלשהי של משתנים מקריים פשוטים ואי שליליים ששואפת ל־X. אז בהכרח

$$\lim_{n\to\infty} EX_n = \sup\{EY | Y \in \mathcal{S}_+,\ Y \leq X\}$$

הוכחה: מכיוון ש־$X_n \leq X_{n+1}$ הם פשוטים ואי שליליים, אז ראינו כי $0 \leq EX_n \leq EX_{n+1}$. כל סדרת מספרים שואפת לגבול ולכן קיים גבול לסדרה EX_n שנסמנו ב־a (שיכול להיות גם אינסוף). כמו כן, נסמן

$$b = \sup\{EY | Y \in \mathcal{S}_+,\ Y \leq X\}$$

גם b יכול להיות אינסוף. מכיוון ש־$X_n \in \mathcal{S}_+$ לכל $n \geq 1$ נובע בהכרח כי $EX_n \leq b$ ולכן $a \leq b$. לכן, אם נראה כי מתקיים גם $a \geq b$ אז זה יוכיח את מה שרצינו. ובכן, נניח כי

$$Y = \sum_{i=1}^{m} a_i 1_{A_i} \in \mathcal{S}_+$$

אז אם נכפיל את Y באינדיקטור של מאורע כלשהו נקבל שוב משתנה מקרי פשוט ואי שלילי ואם נכפיל זאת בקבוע אי שלילי כלשהו שוב נקבל משתנה מקרי פשוט ואי שלילי. אם כן נכפיל את Y ב־

$$1_{\{(1-\epsilon)Y \leq X_n\}}$$

ולכאחר מכן נכפיל ב־$(1-\epsilon)$ כאשר $0 < \epsilon < 1$ ונקבל את המשתנה המקרי הפשוט הבא

$$\begin{aligned}
Y_{\epsilon,n} \equiv (1-\epsilon)Y1_{\{(1-\epsilon)Y\leq X_n\}} &= (1-\epsilon)\left(\sum_{i=1}^{m} a_i 1_{A_i}\right) 1_{\{(1-\epsilon)Y\leq X_n\}} \\
&= \sum_{i=1}^{m} ((1-\epsilon)a_i)\left(1_{A_i} 1_{\{(1-\epsilon)Y\leq X_n\}}\right) \\
&= \sum_{i=1}^{m} ((1-\epsilon)a_i)\left(1_{A_i\cap\{(1-\epsilon)Y\leq X_n\}}\right)
\end{aligned}$$

אם נסתכל על צד שמאל אז נשים לב כי אם $\omega \in \{(1-\epsilon)Y \leq X_n\}$ אז

$$(1-\epsilon)Y(\omega)1_{\{(1-\epsilon)Y\leq X_n\}}(\omega) = (1-\epsilon)Y(\omega) \leq X_n(\omega)$$

ואם $\omega \notin \{(1-\epsilon)Y \leq X_n\}$ אז

$$(1-\epsilon)Y(\omega)1_{\{(1-\epsilon)Y\leq X_n\}}(\omega) = 0 \leq X_n(\omega)$$

ומכאן שבכל מקרה

$$Y_{\epsilon,n}(\omega) \leq X_n(\omega)$$

לכל ω ומכאן ש־$EY_{\epsilon,n} \leq EX_n \leq a$ לכל $n \geq 1$. עכשיו, אם $X(\omega) = 0$ אז גם $X_n(\omega) = 0$ ולכן גם $Y_{\epsilon,n}(\omega) = 0$ לכל $n \geq 1$. אם לעומת זאת $X(\omega) > 0$ אז מכיוון ש־$Y(\omega) \leq X(\omega)$ נובע כי או ש־$Y(\omega) > 0$ ואז $(1-\epsilon)Y(\omega) < Y(\omega) \leq X(\omega)$ או ש־$Y(\omega) = 0$ ואז בודאי ש־$(1-\epsilon)Y(\omega) = 0 < X(\omega)$. מכיוון ש־$X_n(\omega) \to X(\omega)$ אז לכל ω יתקיים מ־n מסויים והילך כי $(1-\epsilon)Y(\omega) \leq X_n(\omega)$. עכשיו, מכיוון ש־$X_n(\omega) \leq X_{n+1}(\omega)$ נובע כי

$$\{(1-\epsilon)Y \leq X_n\} \subset \{(1-\epsilon)Y \leq X_{n+1}\}$$

וממה שהסברנו כרגע נובע כי

$$\bigcup_{n=1}^{\infty} \{(1-\epsilon)Y \leq X_n\} = \Omega$$

מכאן שגם

$$A_i \cap \{(1-\epsilon)Y \leq X_n\} \subset A_i \cap \{(1-\epsilon)Y \leq X_{n+1}\}$$

ומתקיים כי

$$\bigcup_{n=1}^{\infty} (A_i \cap \{(1-\epsilon)Y \leq X_n\}) = A_i \cap \left(\bigcup_{n=1}^{\infty} \{(1-\epsilon)Y \leq X_n\}\right) = A_i \cap \Omega = A_i$$

ואז מרציפות ההסתברות נובע כי

$$\lim_{n\to\infty} P\left(A_i \cap \{(1-\epsilon)Y \leq X_n\}\right) = P\left(\bigcup_{n=1}^{\infty} (A_i \cap \{(1-\epsilon)Y \leq X_n\})\right) = P(A_i)$$

נעזר בכך בפיתוח הבא:

$$EY_{\epsilon,n} = \sum_{i=1}^{m} ((1-\epsilon)a_i) P\left(A_i \cap \{(1-\epsilon)Y \leq X_n\}\right)$$

נזכור כי $EY_{\epsilon,n} \leq EX_n \leq a$ ולכן, אם ניקח גבול (שימו לב כי הסכום הוא סופי ולכן מותר להכניס את הגבול לתוך הסכום) נקבל כי

$$a \geq \lim_{n\to\infty} EY_{\epsilon,n} = \sum_{i=1}^{m} ((1-\epsilon)a_i) P(A_i) = (1-\epsilon)\sum_{i=1}^{m} a_i P(A_i) = (1-\epsilon)EY$$

קיבלנו איפה כי לכל $Y \in \mathcal{S}_+$ ולכל $0 < \epsilon < 1$ מתקיים כי

$$a \geq (1-\epsilon)EY$$

לכן אם נשאיף את ϵ לאפס נקבל כי

$$a \geq EY$$

לכל $Y \in \mathcal{S}_+$. מכאן שהסופרמום (החסם העליון הקטן ביותר) של צד ימין על פני כל $Y \in \mathcal{S}_+$ הוא לכל היותר a ומכאן גם ש־

$$a \geq b$$

וזה מה שנשאר להראות. מ.ש.ל.

שימו לב כי מהתוצאה לעיל נובע כי לא חשוב איזו סדרה לא יורדת של משתנים מקריים פשוטים ואי שליליים השואפת ל־X שניקח תמיד נקבל את אותו גבול של תוחלות ומכאן שגבול זה היא הגדרה הגיונית לתוחלת של X. דהיינו ההגדרה של תוחלת של משתנה מקרי אי שלילי היא

$$EX = \sup\{EY | Y \in \mathcal{S}_+,\ Y \leq X\}$$

סיבה נוספת להגדרה זו היא שאם X הוא משתנה מקרי אי שלילי וחסום (קיים $c \geq 0$ כך ש־$X(\omega) \leq c$ לכל $\omega \in \Omega$) אז אפשר היה לקחת

$$\underline{X}_n = \frac{\lfloor 2^n X \rfloor}{2^n}, \qquad \bar{X}_n = \frac{\lceil 2^n X \rceil}{2^n}$$

ואז היה מתקיים כי

$$\underline{X}_n \leq X \leq \bar{X}_n$$

וכן

$$0 \leq \bar{X}_n \leq \underline{X}_n + \frac{1}{2^n}$$

מכאן היה נובע כי

$$0 \leq E\bar{X}_n \leq E\underline{X}_n + \frac{1}{2^n}$$

ומכיוון ש־$E\underline{X}_n$ שואף ל־EX היינו מקבלים גם כי $E\bar{X}_n$ שואף ל־EX. אם מגדירים ל־X תוחלת זה הגיוני שתוחלת זו תהיה בין $E\underline{X}_n$ לבין $E\bar{X}_n$ לכל $n \geq 1$ ואם כך אז תוחלת זו חייבת להיות EX כפי שהוגדרה.

אם $P(X=\infty) > 0$ אז ניקח למשל $X_n = \frac{\lfloor 2^n X \rfloor}{2^n} \wedge n$ ונקבל כי

$$\begin{aligned} EX_n &= EX_n 1_{\{X \leq n\}} + EX_n 1_{\{X > n\}} = EX_n 1_{\{X \leq n\}} + nP(X > n) \\ &\geq nP(X>n) \geq nP(X=\infty) \to \infty \end{aligned}$$

כאן אנחנו רוצים להגדיר את התוחלת למשתנה מקרי ש $Supp = \mathbb{N}$ ולא סופי למשל אם $X \sim Geo(p)$

אם X הוא משתנה מקרי אי שלילי שמקבל אוסף בן מניה (אינסופי) של ערכים שונים וסופיים $a_1, a_2, \ldots,$ אז ברור כי

$$X_n = \sum_{i=1}^{n} a_i 1_{X^{-1}\{a_i\}}$$

היא סדרה לא יורדת של משתנים פשוטים ואי שליליים ששואפת ל־X . עכשיו, לכל $n \geq 1$

$$EX_n = \sum_{i=1}^{n} a_i P(X = a_i)$$

ואם נשאיף $n \to \infty$ נקבל כמצופה כי

$$EX = \sum_{i=1}^{\infty} a_i P(X = a_i)$$

מהגדרה זו של תוחלת מיד אפשר להסיק את הטענה הבאה.
טענה: אם X_1, X_2 משתנים מקריים אי שליליים ומתקיים לכל ω כי $X_1(\omega) \leq X_2(\omega)$ אז $EX_1 \leq EX_2$.
הוכחה: מכיוון ש־

$$\{EY | Y \in \mathcal{S}_+,\ Y \leq X_1\} \subset \{EY | Y \in \mathcal{S}_+,\ Y \leq X_2\}$$

נובע כי ה־sup על אוסף המספרים השמאלי קטן או שווה מהימני, דהיינו $EX_1 \leq EX_2$.
טענה: $X_n \leq X_{n+1}$ $0 \leq X_n$ לכל $n \geq 1$ ו־$X \uparrow X_n$, כאשר X_n לא בהכרח פשוטים. אז $\lim_{n\to\infty} EX_n = EX$.
הוכחה: למעשה, בעזרת הטענה האחרונה, ניתן להשתמש באותה הוכחה בה השתמשנו קודם כאשר X_nהיו פשוטים ולהראות כי גבול זה שווה ל־

$$\sup \{EY | 0 \leq Y \leq X\} = EX$$

בשביל הגיוון נראה הוכחה אחרת.
נניח כי לכל $n \geq 1$ $Y_{n,k} \in \mathcal{S}_+$ היא סדרה לא יורדת ב־k עם $Y_{n,k} \uparrow X_n$. אז $Z_k = \max_{1 \leq n \leq k} Y_{n,k}$ הוא משתנה מקרי פשוט ואי שלילי ומתקיים כי

$$Z_k = \max_{1 \leq n \leq k} Y_{n,k} \leq \max_{1 \leq n \leq k} Y_{n,k+1} \leq \max_{1 \leq n \leq k+1} Y_{n,k+1} = Z_{k+1}$$

מכיוון ש־

$$Z_k \geq Y_{n,k}\ \forall 1 \leq n \leq k$$

אז לכל $n \geq 1$‘

$$\lim_{k\to\infty} Z_k \geq \lim_{k\to\infty} Y_{n,k} = X_n$$

ומכאן שגם

$$\lim_{k\to\infty} Z_k \geq \lim_{n\to\infty} X_n = X$$

כמו כן מכיוון ש־$X_k \leq X_n \leq Y_{n,k}$ לכל $k \leq n$ אז גם

$$Z_k = \max_{1\leq n\leq k} Y_{n,k} \leq X_k$$

ולכן

$$\lim_{k\to\infty} Z_k \leq \lim_{k\to\infty} X_k = X$$

ולכן קבלנו ש־Z_k היא סדרה לא יורדת של משתנים מקריים פשוטים ואי שליליים השואפת ל־X. מהמשפט שהראינו קודם נובע כי

$$\lim_{n\to\infty} EZ_n = EX$$

מכיוון ש־$Z_n \leq X_n \leq X$ אז גם $EZ_n \leq EX_n \leq EX$ ולכן גם

$$\lim_{n\to\infty} EX_n = EX$$

טענה: אם X, Y משתנים מקריים אי שליליים (לא בהכרח סופיים או עם תוחלת סופית) אז

$$E(X+Y) = EX + EY$$

ניקח משתנים מקריים פשוטים X_n, Y_n כך ש־$X_n \uparrow X$ ו־$Y_n \uparrow Y$. אז גם $X_n + Y_n$ פשוטים ומתקיים כי $X_n + Y_n \uparrow X + Y$. לכן

$$\begin{aligned} E(X+Y) &= \lim_{n\to\infty} E(X_n+Y_n) = \lim_{n\to\infty} \{EX_n + EY_n\} \\ &= \lim_{n\to\infty} EX_n + \lim_{n\to\infty} EY_n = EX + EY \end{aligned}$$

נניח עכשיו כי X הוא משתנה מקרי אי שלילי המקיים $P(X=0)=1$ או באופן שקול, $P(X>0)=0$. משתנה כזה אינו בהכרח משתנה מקרי פשוט. עכשיו, לכל $Y \in \mathcal{S}_+$ המקיים $Y \leq X$ אנו יכולים לרשום

$$Y = Y1_{X^{-1}(\{0\})} + Y1_{X^{-1}((0,\infty])}$$

ומכיוון שלכל $\omega \in X^{-1}(\{0\})$ מתקיים כי $X(\omega)=0$ אז גם $Y(\omega)=0$ ולכן

$$Y = Y1_{X^{-1}((0,\infty])} = \sum_{i=1}^{n} a_i 1_{A_i \cap X^{-1}((0,\infty])}$$

עכשיו

$$P(A_i \cap X^{-1}((0,\infty])) \leq P(X^{-1}((0,\infty])) = P(X>0) = 1 - P(X=0) = 0$$

ולכן

$$EY = \sum_{i=1}^{n} a_i P(A_i \cap X^{-1}((0,\infty])) = 0$$

מכך נובע כי

$$EX = \sup\{EY | Y \in \mathcal{S}_+,\ Y \leq X\} = 0$$

כלומר, אם $X \geq 0$ ומתקיים כי $P(X = 0) = 1$ אז $EX = 0$.
נסמן $a^+ = \max(a, 0)$ ו־$a^- = -\min(a, 0) = (-a)^+$ ועבור משתנה מקרי כלשהו נגדיר

$$EX = EX^+ - EX^-$$

כאשר $EX^+ < \infty$ או $EX^- < \infty$. אם שניהם אינסופיים, נאמר כי התוחלת לא מוגדרת.
אם למשל $EX^+ = \infty > EX^-$ אז נאמר כי התוחלת מוגדרת אבל שווה ל־∞. אם $EX^+ < \infty = EX^-$ אז נאמר כי התוחלת מוגדרת ושווה ל־$-\infty$. כאשר שניהם סופיים אז נאמר כי התוחלת מוגדרת וסופית.
עכשיו, נשים לב כי אם $P(X \geq 0) = 1$ אז $P(X^- > 0) = P(X < 0) = 0$ ולכן $EX^- = 0$. במקרה זה נקבל כי $EX = EX^+$. באותו אופן, אם $P(X \leq 0) = 1$ אז $EX = -EX^-$. אם $P(X = 0) = 1$ אז $P(X > 0) = P(X < 0) = 0$ ולכן גם $P(X^+ = 0) = P(X^- = 0) = 1$ ומכאן ש־

$$EX = EX^+ - EX^- = 0 - 0 = 0$$

בבית אתם תוכיחו את התכונות הבאות של התוחלת:

1. אם לא קיים ω עבורו $X(\omega) = \infty$ ו־$Y(\omega) = -\infty$ או $X(\omega) = -\infty$ וגם $Y(\omega) = \infty$ אז $X + Y$ מוגדר היטב (יכול להיות אינסופי). אם בנוסף EX^+, EY^+ הם סופיים או EX^-, EY^- הם סופיים, אז

$$E(X + Y) = EX + EY$$

בדרך כלל לא נהוג לרשום את התנאי הראשון ש־X, Y לא יכולים לקבל את ערכים הפוכים של אינסוף. במקרה כזה, תחת התנאי על התוחלות מתקיים כי $X + Y$ מוגדר בהסתברות אחת וגם אז רושמים $E(X+Y) = EX + EY$. הכוונה בכתיבה כזו היא שכאשר $X + Y$ לא מוגדר, מגדירים אותו באופן מלאכותי להיות איזה ערך שרוצים (או אפילו משתנה מקרי סופי כלשהו). למשל

$$Z(\omega) = \begin{cases} 0 & (X(\omega), Y(\omega)) \in \{(\infty, -\infty), (-\infty, \infty)\} \\ X(\omega) + Y(\omega) & \text{אחרת} \end{cases}$$

אנו לא נבחין בין Z לבין $X + Y$ ולא נתרגש מכך ש־$X + Y$ אינו מוגדר על מאורע שהסתברותו היא אפס.

2. אם התוחלת של X מוגדרת אז לכל קבוע סופי c מתקיים כי $E(cX) = cEX$.

3. אם $X \leq Y$ בהסתברות אחת ולפחות אחת מההנחות הבאות מתקיימת אז $EX \leq EY$

   (א) $EY^+ < \infty$

   (ב) $EX^- < \infty$

(ג) $EY^- < \infty$ וגם $EX^+ < \infty$

4. אם $P(X \geq 0) = 1$ אז $EX = 0$ אם ורק אם $P(X = 0) = 1$.

5. אם $E|X| < \infty$ אז EX מוגדר וסופי ומתקיים כי $|EX| \leq E|X|$.

6. אם $EX^2 < \infty$ אז גם $E|X| < \infty$. אם בנוסף מתקיים כי $\mathrm{Var}(X) = 0$ אז $P(X = EX) = 1$.

כאן אנחנו מוכיחים את משפט ההכנסות המונוטונית

עכשיו, נניח כי $P(X_n \geq 0) = 1$ לכל $n \geq 1$ וכי $P(X_n \leq X_{n+1}) = 1$ לכל $n \geq 1$. נסמן

$$A = \left(\bigcap_{n=1}^{\infty} \{X_n \geq 0\}\right) \cap \left(\bigcap_{n=1}^{\infty} \{X_n \leq X_{n+1}\}\right)$$

מכיוון ש־A הוא חיתוך בן מניה של מאורעות שהסתברותם אחת אז גם $P(A) = 1$. הסיבה ש־$\{X_n \leq X_{n+1}\} = \{X_{n+1} - X_n \geq 0\}$ הוא מאורע היא כי $X_{n+1} - X_n$ הוא גם משתנה מקרי. עכשיו ברור כי

$$X_n 1_A \geq 0,\ X_n 1_A \leq X_{n+1} 1_A$$

נסמן

$$Z = \lim_{n\to\infty} X_n 1_A$$

הגבול קיים ומתקיים גם כי $Z(\omega) = 0$ לכל $\omega \notin A$ ו־$Z(\omega) \geq 0$ לכל $\omega \notin A$. נניח כי $P(X = Z) = 1$ ונסמן $B = A \cap \{X = Z\}$. אז $P(B) = 1$ ומתקיים כי

$$\lim_{n\to\infty} X_n 1_B = X 1_B$$

מכיוון ש־$X_n 1_{B^c}$ ו־$X 1_{B^c}$ הם אפס בהסתברות אחת, אז גם $EX_n 1_{A^c} = EX 1_{A^c} = 0$. קבלנו אם כן כי $0 \leq X_n \uparrow X$ על B כאשר $P(B) = 1$. כמו כן

$$EX_n^+ 1_{B^c} = EX_n^- 1_{B^c} = 0$$

ולכן גם

$$EX_n 1_{B^c} = 0$$

ובאופן דומה עבור X. מכאן ש־

$$\begin{aligned}
\lim_{n\to\infty} EX_n &= \lim_{n\to\infty} (EX_n 1_B + EX_n 1_{B^c}) = \lim_{n\to\infty} EX_n 1_B \\
&= EX 1_B = EX 1_B + EX 1_{B^c} = EX
\end{aligned}$$

וקבלנו את המשפט הבא (משפט ההתכנסות המונוטונית):

נניח כי לכל $n \geq$ מתקיים כי $P(X_n \leq X_{n+1}) = P(X_n \geq 0) = 1$. אז קיים משתנה מקרי X עם $P(X \geq 0) = 1$ ו־

$$P\left(\exists \lim_{n\to\infty} X_n = X\right) = 1$$

ומתקיים כי

$$\lim_{n\to\infty} EX_n = EX$$

(EX_n, EX יכולים להיות סופיים או אינסופיים).

משפט התכנסות מונוטונית אלטרנטיבי:

נניח כי $P(X_n \geq 0) = 1$. אז קיים משתנה מקרי השווה בהסתברות אחת ל־

$$\sum_{n=1}^{\infty} X_n$$

ומתקיים כי

$$E\sum_{n=1}^{\infty} X_n = \sum_{n=1}^{\infty} EX_n$$

הוכחה: ניקח $Y_m = \sum_{n=1}^{m} X_n$ ואז Y_m מקיימים את תנאי המשפט הקודם עם $Y = \sum_{n=1}^{\infty} X_n$ (בהסתברות אחת).

הלמה של Fatou:

נניח כי $X_n \geq 0$ בהסתברות אחת לכל $n \geq 0$. אז

$$E\liminf_{n\to\infty} X_n \leq \liminf_{n\to\infty} EX_n$$

הוכחה: כפי שהראינו עבור משפט ההתכנסות המונוטונית, מספיק להראות זאת עבור המקרה שבו $X_n \geq 0$ (לכל ω ולא רק בהסתברות אחת). לכל $n \geq 1$ מתקיים כי

$$\inf_{m\geq n} X_m \leq X_k$$

לכל $k \geq n$ ולכן גם

$$E\inf_{m\geq n} X_m \leq EX_k$$

מכאן שגם

$$E\inf_{m\geq n} X_m \leq \inf_{k\geq n} EX_m \leq \sup_{n\geq 1}\inf_{k\geq n} EX_m = \liminf_{n\to\infty} EX_n$$

ולכן

$$\sup_{n\geq 1} E\inf_{m\geq n} X_m \leq \liminf_{n\to\infty} EX_n$$

מכיוון ש־$E\inf_{m\geq n} X_m$ היא סדרה לא יורדת נובע כי

$$\sup_{n\geq 1} E\inf_{m\geq n} X_m = \lim_{n\to\infty} E\inf_{m\geq n} X_m$$

וממשפט ההתכנסות המונוטונית נובע כי

$$\lim_{n\to\infty} E\inf_{m\geq n} X_m = E\lim_{n\to\infty}\inf_{m\geq n} X_m = E\liminf_{n\to\infty} X_n$$

ומכאן נובעת הטענה.

משפט ההתכנסות הנשלטת (dominated convergence theorem):
נניח כי X_n, X הם משתנים מקריים ו־$X_n \to X$ בהסתברות אחת. נניח בנוסף כי קיים משתנה מקרי אי שלילי Y עם $EY < \infty$ כך ש־$P(|X_n| \le Y) = 1$ לכל $n \ge 1$. אז

$$\lim_{n\to\infty} EX_n = EX$$

אם Y הוא קבוע, אז משפט זה נקרא משפט ההתכנסות החסומה.

הוכחה: מכיוון ש־$Y - X_n$ ו־$Y + X_n$ הם אי שליליים בהסתברות אחת, אז מהלמה של Fatou ומכך ש־$EY < \infty$ נובע כי

$$\begin{aligned} EY - E\limsup_{n\to\infty} X_n &= E\liminf_{n\to\infty}(Y - X_n) \le \liminf_{n\to\infty} E(Y - X_n) = EY - \limsup_{n\to\infty} EX_n \\ EY + E\liminf_{n\to\infty} X_n &= E\liminf_{n\to\infty}(Y + X_n) \le \liminf_{n\to\infty} E(Y + X_n) = EY + \liminf_{n\to\infty} EX_n \end{aligned}$$

עכשיו בהסתברות אחת מתקיים כי

$$X = \limsup_{n\to\infty} X_n = \liminf_{n\to\infty} X_n$$

ולכן

$$\begin{aligned} EY - EX &\le EY - \limsup_{n\to\infty} EX_n \Rightarrow \limsup_{n\to\infty} EX_n \le EX \\ EY + EX &\le EY + \liminf_{n\to\infty} EX_n \Rightarrow \liminf_{n\to\infty} EX_n \ge EX \end{aligned}$$

מכך נובע כי $\liminf_{n\to\infty} EX_n = \limsup_{n\to\infty} EX_n = EX$ ולכן קיים הגבול $\lim_{n\to\infty} EX_n$ והוא שווה ל־EX.

נניח עכשיו כי $EX^2 < \infty$ וגם $EY^2 < \infty$. אז X, Y סופיים בהסתברות אחת ולכן XY מוגדר וסופי בהסתברות אחת. הבעיה היחידה שיכולה להיות בהגדרת XY היא כאשר אחד מהמשתנים הוא אפס והשני הוא פלוס או מינוס אינסוף. מכיוון שזה קורה בהסתברות אפס אז לא נבחין בין XY לבין משתנה מקרי Z שמקבל את הערך XY כשזה מוגדר היטב ואיזשהו ערך שרירותי (למשל אפס) כאשר זה לא מוגדר. אם כן לכל a, b (סופיים) מתקיים כי

$$0 \le (|a| - |b|)^2 = a^2 + b^2 - 2|ab|$$

ולכן

$$|ab| \le \frac{a^2 + b^2}{2}$$

ואז מתקיים גם כי

$$|XY| \le \frac{X^2 + Y^2}{2}$$

(לכל ω עבורו צד שמאל מוגדר). ומכאן

$$E|XY| \le \frac{EX^2 + EY^2}{2} < \infty$$

שימו לב כי אם נחליף את Y ב־1 נקבל כי $E|X| < \infty$ ואם נחליף את X ב־1 נקבל $E|Y| < \infty$. זה לא כל כך חשוב לנו כאן, אך טוב לשים לב כי כאשר המומנט השני סופי התוחלת מוגדרת היטב וסופית.

עכשיו, נניח כי $P(X = 0) < 1$. זה שקול לכך ש־$EX^2 > 0$ (הנחנו קודם שערך זה הוא סופי). נבצע עכשיו את הפיתוח הבא

$$
\begin{aligned}
0 &\leq E(Y - tX)^2 = E(Y^2 - 2tXY + t^2X^2) \\
&= EY^2 - 2tEXY + t^2EX^2 \\
&= EX^2(t^2 - 2tEXY) + Y^2 \\
&= EX^2\left(t^2 - 2t\frac{EXY}{EX^2} + \left(\frac{EXY}{EX^2}\right)^2 - \left(\frac{EXY}{EX^2}\right)^2\right) + Y^2 \\
&= EX^2\left(t - \frac{EXY}{EX^2}\right)^2 + \frac{EX^2EY^2 - (EXY)^2}{EX^2}
\end{aligned}
$$

מכיוון שזה מתקיים לכל $t \in \mathbb{R}$ אפשר להציב $t = \frac{EXY}{EX^2}$ ולקבל כי

$$0 \leq E\left(Y - \frac{EXY}{EX^2}X\right)^2 = \frac{EX^2EY^2 - (EXY)^2}{EX^2}$$

מכך נובעות שתי מסקנות. הראשונה היא שצד ימין הוא אי שלילי וזה שקול ל־

$$|EXY| \leq \sqrt{EX^2EY^2}$$

שימו לב כי אם $P(X = 0) = 1$ אז $EXY = 0$ וגם $EX^2 = 0$ ומכאן שבשני הצדדים כתוב אפס. מכאן שאי השוויון הזה מתקיים לכל X, Y עבורם $EX^2 < \infty$ וגם $EY^2 < \infty$. אי שוויון זה הוא מקרה פרטי של "אי שוויון קושי־שוורץ".

המסקנה השניה היא שאם

$$|EXY| = \sqrt{EX^2EY^2}$$

אז

$$E\left(Y - \frac{EXY}{EX^2}X\right)^2 = 0$$

ולכן

$$P\left(Y = \frac{EXY}{EX^2}X\right) = 1$$

אם נסמן $a = -\frac{EXY}{EX^2}$ ו־$b = 1$, אז (a, b) הם לא שניהם אפס ומתקיים כי

$$P(aX + bY = 0) = 1$$

כלומר, X, Y תלויים לינארית בהסתברות אחת. כאשר $P(X = 0) = 1$ אז אפשר לקחת $a = 1$ ו־$b = 0$ ושוב לקבל כי

$$P(aX + bY = 0) = 1$$

כאשר (a,b) אינם שניהם אפס.

המסקנה היא שאם אי שוויון קושי שוורץ מתקיים עם שוויון אז X,Y תלויים לינארית בהסתברות אחת. מתברר שגם ההיפך נכון, דהיינו, אם הם תלויים לינארית בהסתברות אחת אז אי שוויון קושי שוורץ מתקיים עם שוויון. כדי לראות זאת, נניח כי $aX+bY=0$ בהסתברות אחת. נניח בלי הגבלת הכלליות כי $b\neq 0$ (אחרת $a\neq 0$ ואז עושים את אותו הדבר כאשר הופכים את התפקידים בין X ו־Y). נסמן $t=-a/b$ ונקבל כי $Y=tX$ בהסתברות אחת. אז

$$\begin{aligned} |EXY| &= |EX(tX)| = |t|EX^2 \\ \sqrt{EX^2EY^2} &= \sqrt{EX^2E(tX)^2} = |t|EX^2 \end{aligned}$$

ומכאן שמתקיים שוויון.

אם במקום X,Y נציב $X-EX$ ו־$Y-EY$ נקבל כי

$$|E(X-EX)(Y-EY)| \le \sqrt{E(X-EX)^2E(Y-EY)^2}$$

כלומר

$$|\mathrm{Cov}(X,Y)| \le \sigma(X)\sigma(Y)$$

כאשר שויון מתקיים אם ורק אם קיימים (a,b) שאינם שניהם אפס כך ש־$a(X-EX)+b(Y-EY)=0$ בהסתברות אחת. זה שקול לכך שקיימים (a,b,c) כאשר (a,b) אינם שניהם אפס, כך ש־$aX+bY=c$ בהסתברות אחת. הסיבה שזה שקול היא שאם ניקח תוחלת אז בהכרח נקבל כי $aEX+bEY=c$ ואז נקבל כי $aX+bY=aEX+bEY$ וזה שקול ל־$a(X-EX)+b(Y-EY)=0$ (בהסתברות אחת).

אם $\sigma(X)>0$ ו־$\sigma(Y)>0$ (X,Y אינם קבועים בהסתברות אחת) אז קיבלנו כי מקדם המתאם

$$\rho(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}$$

מקיים

$$-1\le\rho(X,Y)\le 1$$

עם שוויון ל־1 או -1 אם ורק אם קיימים (a,b,c) כך ש־a,b אינם שניהם אפס כך ש־$aX+bY=c$ בהסתברות אחת.

למעשה, לומר כי (a,b) אינם שניהם אפס שקול לכך ש־(a,b,c) לא כולם אפס. הסיבה היא שאם $c=0$ אז בהכרח a,b אינם שניהם אפס (אחרת הם כולם אפס). אם $c\neq 0$ אז לא יתכן כי $a=b=0$ כי אז נקבל $c=0$.

לסיכום:

אי שוויון קושי שוורץ

נניח כי X,Y הם משתנים מקרים המקיימים $EX^2<\infty$ וגם $EY^2<\infty$. אז EXY מוגדר וסופי ומתקיים כי

$$|EXY|\le\sqrt{EX^2EY^2}$$

עם שוויון אם ורק אם X,Y תלויים לינארית בהסתברות אחת.

כמו כן מתקיים כי

$$|\mathrm{Cov}(X,Y)| \le \sigma(X)\sigma(Y)$$

עם שוויון אם ורק אם $X,Y,1$ תלויים לינארית בהסתברות אחת. דהיינו קיימים (a,b,c) לא כולם אפס כך ש־$aX+bY+(-c)1=0$.
כאשר X,Y אינם קבועים בהסתברות אחת אי השוויון האחרון שקול ל־$-1 \le \rho(X,Y) \le 1$.

נסיים בכך שנשים לב כי אוסף הנקודות (x,y) המקיימות $ax+by=c$ הוא קו ישר שאינו בהכרח עובר דרך הראשית.