# תורת ההסתברות I 80420

אור דגמי, or@digmi.org

30 בינואר 2012

אתר אינטרנט: http://digmi.org

**תקציר**

סיכום ההרצאות של ד"ר מיכאל הוכמן משיעורי תורת ההסתברות I באוניברסיטה העברית בירושלים בשנת לימודים 2012.

# תוכן עניינים


1 מבוא 4
1.1 מה זה הסתברות? 4
1.2 ספרות 4
2 מרחב מדגם 4
2.1 מאורעות 5
2.1.1 אינטואיציה 5
2.1.2 למה לא כל תת קבוצה היא מאורע? 6
2.2 מרחב הסתברות 9
2.3 מרחבי מידה סופיים 11
2.4 מרחבי מדגם לא אחידים 14
2.5 מרחבי מידה אינסופיים 17
2.5.1 $\limsup / \liminf$ 18
3 ניסויים תלויים 27
3.1 בניית מרחבי הסתברות לסדרת ניסויים 28
4 אי תלות 32
5 סדרות אינסופיות של מאורעות 39
5.1 בורל־קנטלי 41
6 משתנים מקריים 47
6.1 משתנה ברנולי 50
6.2 משתנה בינומי 51
6.3 משתנה פואסון 53
6.4 משתנה גיאומטרי 54
6.5 תכונת חוסר זיכרון 55
6.6 אי תלות 59
6.6.1 אי תלות של סדרות של משתנים מקריים 61
6.7 פונצקיות של משתנים מקריים 62
6.8 התפלגויות מותנות 66
7 תוחלת 67
7.1 תוחלת של התפלגויות קלאסיות 69
7.1.1 משתנה ברנולי 69
7.1.2 משתנה בינומי 69
7.1.3 משתנה פואסוני 70
7.1.4 משתנה גיאומטרי 70
7.2 תוחלת של פונקציות על משתנים מקריים 70
7.3 תכונות של תוחלת 71
7.4 שונות 75
7.4.1 תכונות של שונות 75
7.5 תוחלת מותנית 80
8 משתנים מקריים "רציפים" (לא בדידים) 82
8.1 פונקציות צפיפות (לא כללי) 84
8.1.1 הקשר בין $f_X$ ו $F_X$ 87
9 תוחלת של משתנה רציף 87
9.1 שונות 92
9.2 התפלגות משותפת של משתנים מקריים ממשיים 94
9.3 אי תלות 98
9.4 צפיפות מותנית 103
10 התכנסות של סדרות של משתנים מקריים ממשיים 104
10.1 חוק המספרים הגדולים 106
10.1.1 תאור של הסתברות אחידה על $[0,1]$ 110
10.2 התכנסות בהתפלגות 113

תוכן עניינים

11 משפט הגבול המרכזי - Central Limit Theorem . . . . . . . . . . . . . . . . . . . . . . . . . . . 118
11.1 ecalpaL-ervioM ed . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 121

30/10/20

# 1 מבוא

## 1.1 מה זה הסתברות?

הסתברות היה מקצוע מזולזל במשך שנים, וקיבל push במלחמת העולם השניה עקב חקר ביצועים וכו'.
אז מה זה הסתברות?
לפי Kolmogorov אפשר להתייחס להסתברות כענף בתורת המידה.

## 1.2 ספרות

אין ספר שמלווה את הקורס, אבל בקירוב טוב נסתמך על הרשימות של: Raz Kupferman

# 2 מרחב מדגם

**הגדרה 2.1 מרחב מדגם:** מרחב מדגם הוא קבוצה אשר נסמנה ב-$\Omega$. כל נקודה $\omega \in \Omega$ מייצגת מצב של העולם (ניסוי, התרחשות מקרית, וכו')

**דוגמה 2.2 הטלת מטבע:** מטילים מטבע ויוצא עץ או פלי. נסמן: $H = \text{head}$ ו: $T = \text{tail}$.
מרחב המדגם שלנו הוא:

$$\Omega = \{H, T\}$$

**דוגמה 2.3 הטלה מרובה:** מטילים מטבע שלוש פעמים.

$$\Omega = \{\text{קבוצות באורך 3 של } H,T\} = \{(x_1, x_2, x_3)\ x_i \in \{H, T\}\} = \{H, T\}^3 =$$
$$\{(H,H,H), (H,H,T), (H,T,H), (H,T,T), (T,H,T), (T,T,H), (T,T,T)\}$$

**דוגמה 2.4 כמות אינסופית של הטלות:** מטילים מטבע שוב ושוב ($\infty$ פעמים):

$$\Omega = \{\text{סדרות אינסופיות } (x_i)_i^\infty \text{ של } H,T\} = \{H, T\}^N$$

**הערה 2.5** $\{f : y \to x\} = x^y$

$$\{H, T\}^N = \{x : N \to \{H, T\}\}$$

(אוסף הסדרות)

**דוגמה 2.6** מטילים מטבע אם יוצא $H$ אז מטילים עוד מטבע, אם יוצא $T$ מטילים קוביה.

$$\Omega = \left\{(x, y)\ x \in \{H, T\},\ y \in \begin{cases} \{H, T\} & x = H \\ \{1, \ldots, 6\} & x = T \end{cases}\right\}$$

**דוגמה 2.7 לוטו:** מנחשים 7 מספרים. ה6 הראשונים מייצגים בחירה של 6 כדורים מתוך אוסף של 37 כדורים. ה7 נבחר מאוסף אחר של 7 כדורים.

$$\Omega = \{(x_1,\dots,x_7)\ (x_1,\dots,x_6)\ \text{סדרה של מספרים מ1 עד 37 בלי חזרות},\ x_7\in\{1,\dots,7\}\}$$

**דוגמה 2.8 ביקוע גרעיני:** בזמן 0 נתון חלקיק רדיואקטיבי המתפרק בזמן מקרי $T$.

$$\Omega = [0,\infty)$$

**דוגמה 2.9 קוביה "פיזית":** מטילים קוביה "פיזית" מודל אפשרי:

$$\Omega = \left\{(p,u,v,w)\ \ p,u,v,w\in\mathbb{R}^3\right\}$$

כאשר $u,v,w$ וקטורי יחידה ניצבים ונמצא בשיווי משקל. ו $p$ מייצג את מרחק הקובייה מהראשית

**הערה 2.10** גם מצבים כמו קוביה הנופלת על הפינה הם אפשריים, ובמקרה של קובייה סטנדרטית $\Omega=\{1,\dots,6\}$ זה לא מיוצג.

**הערה 2.11** במודל של מרחב מדגם אין פרמטר זמן או מרחב.
כלומר, במקרה שמטילים מטבעות 3 פעמים אמרנו כי $\Omega=\{H,T\}^3$. כל $\omega\in\Omega$ מייצג 3 הטלות בזמנים שונים (זהו פירוש אבל אין זה חלק מההגדרה).

## 2.1 מאורעות

### 2.1.1 אינטואיציה

**הגדרה 2.12 מאורע:** (הגדרה לא פורמלית)
המאורע שבהטלה 3 מטבעות יצאו פעמיים ראש. המאורע בבדוגמה 6 החלקיק מפתרק לפני זמן 1.

מאורעות מעין אלו מתאימים לתת קבוצות של מרחב המדגם.
בדוגמה 2.3: $\{(H,H,T),(H,T,H),(T,H,H)\}\subseteq\{H,T\}^3=\Omega$
בדוגמה 2.8: $[0,1)\subseteq[0,\infty)=\Omega$
מאורע מסויים מתאים ל:

$$\{\omega\in\Omega\ \text{מייצג מצב עולם שבו המאורע קרה}\ \omega\}\ \Omega$$

**דוגמה 2.13** המאורע שהטלנו שתי קוביות וסכום התוצאות זוגי:

$$\Omega=\{1,\dots,6\}^2=\{(x_1,x_2)\ \ x_1\in\{1,\dots,6\}\}$$

המאורע:

$$A=\{(x_1,x_2)\in\Omega\quad x_1+x_2\ \text{זוגי}\}=$$

קבוצות כל המצבים שבהם סכום התוצאות זוגי

המאורע $\Omega$ $\Longleftrightarrow$ ייצא תוצאה כלשהיא בניסוי.
כלומר: $\omega\in\Omega$ המאורע שעולם $\{\omega\}$ נמצא במצב $\omega$.
אם $\omega_1,\omega_2\in\Omega$ אז $\{\omega_1,\omega_2\}$ הוא המאורע שהעולם נמצא באחד משני המצבים $\omega_1,\omega_2$.
ולסיכום: בהנתן $A\subseteq\Omega$ אפשר לדבר על המאורע שהעולם נמצא במצב כלשהו ב$A$.
בהנתן מאורעות $A,B\in\Omega$ אזי $A\cup B$ הוא המאורע שאו $A$ או $B$ קרה. כי לומר שהעולם נמצא במצב השייך ל $A\cup B$ $\Longleftrightarrow$ במצב השייך ל$A$ או במצב ששייך ל$B$.

**דוגמה 2.14** מטילים מטבע 3 פעמים.

$$A = \{\text{יצא } H \text{ בהטלה הראשונה}\} = \{(H, ?, ?)\}$$

$$B = \{\text{יצא } T \text{ בהטלה השניה}\} = \{(?, T, ?)\}$$

אז:

$$A \cup B = \{\text{יצא } H \text{ בהטלה הראשונה או יצא } T \text{ בהטלה השניה}\}$$

$$A \cap B = \{\text{יצא } H \text{ בהטלה הראשונה ויצא } T \text{ בהטלה השניה}\} = \{(H, T, ?)\}$$

אם $A \subseteq \Omega$ אז:

$$A^c = \Omega \backslash A = \{\text{העולם לא נמצא במצב השייך ל} A\} = \text{"} A \text{ לא קרה"}$$

בדוגמה הקודמת:

$$A^c = \{\text{לא יצא } H \text{ בהטלה הראשונה}\} = \{(T, ?, ?)\}$$

**הגדרה 2.15 מאורע:** (חצי פורמלית)
תת־קבוצה $A \subseteq \Omega$ ניתן לפירוש כארוע נאמר ש $A$ התרחש אם מצב העולם שייך ל$A$ (אז נאמר שהעולם נמצא במצב ב $A$)

ראינו שפעולות קבוצתיות חופפות לפעולות לוגיות על "מאורעות".
נניח שנתונים מאורעות

$$A_1, A_2, \ldots \subseteq \Omega$$

המאורע שלפחות אחד מהמאורעות האלה קרה?

$$\bigcup_{n=1}^{\infty} A_n$$

המאורע שכולם קרו הוא כמובן:

$$\bigcap_{n=1}^{\infty} A_n$$

### 2.1.2 למה לא כל תת קבוצה היא מאורע?

בגלל דברים שנלמד בתורת המידה. כאשר אנו מדברים על קבוצות סופיות אין בעיה, אבל במידה ואנחנו מדברים על קבוצות אינסופיות דברים משתבשים.

**הגדרה 2.16 $\sigma$-אלגברה:** תהי $\Omega$ קבוצה. משפחה (קבוצה של תת קבוצות) $\mathcal{F}$ של תת קבוצות של $\Omega$ נקראת $\sigma$-אלגברה על $\Omega$ אם:

1. $\Omega \in \mathcal{F}$
2. אם $A \in \mathcal{F}$ אז:

$$A^c = \Omega \backslash A \in \mathcal{F}$$

3. אם $A_1, A_2, A_3, \ldots \in \mathcal{F}$ אז:

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$$

(חשוב מאוד שזו תהא סדרה בת מנייה!)

**דוגמה 2.17** אם $\Omega$ קבוצה אז אוסף (קבוצת) כל התת קבוצות של $\Omega$ (נסמן ב: $2^\Omega$) היא $\sigma$-אלגברה על $\Omega$:

$$2^\Omega = \{A,\ A \subseteq \Omega\}$$

למה זה $\sigma$-אלגברה? כיוון ש:

1. $\Omega \subseteq \Omega$ אזי: $\Omega \in 2^\Omega$
2. אם $A \in 2^\Omega$ אז: $A \subseteq \Omega$ לכן: $\Omega \backslash A \subseteq \Omega$ ולכן: $\Omega \backslash A \in 2^\Omega$
3. כל איחוד של תת קבוצה של $\Omega$ הוא תת קבוצה של $\Omega$.

**דוגמה 2.18** בהנתן $\Omega$, $\mathcal{F} = \{\varnothing, \Omega\}$ זו $\sigma$-אלגברה כי:

1. $\Omega \in \mathcal{F}$
2. אם $A \in \mathcal{F}$ אז או: $A = \Omega$ ואז:

$$A^c = \varnothing \in \mathcal{F}$$

או $A = \varnothing$ ואז:

$$A^c = \Omega \in \mathcal{F}$$

3. אם $A_n \in \mathcal{F}$ אז או כל $A_n = \varnothing$ ואז: $\bigcup_{n=1}^{\infty} A_n = \varnothing \in \mathcal{F}$ או יש $n$ כך ש $A_n = \Omega$ ואז: $\bigcup_{n=1}^{\infty} A_n = \Omega \in \mathcal{F}$

**דוגמה 2.19** אם $\Omega$ קבוצה ו $A \subseteq \Omega$ אז:

$$\mathcal{F} = \{\varnothing, A, A^c, \Omega\}$$

זו $\sigma$-אלגברה

**למה 2.20**

אם $\mathcal{F}$ $\sigma$-אלגברה על $\Omega$ אז:

1. $\varnothing \in \mathcal{F}$
2. אם $A_1,\dots,A_n \in \mathcal{F}$ אז $\bigcup\limits_{n=1}^{N} A_n \in \mathcal{F}$
3. אם $A_1,A_2,\dots \in \mathcal{F}$ אז: $\bigcap\limits_{n=1}^{\infty} A_n \in \mathcal{F}$ וגם:

$$\forall N \in \mathbb{N} \ \bigcap_{n=1}^{N} A_n \in \mathcal{F}$$

4. אם $A,B \in \mathcal{F}$ אז:

$$B \backslash A \in \mathcal{F}$$

**הוכחה:**

1. $\Omega \in \mathcal{F}$ (לפי תכונה ראשונה של $\sigma$-אלגברה) לכן:

$$\varnothing = \Omega^c \in \mathcal{F}$$

לפי תכונה שניה של $\sigma$-אלגברה.

2. נתון $A_1,\dots,A_N$. עבור $n > N$ נגדיר:

$$A_n = \varnothing \in \mathcal{F}$$

אז:

$$\bigcup_{n=1}^{N} A_n = \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$$

לפי תכונה שלישית של $\sigma$-אלגברה.

3. נשים לב כי:

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c$$

זו זהות מתורת הקבוצות. דוגמה קלה לחשוב עליה היאף $(B \cup C)^c = B^c \cap C^c$.
ניתון $A_n \in \mathcal{F}$, אנו יודעים כי $A_n^c \in \mathcal{F}$ וכמו כן כי: $\bigcup\limits_{n=1}^{\infty} A_n^c \in \mathcal{F}$ ולכן נקבל מהתכונה הראשונה שוב כי:

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \in \mathcal{F}$$

לגבי החיתוכים סופיים נובע באופן דומה מסעיף 2 של הלמה.

4. נשים לב ש: $B \backslash A = B \cap (A^c)$ נתון $A,B \in \mathcal{F}$ ולכן: $A^c \in \mathcal{F}$ ומסעיף 3: $B \cap A^c \in \mathcal{F}$

■

## 2.2 מרחב הסתברות

**הגדרה 2.21 מרחב מדיד:** מרחב מדיד הוא זוג $(\Omega,\mathcal{F})$ כאשר $\Omega$ קבוצה, $\mathcal{F}$ $\sigma$־אלגברה על $\Omega$. הקבוצה ב$\mathcal{F}$ נקראות מאורעות.

**הגדרה 2.22** סדרת קבוצות $A_1,A_2,\ldots$ נקראת זרה אם: $A_i\cap A_j=\varnothing$ לכל $i\neq j$

**הגדרה 2.23** מרחב הסתברות הוא שלשה $(\Omega,\mathcal{F},P)$כאשר $(\Omega,\mathcal{F})$מרחב מדיד ו: $P:\mathcal{F}\to\mathbb{R}$ היא פונקצית הסתברות (מידת הסתברות) כלומר $P$ מקיימת:

1. $0\leq P\left(A\right)\leq 1$ ($A$ מאורע ב$\mathcal{F}$)
2. $P\left(\Omega\right)=1$
3. אם $A_1,A_2,\ldots\in\mathcal{F}$ סדרה זרה אז:

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

(טור חיובי המתכנס במובן רחב והסדר לא משנה)

$P\left(A\right)$ נקרא הסתברות של $A$.

**למה 2.24**

אם $(\Omega,\mathcal{F},P)$ מרחב הסתברות אז:
1. $P\left(\varnothing\right)=0$
2. אם $A_1,\ldots,A_n\in\mathcal{F}$זרות אז: $P\left(\bigcup_{n=1}^{N}A_n\right)=\sum_{n=1}^{N}P\left(A_n\right)$
3. $P\left(A^c\right)=1-P\left(A\right)$ עבור $A\in\mathcal{F}$
4. אם $A\subseteq B$ אז: $P\left(A\right)\leq P\left(B\right)$עבור $A,B\in\mathcal{F}$

**הוכחה:**

1. נגדיר $A_n=\varnothing$ לכל $n$. כלומר: $\bigcup_{n=1}^{\infty}A_n=\varnothing$ וגם: $A_i\cap A_j=\varnothing$ לכל $i,j$ אז זו סדרה זרה. לכן:

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

לכן: $P\left(\varnothing\right)=0$.

2. בהנתן סדרה סופית זרה של $A_1,\ldots,A_n\in\mathcal{F}$ נגדיר $A_n=\varnothing$ עבור $n>N$ אז אפשר לבדוקק שהסדרה עדין זרה (כתרגיל) אז:

$$\bigcup_{n=1}^{N}A_n=\bigcup_{n=1}^{\infty}A_n$$

ולכן:

$$P\left(\bigcup_{n=1}^{N}A_n\right)=P\left(\bigcup_{n=1}^{\infty}A_n\right)\underset{\text{סדרה זרה}}{=}\sum_{n=1}^{\infty}P\left(A_n\right)=\sum_{n=1}^{N}P\left(A_n\right)$$

3. בהנתן $A\in\mathcal{F}$: $\Omega=A\cup A^{c}$ והזוג: $A,A^{c}$ זר. לכן לפי סעיף 2:

$$1=P\left(\Omega\right)=P\left(A\right)+P\left(A^{c}\right)$$

נעביר אגפים ונקבל:

$$P\left(A^{c}\right)=1-P\left(A\right)$$

4. נתון $A\subseteq B$ לכן:

$$B=A\cup B\backslash A$$

הזוג $A,B\backslash A$ זר לכן:

$$P\left(B\right)=P\left(A\right)+P\left(B\backslash A\right)\geq P\left(A\right)\geq0$$

■

**הערה 2.25** כדאי לחשוב על $\Omega$ כקבוצה משטח 1, ועל $P\left(A\right)$ בתור השטח של $A$ אז התכונה השלישית של $P$ אומרת ששטח האיחוד = איחוד השטחים.

31/10/20

**דוגמה 2.26** דיסק $D$ בשטח 1 משמש כמטרה, פוגעים בנקודה אקראית. מרחב המדגם:

$$\Omega=\left\{ \left(x,y\right)\in D\right\}$$

נרצה להגדיר פונקצית הסתברות $P$ על $\sigma$־אלגברה גדולה כך ששטח של $A=P\left(A\right)$.
נרצה ש$P$ תהא מוגדרת על $\sigma$־אלגברה $\mathcal{F}$ המכילה את כל המלבנים: $\left[a,b\right]\times\left[c,d\right]\subseteq D$

**משפט 2.27**

קיימת $\sigma$־אלגברה $\mathcal{F}$ כזאת ופונקצית ההסתברות $P$ עליה כך ש:

$$P\left(\left[a,b\right]\times\left[c,d\right]\right)=\text{שטח}=\left(b-a\right)\cdot\left(d-c\right)$$

**הערה 2.28** משפט זה לומדים בתורת המידה.
$\mathcal{F}$ נקראת $\sigma$־אלגברה בורל (Borel) והיא ה$\sigma$־אלגברה הקטנה ביותר המכילה את כל המלבנים:

$$\mathcal{F}=\bigcap\left\{ \mathcal{G}:\text{המכילה את כל המלבנים } D \text{ אלגברה על } - \sigma\ \mathcal{G}\right\} =$$
$$\left\{ A\subseteq D|\text{ שייכת לכל } \sigma - \text{ אלגברה } \mathcal{G} \text{ אשר מכילה את המלבנים } A\right\}$$

**משפט 2.29**

$\mathcal{F}\neq2^{D}$ ואי אפשר להגדיר $P$ כפי שאנחנו רוצים על $2^{D}$.

**הערה 2.30** למה חילקנו את הדיסק למלבנים ולא הסתכלנו על נקודות?
תהי $(x,y)\in D$.

$$\mathcal{F}\ni\{(x,y)\}=\bigcap_{n=1}^{\infty}\left[x-\frac{1}{n},x+\frac{1}{n}\right]\times\left[y-\frac{1}{n},y+\frac{1}{n}\right]$$

$$\{(x,y)\}\subseteq\left[x-\frac{1}{n},x+\frac{1}{n}\right]\times\left[y-\frac{1}{n},y+\frac{1}{n}\right]$$

ולכן:

$$P\left(\{(x,y)\}\right)\leq P\left(\left[x-\frac{1}{n},x+\frac{1}{n}\right]\times\left[y-\frac{1}{n},y+\frac{1}{n}\right]\right)=$$

$$\text{שטח של ריבוע עם צלע } \frac{2}{n}=\frac{4}{n^2}\underset{n\to\infty}{\longrightarrow}0$$

לכן, $P\left(\{(x,y)\}\right)=0$. כלומר ההסתברות של פגיעה בנקודה היא אפס.

**הערה 2.31** לכן חשוב להגביל איחודים לאיחודים בני מנייה (כלומר את התכונה השלישית של פונקציית ההסתברות לאיחודים.) כי אילו $P\left(\bigcup\limits_{i\in I}A_i\right)=\sum\limits_{i\in I}P\left(A_i\right)$ אז לכל קבוצה $A\in\mathcal{F}$ נקבל:

$$P\left(A\right)=P\left(\bigcup_{\omega\in A}\left(\{\omega\}\right)\right)=\sum_{\omega\in A}P\left(\{\omega\}\right)=\sum0=0$$

## 2.3 מרחבי מידה סופיים

$\Omega$ קבוצה סופית
מרחב ההסתברות האחיד על $\Omega$ הוא: $(\Omega,\mathcal{F},P)$ כאשר $\mathcal{F}=2^{\Omega}$ ו: $P\left(A\right)=\frac{|A|}{|\Omega|}$. בפרט: $P\left(\{\omega\}\right)=\frac{|\omega|}{\Omega}=\frac{1}{|\Omega|}$

**טענה 2.32**

זהו אכן מרחב הסתברות. כלומר: $P$ פונקציית הסתברות.

**הוכחה:**

$$P\left(\Omega\right)=\frac{|\Omega|}{|\Omega|}=1$$

$$A\subseteq\Omega\Rightarrow0\leq|A|\leq|\Omega|\Rightarrow P\left(A\right)=\frac{|A|}{|\Omega|}\in[0,1]$$

במידה ו$A_1,A_2,\ldots\subseteq\Omega$ קבוצות זרות אזי:

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=\frac{\left|\bigcup\limits_{n=1}^{\infty}A_n\right|}{\Omega}=\frac{\sum\limits_{n=1}^{\infty}|A_n|}{|\Omega|}=\sum_{n=1}^{\infty}\frac{|A_n|}{|\Omega|}=\sum_{n=1}^{\infty}P\left(A_n\right)$$

■

**דוגמה 2.33** מטילים מטבע 6 פעמים. מה ההסתברות שיצא $H$ בהטלה השניה ו $T$ בהטלה השלישית? (בהנחה להסתברות אחידה)

$$\Omega=\{(x_1,\ldots,x_6)\quad x_i\in\{H,T\}\}$$

המאורע המבוקש הוא:

$$A=\{(x_1,H,T,x_4,x_5,x_6)\;\; x_i\in\{H,T\}\}$$

$$P(A)=\frac{|A|}{|\Omega|}=\frac{2^4}{2^6}=\frac{1}{4}$$

**דוגמה 2.34** באותו הניסוי, מה ההסתברות שיצאו יותר $H$ מ$T$?
נסמן $B_H$ מאורע שיצא יותר $H$ ו$B_T$ מאורע שיצא יותר $T$.
יש התאמה חח"ע ועל מ$B_T \to B_H$ ע"י החלפת $H$ ל$T$ ולהפך. לכן

$$|B_H|=|B_T|\Rightarrow P(B_H)=\frac{|B_H|}{|\Omega|}=\frac{|B_T|}{|\Omega|}=P(B_T)$$

כמו כן: $B_H\cap B_T=\varnothing$ לכן:

$$P(B_H\cup B_T)=P(B_H)+P(B_T)=2P(B_H)$$

במקרה זה יותר קל לחשב :

$$\overbrace{\underbrace{P(\Omega\backslash(B_H\cup B_T))}_{\text{ההטלות שבהן יצא מספר שווה של } H \text{ ו-} T}}^{C}\quad=1-P(B_H\cup B_T)$$

ולכן:

$$|C|=H\text{ מספר הדרכים לבחור 3 איברים מתוך 6 עם ערך }=$$

$$\binom{6}{3}=\frac{6!}{3!3!}=20\Rightarrow P(C)=\frac{20}{64}\Rightarrow P(B_H)=\frac{1-P(C)}{2}=\frac{1-\frac{20}{64}}{2}$$

**דוגמה 2.35** מה הסיכוי לזכות בלוטו עם חשיבות לסדר בהנחה להסתברות אחידה.

$$\Omega=\left\{(x_1,\ldots,x_6,y)\,,\,\underset{i\neq j\Rightarrow x_i\neq x_j}{x_i\in\{1,\ldots,37\}},\;y\in\{1,\ldots,7\}\right\}$$

אם בבעלותי כרטיס אחד מסויים:

$$\hat{\omega}=(\hat{x_1},\ldots,\hat{x_6},\hat{y})$$

$$|\Omega|=\underbrace{(\text{מספר דרכים לבחור } x_1,\ldots,x_6)}_{37\cdot36\cdot35\cdot\ldots\cdot32}\times7$$

$$P(\{\hat{\omega}\})=\frac{1}{|\Omega|}=\frac{1}{1673844480}$$

במידה ונרצה להוסיף חשיבות לסדר, $|\hat{\omega}|$ מתאים ל $6!$ כרטיסים רגילים.　　06/11/20

**תזכורת 2.36** מרחב מדגם אחיד על קבוצה סופית $\Omega$ הוא:

$$P=\frac{|A|}{|\Omega|}$$

מוגדר על כל תת הקבוצות של $\Omega$.

$$\mathcal{F}=2^{\Omega}$$

**דוגמה 2.37 פרדוקס יום ההולדת:** בוחרים $n$ אנשים באקראי. מה ההסתברות שיש שניים מהם שנולדו באותו תאריך (יש 365 ימים בכל שנה)

$$\Omega=\{(d_1,\ldots,d_n)\ d_i\in\{1,\ldots,365\}\}$$

$$A=\{d_i=d_j,\ \text{כלשהם}\ i\neq j\ \text{עבור}\}$$

$$A^c=\{(d_1,\ldots,d_n)\in\Omega:\ d_i\neq d_j,\ \forall i\neq j\}$$

$$|A^c|=365\cdot 364\cdot\ldots\cdot(365-n+1)$$

$$|\Omega|=365^n$$

בהנחה של הסתברות אחידה:

$$P\left(A^c\right)=\frac{|A^c|}{|\Omega|}=\frac{365}{365}\cdot\frac{364}{365}\cdot\ldots\cdot\frac{365-n+1}{365}=\prod_{k=0}^{n-1}\left(1-\frac{k}{365}\right)$$

$$\approx 1-\frac{\sum\limits_{k=0}^{n-1}k}{365^n}\approx 1-\frac{n^2}{365}$$

ולכן עבור $n=23$ יצא כי:

$$P\left(A^c\right)<\frac{1}{2}$$

דהיינו:

$$P\left(A\right)=1-P\left(A^c\right)>\frac{1}{2}$$

**הערה 2.38** זה מכונה פרדוקס כי זה שובר את האינטואיציה שלנו, כאשר שואלים מה המסיכוי שלשני אנשים תהיה את אותו תאריך יום הולדת אנשים נוטים להסתכל על התאריך שלהם פרטנית ולחשב את הסיכוי, ולא להסתכל על סיכוי של זוג אפשרי בין כל האנשים.

## 2.4 מרחבי מדגם לא אחידים

**דוגמה 2.39** מטילים מטבע לא הוגן. יוצא $H$ בהסתברות $\frac{2}{3}$ ואילו $T$ בהסתברות $\frac{1}{3}$.
נבחר $\Omega=\{H,T\}$ נגדיר:

$$P(A)=\begin{cases}0 & A=\varnothing\\ \frac{2}{3} & A=\{H\}\\ \frac{1}{3} & A=\{T\}\\ 1 & A=\Omega\end{cases}$$

(בדיקה מראה שזה מרחב הסתברות)

**הגדרה 2.40 וקטור הסתברות:** וקטור $(p_\omega)_{\omega\in\Omega}$ נקרא וקטור הסתברות אם

- $$\forall\omega\in\Omega\ 0\le p_\omega$$
- $$\sum_{\omega\in\Omega}p_\omega=1$$

נניח ש$\Omega$ בן מנייה אם איננו יודעים לסכום יותר מכך

**דוגמה 2.41** בהינתן וקטור הסתברות

$$(q_\omega)_{\omega\in\Omega}=\left(\frac{1}{2},\frac{1}{4},\frac{1}{4},0\right)$$

עבור מרחב המדגם:

$$\Omega=\{\omega_0,\omega_1,\omega_2,\omega_3\}$$

אזי:

$$P(\omega_0)=\frac{1}{2},\ P(\omega_1)=\frac{1}{4},\ P(\omega_2)=\frac{1}{4},\ P(\omega_3)=0$$

**הערה 2.42** לא באמת מציינים את הוקטור בצורה מסודרת, אלא כל קואורדינטה נגישה בעזרת היחידון ולא בעזרת "אינדקסים".

**הגדרה 2.43** תהי $\Omega$ קבוצה סופית או בת־מנייה. $q=(q_\omega)_{\omega\in\Omega}$ וקטור הסתברות.
נגדיר לכל קבוצה $A\subseteq\Omega$:

$$P_q(A)=\sum_{\omega\in A}q_\omega$$

**טענה 2.44**

$P_q$ היא פונקציית הסתברות על $(\Omega,2^\Omega)$

**הוכחה:** ראשית נבחין כי:

$$P_q(\Omega)=\sum_{\omega\in\Omega}q_\omega=1$$

כי $q$ הוא וקטור הסתברות. כמו כן:

$$P_q(A)=\underbrace{\sum_{\omega\in A}q_\omega}_{\text{כל המחוברים}\geq 0}\geq 0$$

וכמו כן:

$$P_q(A)=\sum_{\omega\in A}q_\omega\underbrace{\leq}_{\text{הוספת איברים}}\sum_{\omega\in\Omega}q_\omega=1$$

בהנתן סדרה זרה $A_1,A_2,\ldots$ של תת קבוצות של $\Omega$:

$$P_q\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{\omega\in\bigcup_{n=1}^{\infty}A_n}q_\omega\underbrace{=}_{\text{אסוציאטיביות (כי } A_n \text{ זרים)}}\sum_{n=1}^{\infty}\left(\underbrace{\sum_{\omega\in A_n}q_\omega}_{P_q(A_n)}\right)=\sum_{n=1}^{\infty}P_q(A_n)$$

לכן $P_q$ מקיימת את כל הדרישות להיותה פונקציית הסתברות כנדרש. ■

**דוגמה 2.45**
נקבע $\omega_0\in\Omega$ נבחר:

$$q_\omega=\begin{cases}1 & \omega=\omega_0\\ 0 & \omega\neq\omega_0\end{cases}$$

$$P_q(A)=\sum_{\omega\in A}q_\omega=\begin{cases}1 & \omega_0\in A\\ 0 & \omega_0\notin A\end{cases}$$

$P_q$ תסומן כ $\delta_{\omega_0}$ תקרא מידת דלתא (פונקציית דלתא)

**דוגמה 2.46** מטילים מטבע לא הוגן $n$ פעמים, ההסתברות של כל הטלה להיות $H$ היא $\frac{2}{3}$. עבור:

$$\Omega=\{H,T\}^n$$

$$(x_1,\ldots,x_n)=\omega\in\Omega$$

כאשר $x_1,\ldots,x_n\in\{H,T\}$

$$q_\omega=\left(\frac{2}{3}\right)^{\#\{1\leq i\leq n:\, x_i=H\}}\cdot\left(\frac{1}{3}\right)^{\#\{1\leq i\leq n,\, x_i=T\}}$$

**טענה 2.47**

$(q_\omega)_{\omega\in\Omega}$ הוא וקטור הסתברות לכן $P_q$ היא פונקצית הסתברות

**הערה 2.48** ראינו בשיעור שעבר שבמרחבים לא בני מניה יש פונקצית הסתברות שאינן מהטיפוס הזה ( למשל $P(\{\omega\})=0$ לכל $\omega\in\Omega$)

**טענה 2.49**

אם $\left(\Omega,2^{\Omega},P\right)$ מרחב הסתברות, ו־ $\Omega$ סופית או בת מנייה אז $P=P_q$ עבור וקטור הסתברות $q$ מתאים.

**הוכחה:**

**הערה 2.50** אם $P=P_q$ אז לכל $\omega_0\in\Omega$:

$$P\left(\underbrace{\{\omega_0\}}_{\text{ההנחה שיחידונים מדידים חשובה}}\right)=P_q\left(\{\omega_0\}\right)=\sum_{\omega\in\{\omega_0\}}q_\omega=q_{\omega_0}$$

נבדוק שהוקטור $q=\left(q_\omega\right)_{\omega\in\Omega}$ המוגדר כך הוא וקטור הסתברות ושמתקיים $P=P_q$
מכיוון ש $P\geq 0$ גם:

$$q_\omega=P\left(\{\omega\}\right)\geq 0$$

$$\sum_{\omega\in\Omega}q_\omega=\sum_{\omega\in\Omega}P\left(\{\omega\}\right)\underbrace{=}_{\text{סכום בן־מנייה והיחידונים זרים}}P\left(\bigcup_{\omega\in\Omega}\{\omega\}\right)=P\left(\Omega\right)=1$$

ראינו ש $q$ אכן וקטור הסתברות.
לכל $A\subseteq\Omega$:

$$P_q\left(A\right)=\sum_{\omega\in A}q_\omega=\sum_{\omega\in A}P\left(\{\omega\}\right)=P\left(\bigcup_{\omega\in A}\{\omega\}\right)=P\left(A\right)$$

לכן $P_q\left(A\right)=P\left(A\right)$ לכל $A\subseteq\Omega$ כלומר: $P_q=P$ כנדרש. ■

**טענה 2.51**

אם $\left(\Omega,\mathcal{F},P\right)$ מרחב הסתברות ו: $A,B\in\mathcal{F}$ אז:

$$P\left(A\cup B\right)=P\left(A\right)+P\left(B\right)-P\left(A\cap B\right)$$

**הערה 2.52** הוכחנו בתרגיל 1

**טענה 2.53**

אם $A_n$ סדרת מאורעות אז:

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

**הוכחה:** מגדירים:

$$B_n=A_n\backslash\bigcup_{k=1}^{n-1}A_k$$

אפשר לבדוק באינדוקציה כי:

$$\bigcup_{n=1}^{N} B_n = \bigcup_{n=1}^{N} A_n$$

לכן:

$$\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$$

לכן:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right)$$

$B_n$ זרות כי $B_n \subseteq A_n$ אם $k < n$ אז $B_n \subset A_n \backslash B_k$ לכן:

$$= \sum_{n=1}^{\infty} P\left(B_n\right)$$

כיוון ש $B_n \subseteq A_n$ אזי $P\left(B_n\right) \le P\left(A_n\right)$ ולכן:

$$\le \sum_{n=1}^{\infty} P\left(A_n\right)$$

■

## 2.5 מרחבי מידה אינסופיים

**הגדרה 2.54 סדרת קבוצות עולה\יורדת:** סדרת קבוצות $A_n$ היא עולה אם $A_n \subseteq A_{n+1}$ ויורדת אם: $A_n \supseteq A_{n+1}$

**משפט 2.55**

אם $A_n$ סדרת מאורעות עולה אז:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$$

ואם $A_n$ סדרה יורדת אז:

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$$

**הוכחה:** <u>**מקרה עולה**</u>: כמו קודם, נגדיר:

$$B_n = A_n \backslash \bigcup_{k=1}^{n-1} A_k = A_n \backslash A_{n-1}$$

הזכרנו ש: $B_n$ סדרה זרה ו:

$$\underbrace{\bigcup_{n=1}^{N} A_n}_{A_N} = \bigcup_{n=1}^{N} B_n$$

ולכן:

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=P\left(\bigcup_{n=1}^{\infty}B_n\right)\underbrace{=}_{\text{זרים } B_n}\sum_{n=1}^{\infty}P\left(B_n\right)=\lim_{N\to\infty}\sum_{n=1}^{N}P\left(B_n\right)=\lim_{N\to\infty}P\left(\underbrace{\bigcup_{n=1}^{N}B_n}_{A_N}\right)$$

**מקרה יורד:** תרגיל. (רמז: $A_n$ יורד $\iff$ $A_n^c$ עולה) ■

### 2.5.1 $\limsup/\liminf$

**הגדרה 2.56** תהי $A_n$ סדרת מאורעות. הקבוצה:

$$\limsup A_n=\{\omega\in\Omega:\ \omega\in A_n \text{ עבור אינסוף } n\text{־ים}\}$$

$$\liminf A_n=\{\omega\in\Omega:\ \omega\in A_n \text{ לכל } n \text{ גדול מספיק}\}$$

**הערה 2.57** $\limsup A_n$ = "אינסוף מאורעות קרו"
$\liminf A_n$ - "כל המאורעות קרו למעט מספר סופי"

נסיון שגוי לכתוב את $\limsup A_n$:

$$\bigcup_{I\subseteq\mathbb{N}\ |I|=\infty}\underbrace{\{\omega\in A_n,\ n\in I,\ \omega\notin A_n,\ n\notin I\}}_{\bigcap_{i\notin I}A_i\cap\bigcap_{i\in\mathbb{N}\setminus I}A_i^c}$$

הבעיה היא באיחוד הכללי, כיוון שיש מספר לא בן מנייה של $I$־ים המקיימים $|I|=\infty$. לכן מההצגה הזו של $\limsup A_n$ אי אפשר להסיק שהיא מדידה (דהיינו, שהיא מאורע).

הדרך התקנית לכתוב את $\limsup A_n$ היא: $\omega\in A_n$ עבור אינסוף $n$־ים $\iff$ לכל $n_0\in\mathbb{N}$ יש $n>n_0$ כך ש $\omega\in A_n$ $\iff$

$$\iff\omega\in\bigcap_{n_0=1}^{\infty}\{\eta\in\Omega:\ \text{קיים } n>n_0 \text{ כך ש } \eta\in A_n\}\iff$$

$$\iff\omega\in\bigcap_{n_0=1}^{\infty}\underbrace{\bigcup_{n=n_0+1}^{\infty}A_n}_{\text{הקבוצה הפנימית בביטוי הקודם}}$$

וזהו איחוד בן מנייה של איחודים בני מנייה.

**מסקנה 2.58**

$\limsup A_n$ הוא מאורע (בהנחה שכל ה $A_n$־ים מאורעות).

מה המשמעות של $\omega\in\liminf A_n$ $\iff$ קיים $n_0$ כך ש: [$\omega\in A_n$ לכל $n>n_0$] $\iff$

$$\omega\in\bigcup_{n_0=1}^{\infty}\{\eta\in\Omega:\ \forall n>n_0\ \eta\in A_n\}=\bigcup_{n_0=1}^{\infty}\bigcap_{n=n_0+1}^{\infty}A_n$$

**מסקנה 2.59**

גם $\liminf A_n$ מדידה אם כל $A_n$ מדידים.

**הערה 2.60** לכל קבוצה $A \subseteq \Omega$ הפונקצייה המציינת של $A$ היא: $\mathbb{1}_A : \Omega \to \{0,1\}$

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

**דוגמה 2.61** אם $\Omega \in \mathbb{R}$ $A = [a,b]$ אז:

$$\mathbb{1}_A = \text{Some graph here}$$

כל פונקציה: $f : \Omega \to \{0,1\}$ היא מהצורה $\mathbb{1}_A$ (כי לוקחים: $A = \{\omega : f(\omega) = 1\}$ )
אם $A_n$ סדרת מאורעות לכל $\omega$ אפשר להגדיר:

$$\begin{aligned} f(\omega) &= \limsup \mathbb{1}_{A_n}(\omega) \\ g(\omega) &= \liminf \mathbb{1}_{A_n}(\omega) \end{aligned}$$

$$f(\omega), g(\omega) \in \{0,1\}$$

**טענה 2.62**

$f = \mathbb{1}_{\limsup A_n}$, $g = \mathbb{1}_{\liminf A_n}$

**הוכחה:** $f(\omega) = 1 \iff \limsup \mathbb{1}_{A_n}(\omega) = 1 \iff$ יש אינסוף $n$־ים כך ש $\mathbb{1}_{A_n}(\omega) = 1 \iff \omega \in A_n$ עבור אינסוף $n$־יםץ
$g(\omega) = 1 \iff \liminf \mathbb{1}_{A_n}(\omega) = 1 \iff$ יש מספרי סופי של $n$־ים כך ש$\mathbb{1}_\omega(\omega) = 0 \iff \omega \notin A_n$ עובר מספר סופי של $n$ ־ים $\omega \in \liminf A_n$ ■

07/11/20

**הערה 2.63**

$$\liminf A_n \subseteq \limsup A_n$$

נובע ישירות מההגדרה.

**תזכורת 2.64** $f : \mathbb{R} \to \mathbb{R}$ רציפה ב$x_0$ לכל סדרה $x_n \to x_0$ מתקיים $f(x_n) \to f(x_0)$.

ולכן, בהנתן משפחת קבוצות $\mathcal{F}$ ובהנתן פונקציה: $P : \mathcal{F} \to \mathbb{R}$ נאמר ש $P$ רציפה ב $A \in \mathcal{F}$ אם $A_n \to A$ גורר: $P(A_n) \to P(A)$.

**הגדרה 2.65 התכנסות של סדרת קבוצות:** נאמר ש $A_n \to A$ אם לכל $\omega$ מתקיים:

- $\omega \in A \iff \omega \in A_n$ לכל $n$ גדול מספיק
- $\omega \in A^c \iff \omega \in A_n^c$ לכל $n$ גדול מספיק

במילים אחרות, שייכות של $\omega$ ל$A$ מוכרע ע"י השייכות ל $A_n$ עבור $n$ גדול דיו.

לכן, $A_n$ לא מתכנסת אם יש $\omega$ כך ש $\omega \in A_n$ לאינסוף $n$־ים וגם $\omega \in A_n^c$ לאינסוף $n$־ים

**דוגמה 2.66**

1. $A_n=\left(-1+\frac{1}{n},1-\frac{1}{n}\right)$ נקבל כי: $A_n\to(-1,1)$ כיוון שאם $\omega\in(-1,1)$ אזי לכל $n$ גדול מספיק נקבל $\omega\in A_n$ ואם $\omega\notin(-1,1)$ אזי $\omega\notin A_n$ לכל $n$. בפרט לכל $n$ גדול דיו.
2. נתבונן ב $A_{2n}=[-1,0]$ ואילו: $A_{2n+1}=[0,1]$. נניח בשלילה כי $A_n\to A$. נתבונן ב $\omega=\frac{1}{2}$ אז $\omega\notin A$ כי $\omega\notin A_n$ עבור כל $n$ זוגי. אבל גם $\omega\notin A_n^c$ עבור כל $n$ אי־זוגי.

**טענה 2.67**

$A_n\to A$ אס"ס $A=\limsup A_n=\liminf A_n$

**הוכחה:** נניח ש $A_n\to A$. ציינו קודם כי: $\liminf A_n\subseteq\limsup A_n$. אם $\omega\in A$ אזי: $\omega\in A_n$ לכל $n$ גדול מספיק, ולכן גם: $\omega\in\liminf A_n$ לכן $A\subseteq\liminf A_n$
נניח כי $\omega\in\limsup A_n$ אזי $\omega\in A_n$ לאינסוף $n$־ים. לא מתקיים $\omega\in A_n^c$ החל ממקום מסויים. ולכן לפי הגדרת ההתכנסות, לא מתקיים $\omega\in A^c$ ולכן $\omega\in A$. וקיבלנו כי: $\limsup A_n\subseteq A$ ולכן קיבלנו:

$$A\subseteq\liminf A_n\subseteq\limsup A_n\subseteq A\Rightarrow A=\liminf A_n=\limsup A_n$$

**__כיוון שני:__**
נותר להראות כי אם $A_n$ לא מתכנסת, אזי $\limsup A_n\neq\liminf A_n$.
ראינו ש$A_n$ אינה מתכנסת אם יש $\omega$ כך ש $\omega$ כך שהתכונה $\omega\in A_n$ מתקיימת לאינסוף $n$־ים וגם לא מתקיימת לאינסוף $n$־ים ולכן: $\omega\in\limsup A_n$ אבל $\omega\notin\liminf A$ ולכן: $\liminf A_n\neq\limsup A_n$ ■

**למה 2.68**

אם $A_n$ עולה, אז: $A_n\to\bigcup_{n=1}^{\infty}A_n$ ואם יורדת אז: $A_n\to\bigcap_{n=1}^{\infty}A_n$.

**תרגיל:** להוכיח את הלמה הנ"ל

**למה 2.69 פטו (*Fatou*)**

$$\underbrace{P\left(\liminf A_n\right)}_{\text{קבוצות}}\leq\underbrace{\liminf P\left(A_n\right)}_{\text{מספרים}}$$

**הוכחה:** נשים לב כי:

$$\liminf A_n=\bigcup_{n=1}^{\infty}\underbrace{\bigcap_{k=n+1}^{\infty}A_k}_{B_n}$$

נשים לב כי $B_n$ עולה כי ככל ש $n$ גדל, חותכים פחות קבוצות $A_k$ לכן:

$$P\left(\liminf A_n\right)=\lim_{n\to\infty}P\left(B_n\right)$$

אבל:

$$P\left(B_n\right)=P\left(\bigcap_{k>n}a_k\right)\leq\inf_{k>n}P\left(A_k\right)$$

כיוון ש $\bigcap_{j>n}A_j\supseteq A_k$ לכל $j>n$. כלומר:

$$P\left(\liminf A_n\right)\leq\lim_{n\to\infty}\inf_{k>n}P\left(A_k\right)=\liminf P\left(A_n\right)$$

■

**למה 2.70**

באופן דומה: $P\left(\limsup A_n\right) \geq \limsup P\left(A_n\right)$

**משפט 2.71 רציפות של $P$ על $\mathcal{F}$**

בהינתן $(\Omega, \mathcal{F}, P)$ מרחב הסתברות, והסדרה $\{A_n\}_{n=1}^{\infty}$ כך ש $A_n \in \mathcal{F}$ לכל $n$. אזי אם: $A_n \to A$ אז $A \in \mathcal{F}$ ומתקיים:

$$P\left(A_n\right) \to P\left(A\right)$$

**הוכחה:** נתון $A_n \to A$. ראינו $A = \limsup A_n = \liminf A_n$ בפרט, $A \in \mathcal{F}$. כמו כן:

$$\liminf P\left(A_n\right) \underbrace{\geq}_{\text{Fatou}} P\left(\liminf A_n\right) = P\left(A\right) = P\left(\limsup A_n\right) \geq \limsup P\left(A_n\right) \geq \liminf P\left(A_n\right)$$

ולכן כל האי שיוויות הם בעצם שיוויות ולכן:

$$\lim P\left(A_n\right) = P\left(A\right)$$

כנדרש. ■

13/11/20

**דוגמה 2.72** מישהו זורק 2 קוביות. מה הסיכוי ששניהם אי־זוגיים?
נשים לב כי מרחב המדגם:

$$\Omega = \{(i,j) : i,j \in \{1,\ldots,6\}\}$$

בהסתברות אחידה.

$$A = \{(i,j) : i,j \in \{1,3,5\}\}$$

$$P\left(A\right) = \frac{|A|}{|\Omega|} = \frac{3^2}{6^2} = \frac{1}{4}$$

נניח שאנו יודעים שבאחת מהקוביות יצא 5. $B$= מאורע זה.
מה ההסתברות של $A \cap B$?

$$B = \{(1,5),(2,5),(3,5),(4,5),(5,5),(6,5),(5,1),(5,2)\ldots\}$$

לכן:

$$|B| = 11, \quad |A \cap B| = |5|$$

ולכן ההסתברות כעת היא:

$$\frac{|A \cap B|}{|B|} = \frac{5}{11}$$

**הגדרה 2.73** תהי $(\Omega,\mathcal{F},P)$ מרחב הסתברות. נניח ש $B\in\mathcal{F}$ מאורע ו: $P(B)>0$.
ההסתברות המותנית של מאורע $A$ בהנתן $B$ הוא:

$$P(A|B)=\frac{P(A\cap B)}{P(B)}$$

**הערה 2.74** מתקיימת זהות:

$$P(A\cap B)=P(B)\,P(A|B)$$

בהנחה ש $P(B)>0$.

**דוגמה 2.75** נתון כד עם 6 כדורים. 4 לבנים ו2 שחורים. בוחרים שני כדורים ברצף.
נגדיר מאורע:

$A$ = הראשון יצא לבן
$B$ = השני יצא שחור

מהו $P(B|A)$?

$$P(B|A)=\frac{P(B\cap A)}{P(A)}$$

מרחב המדגם הוא:

$$\Omega=\{(i,j):,\ i\neq j,\ i,j\in\{1,\ldots,6\}\}$$

ונסכים שהכדורים הלבנים ממוספרים $1,\ldots,4$.

$$A=\{(i,j):\ i\in\{1,\ldots,4\},\ j\neq i,\ j\in\{1,\ldots,6\}\}$$

$$|A|=4\cdot 5=20$$

$$B=\{(i,j):\ j\in\{5,6\}\ i\neq j,\ i\in\{1,\ldots,6\}\}$$

$$|A\cap B|=\{(i,j):\ i\in\{1,\ldots,4\},\ j\in\{5,6\}\}$$

$$|A\cap B|=4\cdot 2$$

ולכן:

$$P(B|A)=\frac{\frac{|A\cap B|}{|\Omega|}}{\frac{|A|}{|\Omega|}}=\frac{4\cdot 2}{4\cdot 5}=\frac{2}{5}$$

**טענה 2.76**

אם $P(B)>0$ אז אם נגדיר: $P_B(A)=P(A|B)$ אז $P_B$ היא פונקציית הסתברות על $(\Omega,\mathcal{F})$

**הוכחה:** :

$$0 \underbrace{\leq}_{\text{מנסה של מספרים אי שליליים}} P_B(A)=\frac{P(A\cap B)}{P(B)} \underbrace{\leq}_{\substack{A\cap B\subseteq B\\ P(A\cap B)\leq P(B)}} \frac{P(B)}{P(B)}=1$$

כמו כן, $\Omega\cap B=B$ לכן:

$$P_B(\Omega)=\frac{P(\Omega\cap B)}{P(B)}=\frac{P(B)}{P(B)}=1$$

נניח ש $A_n$ סדרה זרה של מאורעות. אז הסדרה: $A'_n=A_n\cap B$ גם כן זרה כי לכל $i\neq j$:

$$A'_i\cap A'_j=(A_i\cap B)\cap(A_j\cap B)=B\cap(A_i\cap A_j)=B\cap\varnothing=\varnothing$$

לכן:

$$P_B\left(\bigcup_{n=1}^{\infty}A_n\right)=\frac{P\left(B\cap\bigcup_{n=1}^{\infty}A_n\right)}{P(B)}=\frac{P\left(\bigcup_{n=1}^{\infty}\overbrace{(B\cap A_n)}^{A'_n}\right)}{P(B)}=\sum_{n=1}^{\infty}\frac{P(A_n\cap B)}{P(B)}=\sum_{n=1}^{\infty}P_B(A_n)$$

כנדרש. ■

**הערה 2.77** אם מגדירים: $\mathcal{F}_B=\{C\cap B: C\in\mathcal{F}\}$ אפשר לבדוק ש: $\mathcal{F}_B$ היא $\sigma$-אלגברה ו $P_B$ היא פונקציית הסתברות על $(B,\mathcal{F}_B)$.

**טענה 2.78 נוסחת ההסתברות השלמה**

נניח ש: $B_1,\dots,B_n$ חלוקה של הקבוצה $B$. כלומר: $B=\bigcup_{i=1}^{n}B_i$ והאיחוד זר. נניח גם: $P(B_i)>0$ אז:

$$P(A\cap B)=\sum_{i=1}^{N}P(A|B_i)P(B_i)$$

בפרט אם: $B_1,\dots,B_n$ היא חלוקה של $\Omega$ אז:

$$P(A)=\sum_{i=1}^{n}P(A|B_i)P(B_i)$$

**הוכחה:**

$$A\cap B=\bigcup_{i=1}^{n}A\cap B_i$$

האיחוד זר כי $B_i$ זרים ולכן לפי הגדרה:

$$P(A\cap B)=\sum_{i=1}^{n}P(A\cap B_i)=\sum_{i=1}^{n}P(A|B_i)P(B_i)$$

במקרה $B=\Omega$ אז זה מיידי מהנ"ל כי: $A\cap B=A$. ■

**דוגמה 2.79** נניח שיש שלושה מפעלים א,ב,ג שמייצרים מוצר מסויים. הסמויות והסיכוי לייצור מוצר פגום נתון בטבלה הבאה:

| | מפעל א | מפעל ב | מפעל ג |
|---|---|---|---|
| מספר יחידות | 1000 | 2000 | 3000 |
| % פגומים | 20% | 10% | 5% |

קניתי מוצר.

1. מה ההההסתברות שהוא יוצא במפעל א'?
2. מה ההסתברות שהוא פגום בהנתן שהוא לא יוצר במפעל ג?
3. מה ההסתברות שהוא פגום.
4. מה ההסתברות שהוא יוצר במפעל א' בהנתן שהוא פגום?

נסמן $A,B,C$ המאורעות שהמוצר יוצר במפעל א,ב,ג בהתאמה. וב$R$ את המאורע שהמוצר פגום.

**פתרון**

.1

$$P\left(A\right)=\frac{1000}{6000}=\frac{1}{6}$$

2. ראשית נשים לב כי: $C^{c}=A\uplus B$ ולכן:

$$P\left(C^{c}\right)=\frac{3000}{6000}=\frac{1}{2}$$

$$P\left(R|C^{c}\right)=\frac{P\left(C^{c}\cap R\right)}{\underbrace{P\left(C^{c}\right)}_{1/2}}$$

$$P\left(R\cap C^{c}\right)=P\left(R|A\right)P\left(A\right)+P\left(R|B\right)P\left(B\right)$$

נתון:

$$P\left(R|A\right)=0.2,\ P\left(R|B\right)=0.1$$

$$P\left(R|C^{c}\right)=\frac{0.2\cdot\frac{1}{6}+0.1\cdot\frac{1}{3}}{\frac{1}{2}}=\frac{2}{15}$$

.3

$$P\left(R\right)=P\left(R|A\right)P\left(A\right)+P\left(R|B\right)P\left(B\right)+P\left(R|C\right)P\left(C\right)=\frac{11}{120}$$

.4

$$P\left(A|R\right)=\frac{P\left(A\cap R\right)}{\underbrace{P\left(R\right)}_{\frac{11}{120}}}=\frac{120}{11}\underbrace{\left(P\left(R|A\right)P\left(A\right)\right)}_{P\left(R\cap A\right)}=\frac{120}{11}\cdot0.2\cdot\frac{1}{6}=\frac{4}{11}$$

**הערה 2.80** הסימון $A\uplus B$ מסמן איחוד זר.

**משפט 2.81** ***Bayse***

נניח ש: $P(A), P(B) > 0$ אז:

$$P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}$$

אם:

$$\Omega = \bigcup_{i=1}^{n} B_i$$

$B_i$זרים ו $P(B_i) > 0$ אז:

$$P(B|A) = \frac{P(A|B_i)\,P(B_i)}{\sum\limits_{j=1}^{n} P(A|B_j)\,P(B_j)}$$

**הוכחה:**

$$P(B) = P(A|B) = P(A \cap B) = P(A)\,P(B|A)$$

מחלקים ב $P(B)$ בחלק השני. יודעים מהחלק הראשון:

$$P(B_i|A) = \frac{P(A|B_i)\,P(B_i)}{P(A)}$$

מנוסחת ההסתברות השלמה נקבל:

$$P(A) = \sum_{j=1}^{n} P(A|B_j)\,P(B_j)$$

ולכן הנדרש. ■

**דוגמה 2.82** מחלה מסויימת בעלת שכחיות באוכלוסיה של $1:100,000$ .

יש בדיקת דם כך ש:

- אם מישהו חולה נבדק התוצאה חיובית בסיכוי $95\%$
- אם מישהו בריא נבדק התוצאה היא בסיכוי של $0.5\%$

מה הסיכוי שמישהו שתוצאות הבדיקה שלו יצא חיובית אכן חולה?

<u>**הפתרון:**</u>

נסמן את מרחב המדגם שלנו:

$$\Omega = \left\{ \underbrace{s}_{\text{sick}}, \underbrace{h}_{\text{healthy}} \right\} \times \{+,-\}$$

נסמן:

$$S = \{(s,+),(s,-)\}$$

ונסמן:

$$\begin{aligned} T^+ &= \{(s,+),(h,+)\} \\ T^- &= \{(s,-),(h,-)\} \end{aligned}$$

נתונים:

$$\begin{aligned} P(S) &= \frac{1}{100,000} \\ P\left(T^+|S\right) &= 0.95 \\ P\left(T^+|S^c\right) &= 0.005 \end{aligned}$$

נרצה לחשב את $P\left(S|T^+\right)$. מנוסחת Bayse:

$$P\left(S|T^+\right) = \frac{P\left(T^+|S\right)P(S)}{P\left(T^+|S\right)P(S)+P\left(T^+|S^c\right)P\left(S^c\right)}$$

(לקחנו את החלוקה $\Omega = S \dot\cup S^c$)

$$= \frac{0.95\cdot\frac{1}{100,000}}{0.95\cdot\frac{1}{100,000}+0.005\left(1-\frac{1}{100,000}\right)} \sim 0.002$$

כלומר, גם ניסוי שנראה יחסית אמין, לא באמת כל־כך טוב.

**טענה 2.83**

נניח: $B = \bigcup_{i=1}^{n} B_i$ ונניח ש: $P(A|B_i) \geq \alpha$ לכל $i$ אז: $P(A|B) \geq \alpha$

**הוכחה:** מנוסחת ההסתברות השלמה:

$$P(A\cap B) = \sum_{i=1}^{n} P(A|B_i)P(B_i) \geq \alpha\sum_{i=1}^{n} P(B_i) = \alpha P(B)$$

ולכן:

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \geq \alpha$$

■

**דוגמה 2.84 פרדוקס סימפסון:**

| | נרשמו לא' | התקבלו לא' | נרשמו לב' | התקבלו לב' | נרשמו סה"כ | התקבלו סה"כ |
|---|---|---|---|---|---|---|
| גברים | 10 | 1 (10%) | 100 | 50 (50%) | 110 | 51 (46%) |
| נשים | 100 | 20 (20%) | 10 | 7 (70%) | 110 | 27 (24.5%) |

$$\Omega = \{m,w\}\times\{a,b\}\times\{+,-\} = \{(m,a,+),(m,a,-),\ldots\}$$

$$\begin{aligned} M &= \{\text{הנרשם גבר}\} = \{(m,?,?)\} \\ W &= \{\text{הנרשמת אישה}\} = \{(w,?,?)\} \\ A &= \{\text{הנרשם נרשם למחלקה א'}\} = \{(?,a,?)\} \\ B &= \{\text{הנרשם נרשם למחלקה ב'}\} = \{(?,a,?)\} \\ U &= \{\text{הנרשם התקבל}\} = \{(?,?,+)\} \end{aligned}$$

$$P(U|M)=\frac{P(U\cap M)}{P(M)}\underbrace{=}_{\Omega=A\cup B}\frac{P(U\cap M\cap A)+P(U\cap M\cap B)}{P(M\cap A)+P(M\cap B)}=$$

$$\frac{\overbrace{P(U|M\cap A)}^{\text{סיכוי קבלה של גבר למחלקה א}}P(M\cap A)+\overbrace{P(U|M\cap B)}^{\text{סיכוי קבלה של גבר למחלקה ב'}}P(M\cap B)}{P(M\cap A)+P(M\cap B)}$$

$$P(U|W)=\frac{\overbrace{P(U|W\cap A)}^{\text{סיכוי קבלה של אישה למחלקה א}}P(W\cap A)+\overbrace{P(U|W\cap B)}^{\text{סיכוי קבלה של אישה למחלקה ב'}}P(W\cap B)}{P(W\cap A)+P(W\cap B)}$$

**הגדרה 2.85 סכום קמור:** סכום קמור של מספרי $x,y$ הוא מספר מהצורה:

$$tx+(1-t)y$$

עבור $0\le t\le 1$

**הערה 2.86** לעבור ברשימות של רז על ה sibiling paradox

## 3 ניסויים תלויים

**דוגמה 3.1** מטילים מטבע שיוצא $H$ בהסתברות $p$.
אם יצא $H$ מטילים שוב את אותו מטבע, אחרת מטילים מטבע הוגן.
אם $A$ הוא המאורע שיצא $H$ בהטלה הראשונה
$B$ שיצא $H$ בהטלה השניה אז נתון:

$$P(B|A)=p$$
$$P(B|A^c)=\frac{1}{2}$$

מרחב המדגם שלנו הוא:

$$\Omega=\{H,T\}^2$$

אנו רוצים $P$ כנ"ל. הנתונים גוררים ש:

$$P(\underbrace{\{(H,H)\}}_{B\cap A})=P(A)P(B|A)=p\cdot p$$

$$P(\{(H,T)\})=P(A\cap B^c)=P(B^c|A)P(A)\underbrace{=}_{\text{פונקציית הסתברות } P(\cdot|A)}\left(1-\underbrace{P(B|A)}_{p}\right)p=(1-p)$$

באותו אופן:

$$P(\{(T,H)\})=P(\{T,T\})=(1-p)\frac{1}{2}$$

ראינו שאם יש הסתברות $P$ כנדרש אז היא נקבעת ע"י הסתברויות היחידונים. צריך לבדוק שסכום ההסתברויות הוא 1:

$$p\cdot p+p(1-p)+(1-p)\frac{1}{2}+(1-p)\frac{1}{2}=1$$

14/11/20

## 3.1 בניית מרחבי הסתברות לסדרת ניסויים

לכל $n$ מתון מרחב מדגם $\Omega_n$ סופי או בן מנייה. מבצעים סדרת ניסויים. הניסוי ה $n$ עם תוצאה ב $\Omega_n$ או $1 \le n \le N$ או: $n = 1, 2, 3, \ldots$
מרחב המדגם הוא:

$$\Omega_1 \times \Omega_2 \times \ldots \times \Omega_N = \Omega^N = \{(\omega_1, \ldots, \omega_N) : \omega_i \in \Omega_i\}$$

או לחילופין:

$$\Omega^\infty = \Omega_1 \times \Omega_2 \times \ldots = \{(\omega_1, \omega_2, \ldots) : \omega_i \in \Omega_i\}$$

בהנתן $\omega \in \Omega_n$ נגדיר:

$$A_{n,\omega} = \left\{ \underbrace{\eta}_{(\eta_1,\ldots,\eta_N)} \in \Omega^N : \eta_n = \omega \right\}$$

**דוגמה 3.2**

$$\Omega_1 \times \Omega_2 \times \Omega_3 = \{H, T\}$$

$$\Omega^3 = \{(x, y, z) : x, y, z \in \{H, T\}\}$$

$$A_{2,H} = \{\text{יצא } H \text{ בהטלה השנייה}\} = \{(H,H,H), (H,H,T), (T,H,H), (T,H,T)\}$$

**הערה 3.3** בהינתן: $\omega_1 \in \Omega_1, \ldots, \omega_n \in \Omega_n$:

$$\bigcap_{i=1}^{n} A_{i,\omega_i} = \{\omega_1, \ldots, \omega_n \text{ יצא בהניסויים הראשונים } n\}$$

**למה 3.4**

בהנתן מאורעות $B_1, \ldots, B_k$ בעלי חיתוך עם הסתברות חיובית במרחב ההסתברות:

$$P\left(\bigcap_{i=1}^{k} B_i\right) = P(B_1) P(B_2|B_1) P(B_3|B_2 \cap B_1) \cdot \ldots \cdot P\left(B_k | \bigcap_{i=1}^{k-1} B_i\right)$$

**הוכחה:** באינדוקציה:
**בסיס:** עבור $k = 2$

$$P(B_1 \cap B_2) = P(B_1) P(B_2|B_1)$$

(נובע מהגדרת: $P(B_2|B_1) = \frac{P(B_1 \cap B_2)}{P(B_1)}$)
עבור $k = 3$:

$$P(B_1 \cap [B_2 \cap B_3]) = P(B_1) P(B_2 \cap B_3 | B_1)$$

אבל אנו יודעים:

$$P_{B_1}\left(B_2\cap B_3\right)=\underbrace{P_{B_1}\left(B_2\right)}_{P(B_2|B_1)}P_{B_1}\left(B_3|B_2\right)$$

אפשר לבדוק:

$$P_{B_1}\left(B_3|B_2\right)=P\left(B_3|B_1\cap B_2\right)$$

■

**משפט 3.5**

בהינתן $\Omega_1,\ldots,\Omega_N$ כנ"ל. נניח שלכל $n$ מתקיים $\omega_1,\ldots,\omega_n\in\Omega_1\times\ldots\times\Omega_n$ נתונה מידת (פונקציית) הסתברות על $\Omega_n$ המסומנת $Q_{\omega_1,\ldots,\omega_{n-1}}\left(\cdot\right)$:

אז:

1. עבור $N<\infty$ קיימת מידת הסתברות על $\left(\Omega^N,2^{\Omega^N}\right)$ $P=P^N$ כך ש:

$$P\left(A_{n,\omega_n}|\bigcap_{i=1}^{n-1}A_{i,\omega_i}\right)=Q_{\omega_1,\ldots,\omega_{n-1}}\left(\{\omega_n\}\right)$$

2. עבור $N=\infty$ קיימת $\sigma$-אלגברה $\mathcal{F}^\infty$ על $\Omega_1\times\Omega_2\times\ldots=\Omega^\infty$ כך ש: $A_{n,\omega}\in\mathcal{F}^\infty$ לכל $n$ ולכל $\omega\in\Omega_n$ קיימת מידת הסתברות $P=P^\infty$ על $(\Omega^\infty,\mathcal{F}^\infty)$כך ש:

$$P\left(A_{n,\omega_n}|\bigcap_{i=1}^{n-1}A_{i,\omega_i}\right)=Q_{\omega_1,\ldots,\omega_{n-1}}\left(\{\omega_n\}\right)$$

**הערה 3.6** אם $N<\infty$ אז $\Omega^N$ בן מנייה (אם $|\Omega_i|<\infty$ לכל $i$ אז $\left|\Omega^N\right|<\infty$)
אבל $\Omega^\infty$ לא בן מנייה (למעט במקרה שכל $\Omega_i$ הם יחידונים פרט למספר סופי של $i$).
ראינו באינפי 2 את הדוגמה של:

$$\{0,1\}^\infty=\{(x_1x_2,\ldots):\,x_i\in\{0,1\}\}$$

כלומר כתיב בינארי, והוכחנו כי הקבוצה אינה בת מנייה.

**הערה 3.7** במקרה בו $n=1$ יש סדרה יחידה $\omega_1,\ldots,\omega_{n-1}$ - הסדרה הריקה. לכן עבור $n=1$ נתון $Q_\varnothing$ יחיד.

**הוכחה:** **חלק ראשון ($N<\infty$):**
במקרה זה, $\Omega^N$ לכל היותר בת־מנייה. ולכן כל מידת הסתברות על $\left(\Omega^N,2^N\right)$נקבעת ע"י הערכים על יחידונים.

**הערה 3.8 מוטיבציה:** אילו $P$ מקיימת את:

$$P\left(A_{n,\omega_n}|\bigcap_{i=1}^{n-1}A_{i,\omega_i}\right)=Q_{\omega_1,\dots,\omega_{n-1}}\left(\{\omega_n\}\right)$$

אז לכל $\omega_1,\dots,\omega_N\in\Omega^N$:

$$P\left(\{\omega\}\right)=P\left(\bigcap_{n=1}^{N}A_{n,\omega_n}\right)\underbrace{=}_{\text{מהלמה}}\prod_{n=1}^{N}P\left(A_{n,\omega_n}|\bigcap_{k=1}^{n-1}A_{k,\omega_k}\right)=\prod_{n=1}^{N}Q_{\omega_1,\dots,\omega_{n-1}}\left(\{\omega_n\}\right)$$

**מסקנה 3.9**

אם $P$ קיימת היא יחזה כי:

$$P\left(A_{n,\omega_n}|\bigcap_{i=1}^{n-1}A_{i,\omega_i}\right)=Q_{\omega_1,\dots,\omega_n}\left(\{\omega_n\}\right)$$

קובע את הערך על היחידונים.

נגדיר וקטור $(q_\omega)_{\omega\in\Omega^N}$ע"י

$$q_{\omega_1,\dots,\omega_N}=\prod_{n=1}^{N}Q_{\omega_1,\dots,\omega_{n-1}}\left(\{\omega_n\}\right)$$

נגדיר $P$ ע"י:

$$P\left(B\right)=\sum_{b\in B}q_b$$

**למה 3.10**

$q=(q_\omega)_{\omega\in\Omega^N}$ הוא וקטור הסתברות

**הוכחה:** $q_\omega\geq 0$ כי: $Q_{\omega_1,\dots,\omega_n}$ מידת הסתברות. צריך להוכיח: $\sum\limits_{\omega\in\Omega^N}q_\omega=1$ . נחשב:

$$\sum_{\omega\in\Omega^N}q_\omega=\sum_{\omega_1\in\Omega_1}\sum_{\omega_2\in\Omega_2}\cdots\sum_{\omega_N\in\Omega_N}q_{\omega_1,\dots,\omega_n}=\sum_{\omega_1\in\Omega_1}\sum_{\omega_2\in\Omega_2}\cdots\sum_{\omega_N\in\Omega_N}\prod_{n=1}^{N}Q_{\omega_1,\dots,\omega_{n-1}}\left(\{\omega_n\}\right)=$$

$$\sum_{\omega_1\in\Omega_1}\sum_{\omega_2\in\Omega_2}\cdots\sum_{\omega_{N-1}\in\Omega_{N-1}}\prod_{n=1}^{N-1}Q_{\omega_1,\dots,\omega_{n-1}}\left(\{\omega_n\}\right)\cdot\underbrace{\sum_{\omega_N\in\Omega_N}Q_{\omega_1,\dots,\omega_{N-1}}\left(\{\omega_N\}\right)}_{=1}$$

הביטוי האחרון שווה 1 כיוון שהוא שווה ל:

$$Q_{\omega_1,\dots,\omega_{N-1}}\left(\bigcup_{\omega_N\in\Omega_N}\{\omega_N\}\right)=1$$

כי $Q_{\omega_1,\dots,\omega_N}$מידת הסתברות על $\Omega_N$. ממשיכים (ברקורסיה) ונעלים את כל האיברים ולבסוף נקבל כי אכן: $\sum\limits_{\omega\in\Omega^N}q_\omega=1$ ■

**למה 3.11**

לכל $n$ וסדרה $\omega_i \in \Omega_i$ כך ש $i = 1, \ldots, n$ אזי:

$$P\left(\bigcap_{i=1}^{N} A_{i,\omega_i}\right) = \sum_{\omega_{n+1}\in\Omega_{n+1}} \cdots \sum_{\omega_N\in\Omega_N} \prod_{k=1}^{n} Q_{\omega_1,\ldots,\omega_{k-1}}\left(\{\omega_k\}\right)$$

20/11/20 **הוכחה:** הוכחה דומה לקודמת (זה טריוויאלי עבור $n = N$)

$$P\left(\bigcap_{i=1}^{N} A_{i,\omega_i}\right) = \sum_{\eta\in\bigcap_{i=1}^{k} A_{i,\omega_i}} q_\eta$$

$$\bigcap_{i=1}^{N} A_{i,\omega_i} = \{\eta |\, \eta_i = \omega_i,\ i \leq n\}$$

■

נוכיח כי:

$$P\left(A_{n,\omega_n} | \bigcap_{i=1}^{n-1} A_{i,\omega_i}\right) = Q_{\omega_1,\ldots,\omega_{n-1}}\left(\{\omega_n\}\right)$$

נשים לב כי

$$P\left(A_{n,\omega_n} | \bigcap_{i=1}^{n-1} A_{i,\omega_i}\right) = \frac{P\left(\bigcap_{i=1}^{n} A_{i,\omega_i}\right)}{\bigcap_{i=1}^{n-1} A_{i,\omega_i}} = \frac{\sum_{\omega_{n+1},\ldots,\omega_N} \prod_{j=1}^{N} Q_{\omega_1,\ldots,\omega_{j-1}\{\omega_j\}}}{\sum_{\omega_n,\ldots,\omega_N} \prod_{j=1}^{N} Q_{\omega_1,\ldots,\omega_{j-1}\{\omega_j\}}} =$$

$$\text{@\$\%\^{}WRTH@\^{}\%h} = \frac{Q_{\omega_1,\ldots,\omega_{n-1}}\{\omega_n\}}{\sum_{\eta\in\Omega_k} Q_{\omega_1,\ldots,\omega_{n-1}}\left(\{\eta\}\right)} \quad 1$$

■

**הערה 3.12** ביקשנו שיעלו את זה לאתר, אם יהיה לי זמן אני אתקן את זה.

**הערה 3.13** קיבלנו גם מידת הסתברות $P$ על $\Omega_1 \times \Omega_2, \ldots$ כך שמתקיים:

$$P\left(A_{n,\omega_n} | \bigcap_{i=1}^{n-1} A_{i,\omega_i}\right) = Q_{\omega_1,\ldots,\omega_{n-1}}\left(\{\omega_n\}\right)$$

ו$P$ מוגדרת על $\sigma$־אלגברה המכילה את ה $A_{i,\omega_i}$־ ות. במקרה הזה $Pp$ יחיד על ה$\sigma$־אלגברה הנוצרת מה $A_{i,\omega}$

**הערה 3.14** הסתברות מותנית $P\left(A|B\right)$ רציפה ב $A$ ו־$B$ כלומר, אם $A_n \to A$ ו: $B_n \to B$ אזי

$$P\left(A|B\right) = \lim P\left(A_n|B_n\right)$$

**הערה 3.15** כדי $P\left(A|B\right)$ יהיה מוגדר, צריך ש $P\left(B\right) > 0$.
יש שיטות להגדיר גם בחלק מהמקרים ש: $P\left(B\right) = 0$. אבל זה בתורת המידה.

# 4 אי תלות

**הגדרה 4.1 מאורעות בלתי תלויים:** מאורעות $A,B$ הם בלתי תלויים (ב"ת) אם:

$$P(A\cap B)=P(A)\cdot P(B)$$

**הערה 4.2** כאשר $P(B)>0$ אי תלות שקולה ל:

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=P(A)$$

כלומר, אם אנו יודעים ש $B$ התרחש, לא מקבלים "שום מידע" על התרחשות $A$.

**הערה 4.3** כל מאורע $A$ הוא בלתי תלוי מ$\varnothing$ ומ $\Omega$ כי:

$$A\cap\Omega=A,\quad A\cap\varnothing=\varnothing$$

$$\begin{aligned}P(A\cap\Omega)&=P(A)=P(A)\cdot 1=P(A)\cdot P(\Omega)\\ P(A\cap\varnothing)&=P(\varnothing)=0=P(\varnothing)\cdot P(A)\end{aligned}$$

**דוגמה 4.4** בוחרים קלף באופן מקרי (אחיד) מחפיסה

$$\begin{aligned}A&=\{\text{יצא אס}\}\\ B&=\{\text{יצא תלתן}\}\\ \Omega&=\{\text{קלפים}\}\end{aligned}$$

נשים לב כי:

$$|\Omega|=52,\quad |A|=4,\quad |B|=13,\quad |A\cap B|=1$$

$$P(A\cap B)=\frac{1}{52}$$

$$\frac{1}{4}\cdot\frac{1}{13}=P(A)P(B)$$

לכן, $A,B$ בלתי תלויים.

**דוגמה 4.5** מטילים שתי קוביות:

$$\Omega=\{(i,j)\,|\,1\le i,j\le 6\}$$

תחת הסתבות אחידה.

$$\begin{aligned}A&=\{\text{בהטלה הראשונה יצא }4\}\\ B&=\{\text{סכום ההטלות הוא }6\}\end{aligned}$$

$$A\cap B=\{(4,2)\}$$

$$P(A\cap B)=\frac{1}{36}\neq\frac{1}{6}\cdot\frac{5}{36}=P(A)\cdot P(B)$$

**למה 4.6**

$A,B$ בלתי תלויים אס"ם $A^c,B$ בלתי תלויים.

**הוכחה:**

$$P(A^c\cap B)=A^c\cap B=B\backslash A=B\backslash(B\cap A)$$

נשים לב כי:

$$B=(B\cap A)\uplus(B\backslash(B\cap A))$$

ולכן:

$$P(B)-P(B\cap A)=P(B)-P(B)P(A)=P(B)(1-P(A))=P(B)P(A^c)$$

■

**מסקנה 4.7**

אי תלות היא תכונה של החלוקה של $\Omega$ ל$(A,A^c)$ וחלוקה $(B,B^c)$.

נניח נתונים מאורעות: $A_1,\ldots,A_n$ מה זה אי תלות?

אפשרות: הלתי תלויים אם לכל $i\neq j$ אז $A_i,A_j$ בלתי תלויים.

נראה שזה לא נכון.

**דוגמה 4.8** מטילים שתי מטבעות: $\Omega=\{H,T\}^2$ נתונים המאורעות:

$$\begin{aligned}A&=\{\text{הטלה ראשונה }H\}\\B&=\{\text{הטלה שניה }H\}\\C&=\{(H,T),(T,H)\}\end{aligned}$$

נראה שכל זוג הוא בלתי תלוי:

$$P(A)=P(B)=P(C)=\frac{1}{2}$$

$$\begin{aligned}A\cap B&=\{(H,H)\}\\A\cap C&=\{(H,T)\}\\B\cap C&=\{(T,H)\}\end{aligned}$$

$$P(A\cap B)=P(A\cap C)=P(B\cap C)=\frac{1}{4}$$

אבל:

$$P(C|A\cap B)=0$$

כיוון ש:

$$P(C\cap A\cap B)=0$$

כלומר אם אנו יודעים כי $A,B$ קרו, אזי קיבלנו המון מידע על $C$.

**הגדרה 4.9** $A_1,\dots,A_n$ בלתי תלויים אם לכל:

$$I\subseteq\{1,\dots,n\}$$

$$P\left(\bigcap_{i=I}A_i\right)=\prod_{i\in I}P\left(A_i\right)$$

**דוגמה 4.10** למשל עבור 3 קבוצות דורשים אי תלות לכל זוג ($|I|==2$) וגם:

$$P\left(A_1\cap A_2\cap A_3\right)=P\left(A_1\right)P\left(A_2\right)P\left(A_3\right)$$

**הגדרה 4.11 חלוקה של $\Omega$:** חלוקה של $\Omega$ היא סדרת קבוצות זרות שאיחודן $\Omega$.

**הגדרה 4.12** בהינתן מאורעות $A_1,\dots,A_n$ החלוקה שהן מגדירות היא:

$$\pi\left(A_1,\dots,A_n\right)=\left\{\bigcap_{i\in I}A_i\cap\bigcap_{h\in\{1,\dots,n\}\setminus I}A^c\middle|\ I\subseteq\{1,\dots,n\}\right\}$$

**הגדרה 4.13** מאורע $B$ בלתי תלוי במאורעות $A_1,\dots,A_n$ אם:

$$P\left(B|C\right)=P\left(B\right)$$

לכל $C\in\pi\left(A_1,\dots,A_n\right)$

**דוגמה 4.14** הטלה של 2 מטבעות:

$$\begin{aligned}A_1&=\{H\text{ הטלה ראשונה}\}\\A_2&=\{H\text{ הטלה שניה}\}\end{aligned}$$

אזי:

$$\pi\left(A_1,A_2\right)=\begin{cases}\Omega\cap\bigcap_{i=1,2}A_i^c=A_1^c\cap A_2^c & I=\varnothing\\A_1\cap A_2^c & I=\{1\}\\A_2\cap A_1^c & I=\{2\}\\\left(A_1\cap A_2\right)\cap\Omega & I=\{1,2\}\end{cases}$$

**הערה 4.15**

$$P\left(C|\left\{\omega\right\}\right)=\begin{cases}0 & \omega\notin C\\1 & \omega\in C\end{cases}$$

כיוון ש: $\pi\left(A_1,A_2\right)$ היא חלוקה ליחידונים $\Leftarrow$ אזי $C$ בלתי תלויה מ$A_1,A_2$ אס"ם $P\left(C\right)=0,1$.

**טענה 4.16**

$\pi\left(A_1,\dots,A_n\right)$ הוא חלוקה

**הוכחה:** נניח $B, B' \in \pi(A_1, \ldots, A_n)$. אם $B \neq B'$ אז $B$ מוגדר בעזרת קבוצת האינדקסים $I \subseteq \{1, \ldots, n\}$. $B'$ מוגדר ע"י $I' \in \{1, \ldots, m\}$, $I \neq I'$.
אז יש $j \in I$ כך ש $j \notin I'$ או להיפך. נטפל במקרה הראשון:

$$B = \bigcap_{i \in I} A_i \cap \bigcap_{k \in \{1, \ldots, n\} \setminus I} A_k^c$$

$A_j$ מופיע באחד הנחתכים.
כמו כן $A_j^c$ מופיע בביטוי המגדיר את $B'$ לכן:

$$B \cap B' \subset \underbrace{A_j}_{\substack{\cup \\ B}} \cap \underbrace{A_j^c}_{\substack{\cup \\ B'}} = \varnothing$$

המקרה השני דומה.
עכשיו צריך להוכיח כי:

$$\bigcup_{C \in \pi(A_1, \ldots, A_n)} C = \Omega$$

בהנתן: $\omega \in \Omega$ נגדיר:

$$I = \{1 \leq i \leq n, \quad \omega \in A_i baho\} kc$$

שאם $i \notin I$ אז: $\omega \in A_i^c$.
לכן $\omega$ שייך לכל אחד מהמחותכים ב:

$$\bigcap_{i \notin I} A_i \cap \bigcap_{j \in \{1, \ldots, n\} \setminus I} A_j^c \in \pi(A_1 \ldots, A_n)$$

לכן:

$$\omega \in \bigcup_{C \in \pi(A_1, \ldots, A_n)} C$$

אבל $\omega$ שרירותי, לכן זה נכון לכל $\omega \in \Omega$ כלומר:

$$\Omega = \bigcup_{C \in \pi(A_1, \ldots, A_n)} C$$

כנדרש. ■

**למה 4.17**

אם $B_1, \ldots, B_k$ מאורעות זרים ו־ $A$ בלתי תלוי מכל אחד מה־ $B_i$ אז הוא בלתי תלוי מ:

$$B = \bigcup_{i=1}^{n} B_i$$

**הוכחה:**

$$A \cap B = \uplus A_n B_i \Rightarrow P(A \cap B) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A) P(B_i) = P(A) \sum_{i=1}^{n} P(B_i) = P(A) P(B)$$

■

**למה 4.18**

אם $A_1,\ldots,A_n$ בלתי תלויה אז גם: $A_1^c,A_2,\ldots,A_n$ בלתי תלויה

**הוכחה:** כמו במקרה $n=2$. (כתרגיל) ■

**משפט 4.19**

$A_1,\ldots,A_n$ בלתי תלויים אס"ס לכל $A_i$ בלתי תלוי מיתר המאורעות. כלומר $A_i$ בלתי תלוי מ: $\{A_j| \quad j\neq i\}$

**הערה 4.20** בדוגמה 4.8, $C$ לא בלתי תלוי ב$A,B$

**הוכחה:** נניח $A_1,\ldots,A_n$ בלתי תלויות. נרצה להראות ש$A_n$ בלתי תלוי מכל $C\in\pi\left(A_1,\ldots,A_{n-1}\right)$.
(מכיוון שההגדרה של אי־תלות אינווריאנטי תחת שינוי סדר זה בלי הגבלת הכלליות).
צריך להוכיח:

$$P\left(A_n\cap\left(A_1^{x_1}\cap\ldots\cap A_{n-1}^{x_{n-1}}\right)\right)=P\left(A_n\right)\cap P\left(A_1^{x_1}\cap\ldots\cap A_{n-1}^{x_{n-1}}\right)$$

כאשר $x_i$ הוא או "כלום" או "$c$" (כלומר משלים).
מלמה 4.18: $A_1^{x_1},\ldots,A_{n-1}^{x_{n-1}},A_n$ בלתי תלויה שמשפחה של $n$ קבוצות לכן:

$$P\left(A_1^{x_1}\cap\ldots\cap A_{n-1}^{x_{n-1}}\right)=\prod_{i=1}^{n-1}P\left(A_i^{x_i}\right)$$

$$P\left(A_n\cap\left(A_1^{x_1}\cap\ldots\cap A_{n-1}^{x_{n-1}}\right)\right)=P\left(A_n\right)\prod_{i=1}^{n-1}P\left(A_i^{x_i}\right)$$

**<u>כיוון שני:</u>**
נוכיח באינדוקציה, המקרה $n=2$ טריוואלי. נניח שידוע שזה נכון ל$n-1$.

**הערה 4.21** באופן כללי לכל $c\in\left(B_1,\ldots,B_k\right)$ היא איחוד זר של קבוצות ב: $\pi\left(B_1,\ldots,B_{k+1}\right)$.

לכן, מלמה 4.17, לכל $I\subseteq\{1,\ldots,n\}$ גם $\{A_i\}_{i\in I}$ מקיימת שכל אחת $A_i$ $i\in I$ בלתי תלויה מהקבוצות:

$$A_j,\ j\in I\backslash\{i\}$$

כי אם:

$$C\in\pi\left(A_j:\ j\in I,\ j\neq i\right)$$

אז זהו איחוד זר של איברים של $\pi\left(A_j:\ j\neq i\right)$. וכיוון ש $A_i$ בלתי תלויה מכל קבוצה ב:$\pi\left(A_j:\ j\neq i\right)$ היא בלתי תלויה מ$C$.
צריך להוכיח כי: $I\subseteq\{1,\ldots,n\}$

$$P\left(\bigcap_{i\in I}A_i\right)=\prod_{i\in I}P\left(A_i\right)$$

אם $|I|<n$ ע"י הצטמצמות ל $\{A_i\}_{i\in I}$ ושימוש בהנחת האינדוקציה מקבלים את מה שרצינו.
במקרה $I=\{1,\ldots,n\}$ אז:

$$C=A_1\cap\ldots\cap A_{n-1}\in\pi\left(A_1,\ldots,A_{n-1}\right)$$

לפי הנחה, $A_n$ בלתי תלויה מ$C$ ולכן:

$$P\left(A_1\cap\ldots\cap A_n\right)=P\left(C\cap A_n\right)=P\left(C\right)P\left(A_n\right)$$

על $C$ כבר ראינו: $P(C)=\prod_{i=1}^{n-1}P(A_i)$ולכן:

$$=\prod_{i}^{n}P(A_i)$$

■

**הערה 4.22** בדרך כלל מנסחים את המשפט כאשר במקום $\pi(A_j,\, j\neq i)$ מופיעה האלגברה $\mathscr{A}(A_j|\, j\neq i)$ הנוצרת מ $\{A_j,\, j\neq i\}$. כלומר כל הקבוצות שאפשר לרשום בעזרת איחודים, חיתוכים ומשלים מקבוצות המשפחה.
ברור ש:

$$\pi(B_1,\ldots,B_k)\subseteq\mathscr{A}(B_1,\ldots,B_k)$$

אפשר להראות שכל $C\in\mathscr{A}(B_1,\ldots,B_k)$ היא איחוד זר של איברי חלוקה מ$\pi(B_1,\ldots,B_k)$. ולכן מהלמה 4.17 הדברים שקולים.

**הערה 4.23** סדרה אינסופית $A_1,A_2,\ldots$ של מאורעות נקראת בלתי תלויה אם כל תת־משפחה סופית שלה היא בלתי תלויה.

21/11/20

**דוגמה 4.24** הטלת מטבע (ערכים $H,T$) באופן בלתי תלוי.
בכל הטלה הסיכוי ל $H$ הוא: $0\leq p_H\leq 1$.
ואילו הסיכוי ל $T$ הוא: $p_T=1-p_H$
סדרה של $n$ (או $\infty$) הטלות מתואר ע"י מרחב ההסתברות ומסדרת מאורעות $A_1,\ldots,A_n$(או $A_1,A_2,\ldots$) כך ש

1. המאורעות בלתי תלויים
2. $P(A_n)=p_H$

פירוש: $A_n$ המאורע שיצא $H$ בהטלה $n$.
נרצה להראות כי קיים מרחב הסתברות מתאים למאורעות כנ"ל.
בשלב זה נוכיח עבור $N$ סופי.
המרחב מדגם שלנו הוא:

$$\Omega=\{H,T\}^N,\quad \mathcal{F}=2^\Omega$$

נסמן:

$$A_{n,H}=A_n=\{(x_1,\ldots,x_N)\,|\,x_i\in\{H,T\}\,,\,x_n\in H\}$$

נגדיר:

$$P(\{x,\ldots,x_n\})=p_{x_1}p_{x_2}\cdot\ldots\cdot p_{x_N}$$

כרגיל:

$$P(B)=\sum_{\omega\in B}P(\{\omega\})$$

זה נכפה ממה שרוצים להוכיח כי:

$$\{(x_1,\ldots,x_n)\}=\bigcap_{i=1}^{N}A_{i,x_i}$$

נאי תלות:

$$P\left(\bigcap_{i=1}^{N} A_{i,x_i}\right) = P\left(A_{1,x_1}\right)\cdot\ldots\cdot P\left(A_{N,x_N}\right)$$

$$P\left(A_{i,x_i}\right) = P\left(\bigcup_{\omega\in A_{i,x_i}}\{\omega\}\right) = \sum_{\substack{x_j\in\{H,T\}\\ j=1,\ldots,N\\ j\neq i}} P\left(\{(x_1,\ldots,x_N)\}\right) =$$

$$\sum_{x_1\in\{H,T\}}\sum_{x_2\in\{H,T\}}\cdots\sum_{x_N\in\{H,T\}} p_{x_1}p_{x_2}\cdot\ldots\cdot p_{x_N}$$

(לא מופיע סכום על $x_i$)

$$= p_{x_i}\underbrace{\sum_{x_1} p_{x_i}\cdots\underbrace{\sum_{x_{N-1}} p_{x_{N-1}}\underbrace{\sum_{x_N} p_{x_N}}_{p_H+p_T=1}}_{=1}}_{=1} = p_{x_i}$$

לסיכום, ראינו שעם ההגדרה הזו:

$$P\left(A_{i,H}\right) = p_H,\quad P\left(A_{i,T}\right) = p_T = 1-p_H$$

צריך להוכיח כי:

$$\sum_{(x_1,\ldots,x_N)\in\Omega} p\left(\{(x_1,\ldots,x_N)\}\right) = 1$$

הוכחה כמו החישוב הקודם.
נותר להוכיח ש:

$$\{A_{1,H},\ldots,A_i\}\&N$$

מאורעות בלתי תלויים. כלומר לפי המשפט שלכל $i$ $A_i$ בלתי תלוי מ ($A_j$ $j\neq i$)
כלומר שלכל $i$ קבוע, לכל בחירה של $A'_j = A_j$ או: $A'_j = A_j^c$ כאשר $j\neq i$ $A_i$ בלתי תלוי מ:

$$\bigcap_{j\neq 1} A'_j$$

כלומר אם נגדיר $A'_i = A_i$ רוצים:

$$P\left(\bigcap_i A'_i\right) = P\left(A'_1\right)P\left(A'_2\right)\cdot\ldots\cdot P\left(A'_N\right)$$

נשים לב ש: $A_{j,H}^c = A_{j,T}$ לכן בחירה $A_j^c$ או $A_j$ כמוהו לבחור $A_{j,x_j}$ עבור $x_j = H$ או $x_j = T$.
לכן צריך להוכיח, בהנתן $i$ ו: $H = x_i$ לכל בחירה $x_j\in\{H,T\}$ כאשר $j\neq i$:

$$P\left(\bigcap_j A_{j,x_j}\right) = P\left(A_{i,x_i}\right)\cdot\ldots\cdot P\left(A_{N,x_N}\right)$$

כלומר:

$$P\left(\bigcap_{j=1}^{N}A_{j,x_j}\right)=P\left(\{x_1,\dots,x_N\}\right)=p_{x_1}\cdot\dots\cdot p_{x_N}=P\left(A_{i,x_i}\right)\cdot\dots\cdot P\left(A_{N,x_N}\right)$$

**דוגמה 4.25** מטילים מטבע שיוצא ראש בהסתברות $0\le p\le 1$. מה ההסתברות שלא יצא $H$ ב$N$ ההטלות הראשונות?
נבחר מרחב הסתברות ומאורעות $A_n$ ($A_n$ מפורשת - יצא $H$ בהטלה $n$)
כך ש $P\left(A_n\right)=p$ ו: $A_1,\dots,A_n$ בלתי תלויות.
אז:

$$P\left(\text{לא יצא } H \text{ באף הטלה}\right)=P\left(\bigcap_{n=1}^{N}A_i^c\right)=P\left(A_i^c\right)\cdot\dots\cdot P\left(A_N^c\right)\underbrace{=}_{P(A_1)=p\Rightarrow P\left(A_1^c\right)=1-p}\left(1-p\right)^N$$

27/11/20

# 5 סדרות אינסופיות של מאורעות

**דוגמה 5.1 עקרית:** הטלות חוזרות של מטבע (בזמנים $n=1,2,3,\dots$)
בהטלה $n$ ההסתברות ל $H$ היא: $p=p_H$ והסתברות ל $T$ היא: $p_T=1-p_H$.
תיאור מתמטי:
נתון מרחב הסתברות $(\Omega,\mathcal{F},P)$ ומאורעות $A_n$ (מייצגות הטלת $H$ בזמן $n$) כך ש: $P\left(A_n\right)=p$ ו: $A_1,A_2,\dots$ בלתי תלויים
(כלומר לכל $n$ הסדרה $A_1,\dots,A_n$ בלתי תלויה)
ניתן לסמן:

$$\begin{aligned}A_{n,H}&=A_n\\A_{n,T}&=A_n^c\end{aligned}$$

ואז:

$$P\left(A_{n,x}\right)=p_x$$

עבור $x\in\{H,T\}$.

**הערה 5.2** העובדה שמודל זה מתאר בצורה טובה את הסיפור

1. אינטואיטיבת זה מתקבל על הדעת
2. ניסיונית זה מוצלח

**הערה 5.3** במציאות קשה לייצר הרבה הטלות שאמת בלתי תלויות. כלומר, קשה לבנות מכולל אקראי אמיתי, למרות שיש לכך דרישה למשל בסימולציה, הצפנה וכו'

**הערה 5.4** לא משנה מה מרחב המדגם, אפשר לחשוב על: $\Omega=\{H,T\}^{\mathbb{N}}$ לא בהכרח מה שראינו באופן חלקי בשעורים קודמים שקיימית מרחבי הסתברות ומאורעות מתאימים.

בהנתן: $x_m,\dots,x_n\in\{H,T\}$ כאשר $m\le n$:

$$P\left(\text{יצא } x_i \text{ בהטלה } i\right)=P\left(\bigcap_{i=m}^{n}A_{i,x_i}\right)\underbrace{=}_{\text{אי תלות}}\prod_{i=m}^{n}\underbrace{\left(P\left(A_{i,x_i}\right)\right)}_{p_{x_i}}=p_H^{\#\{m\le i\le n,\ x_i=H\}}\cdot p_T^{\#\{m\le i\le n,\ x_i=T\}}$$

מניחים $0 < p_H < 1$:

$$P\left(\text{בדיוק } H \text{ פעמים } k \text{ יצאו } n \text{ עד זמן}\right) = \binom{n}{k} p_H^k p_T^{n-k}$$

$$P\left(H \text{ פעמים } k \text{ יצאו לכל היותר } n \text{ עד זמן}\right) = P\left(\bigcup_{m=0}^{k} \{k \text{ עד זמן } H \text{ פעמים } m \text{ יצאו בדיוק}\}\right) =$$
$$\sum_{m=0}^{k} \binom{n}{m} p_H^m p_T^{n-m}$$

$$P\left(H \text{ פעמים } k \text{ יצאו לכל היותר } n \text{ בכל הזמנים}\right) = P\left(\bigcap_{n=1}^{\infty} B_{k,n}\right)$$

נשים לב כי:

$$B_{1,n} \supseteq B_{k,n+1}$$

אם עד זמן $n+1$ לא קרו יותר מ $k$ פעמים $H$ אז בוודאי זה נכון עד זמן $n$.
כלומר, $\bigcap_{n=1}^{\infty} B_{k,n}$ חיותך יורד, ולכן:

$$P\left(\bigcap_{n=1}^{\infty} B_{k,n}\right) = \lim_{N\to\infty} P\left(B_{k,n}\right)$$

$$0 \le P\left(B_{k,n}\right) \le \sum_{m=1}^{k} \underbrace{\binom{n}{m}}_{\frac{n(n-1)\cdot\ldots\cdot(n-m+1)}{m!}} \left(\frac{p_H}{p_T}\right)^m p_T^n \le \sum_{m=1}^{k} n^m c^m p_T^n$$

כאשר $c = \frac{p_H}{p_T}$.

$$\le k\left(n^k \cdot D^k \cdot p_T^n\right)$$

כאשר $D$ הוא קבוע אחר התלוי אם $c > 1$ אז: $D = c$. אם $c \le 1$ אז $D = \frac{1}{c}$. (תזכורת: $k$ קבוע)

$$= \underbrace{D}_{k \text{ קבוע תלוי ב}} \underbrace{n^k}_{\text{פולינומיאלי ב k}} \underbrace{p_T^n}_{n \text{ אקספוננציאלית ב}} \underset{n\to\infty}{\longrightarrow} 0$$

(בגלל ש $p_T^n$ דועך מהר יותר מכך פולינום).
כלומר ראינו: לכל $k$:

$$\lim_{n\to\infty} P\left(B_{n,k}\right) = 0$$

(כלל סנדוויץ) ולכן כפי שראינו:

$$P\left(\bigcap_{n=1}^{\infty} B_{k,n}\right) = 0$$

וזהו המאורע ששאלנו לגביו. ובאופן כללי:

$$P\left(\text{בכל הזמנים יצאו מספר סופי של } H\right) = P\left(\bigcup_{k=1}^{\infty}\left\{\text{בכל הזמנים יצאו לכל היותר } k \text{ פעמים } H\right\}\right)$$

$$\leq \sum_{k=1}^{\infty}\underbrace{P\left(\text{יצאו לכל היותר } k \text{ פעמים } H\right)}_{0} = 0$$

מה מספר ה$H$ ה"צפוי" עד זמן $n$?

תשובה: $p_H \cdot n \sim$. ראינו שעל מאורע שהסתברות 1

**הערה 5.5** במודל $\Omega = \{H,T\}^{\mathbb{N}}$ יש אינסוף סדרות $\omega \in \Omega$ כך ש $\omega_i = H$ רק מספר סופי של פעמים.
בהסתברות מאורע עם הסתברות 0 נחשב זניח. למשל עדות לכך שאיחוד בן מנייה של כאלה לעולם אינו עם הסתברות חיובית.
באותו אופן, מאורע בעל התסברות 1 נחשב "כמעט הכל"

**הגדרה 5.6 כמעט תמיד:** אומרים על מאורע $A$ שהוא מתרחש כמעט תמיד (כ"ת) אם $P(A) = 1$ (almost surely)
קודם הראינו שכ"ת יש מספר אינסופי של $H$.

**משפט 5.7**

נקבע $0 < \alpha < 1$ אז כמעט תמיד מתקיים שלכל $n$ גדול מספיק יש עד זמן $n$ לפחות $n^{\alpha}$ פעמים $H$.

**הערה 5.8** שימו לב: $\underbrace{p_H \cdot n}_{\text{המספר הצפוי}} \gg n^{\alpha}$ עבור $n$ גדול מספיק.

**הוכחה:**

$$B_{n,k} = \left\{\text{יצאו לכל היותר } k \text{ פעמים } H \text{ עד זמן } n\right\}$$

רוצים להראות שבהסתברות 1 קרו רק מספר סופי של המאורעות: $\underbrace{B_{n,n^{\alpha}}}_{\text{נסמן } C_n}$

כלומר:

$$P\underbrace{\left(\text{קרו רק מספר סופי של } C_n\right)}_{(\limsup C_n)^c} = 1$$

■

## 5.1 בורל־קנטלי

**למה 5.9 הלמה הראשונה של בורל־קנטלי (*Borel-Cantelli*)**

אם $C_n$ מאורעות ו: $\sum_{n=1}^{\infty} P(C_n) < \infty$ אז:

$$P\underbrace{(\limsup C_n)}_{\{\text{קרו } \infty \text{ פעמים } C_n\} = \bigcap_{n=1}^{\infty}\bigcup_{k>n} C_k} = 0$$

**הוכחה:** נסמן:

$$D_n = \bigcup_{k>n} C_k$$

זו סדרה יורדת כיוון שככל ש$n$ גדל, יש באיחוד פחות קבוצות, לכן מרציפות:

$$P\left(\bigcap_n D_n\right) = \lim_{n\to\infty} P\left(D_n\right) = 0$$

אבל:

$$0 \le P\left(D_n\right) \le \sum_{k=n+1}^{\infty} P\left(C_k\right) \underset{n\to\infty}{\longrightarrow} 0$$

כיוון שזהו זנב של טור מתכנס.
מכלל הסנדוויץ:

$$P\left(D_n\right) \to 0$$

כנדרש.
צריך להוכיח כי:

$$\sum_{n=1}^{\infty} \underbrace{P\left(C_n\right)}_{\{\text{עד זמן } n \text{ יצאו לכל היותר } n^\alpha \text{ פעמים } H\}} < \infty$$

כי אז אפשר להשתמש בלמה.

$$P\left(C_n\right) = P\left(\bigcup_{k=0}^{n^\alpha} \{\text{עד זמן } n \text{ יצאו בדיוק } k \text{ פעמים } H\}\right) \le \underbrace{n^\alpha n^{n^\alpha}\left(\frac{p_H}{p_T}\right)^{n^\alpha}}_{\to\infty} \underbrace{p_T^n}_{\to 0}$$

(נוסחה שקיבלנו קודם).
נשים לב כי: $p_T^n$ יורד לאפס אקספוננציאלית, ואילו $n^\alpha n^{n^\alpha}\left(\frac{p_H}{p_T}\right)^{n^\alpha}$ הולך לאינסוף תת אקספוננציאלית ולכן כל הביטוי שואף לאפס (לכן הטור הכללי הולך לאפס אקספוננציאלית ולכן הוא מתכנס).
נראה את זה:

$$= \exp\left(\overbrace{\alpha\ln(n) + n^\alpha \ln(n) + n^\alpha \ln\left(\frac{p_H}{p_T}\right)}^{t_n} + n\underbrace{\ln p_T}_{<0}\right)$$

אז:

$$\frac{t_n}{n} \underset{n\to\infty}{\longrightarrow} 0$$

עבור $n$ גדול מספיק:

$$|t_n| < \frac{|\ln p_T|}{2}$$

עבור אותו $n$ גם:

$$\exp\left(-n\left|\ln p_T\right| + n\frac{t_n}{n}\right) \underbrace{\le}_{\text{כיוון ששואף לאפס, אז קיים } n \text{ מכנ"ל } \frac{t_n}{n} < \frac{|\ln p_T|}{2}} \exp\left(-n\frac{|\ln p_T|}{2}\right) = c^n$$

עבור קבוע $c$ כך ש: $|c|<1$
קיבלנו:

$$P\left(C_n\right)<c^n\Rightarrow\sum P\left(C_n\right)<\infty$$

ולכן לפי בורל־קנטלי

$$P\left(\left\{C_n\text{ מתרחש מספר סופי של פעמים}\right\}\right)=1$$

$$P\left(C_n\text{ קורה מספר אינסופי של פעמים}\right)=0\iff P\left(C_n\text{ קורה מספר סופי של פעמים}\right)=1$$

ולכן:

$$P\left(\text{לכל } n \text{ גדול מספיק יש לכל הפחות } n^\alpha \text{ פעמים } H \text{ עד זמן } n\right)=1$$

כנדרש. ■

**הערה 5.10** בורל קנטלי איננו איפיון. כלומא אם $\sum P\left(C_n\right)=\infty$ אי אפשר להסיק שבהסתברות 1 $C_n$ קורה מספר סופי של פעמים.

**דוגמה 5.11** תבחרו מאורע $C$ שרירותי. $0<P\left(C\right)<1$. נגדיר $C_n=C$ לכל $n$ אז עבור $\omega\in C$ מתקיים $\omega\in C_n$ לאינסוף $n$־ים אבל:

$$\sum_{n=1}^{\infty}P\left(C_n\right)=\sum_{n=1}^{\infty}P\left(C\right)=\infty$$

(זו לא בדיוק דוגמה לשלילה של בורל קונטלי, אבל זו דוגמה לכך שההנחה של הלמה הבאה נכונה)

**למה 5.12 הלמה השניה של בורל־קנטלי (*Borel-Cantelli*)**

אם $C_n$ בלתי תלויים ואם:

$$\sum_{n=1}^{\infty}P\left(C_n\right)=\infty$$

אז בהסתברות 1 קורה אינסוף $C_n$־ים כלומר:

$$P\left(\limsup C_n\right)=1$$

**הערה 5.13** הדוגמה מראה שאי תלות נחוצה עבור הלמה.

**הוכחה:**

$$P\left(C_n\text{ קרו }\infty\text{ פעמים}\right)=P\left(\bigcap_{m=k}^{\infty}\left\{C_m\text{ לא קרה}\right\}\right)=\lim_{T\to\infty}=P\left(\bigcap_{m=k}^{T}C_m^c\right)=\lim_{T\to\infty}\prod_{m=k}^{T}P\left(C_m^c\right)\to 1-P\left(C_m\right)$$

נשתמש ב $1-x\leq e^{-x}$:

$$\prod_{m=k}^{T}\left(1-P\left(C_m\right)\right)\leq\prod_{m=k}^{T}e^{-P(C_m)}=\exp\left(-\underbrace{\sum_{m=k}^{T}P\left(C_m\right)}_{\text{זנב של טור חיובי אשר שואף ל}\infty\ \xrightarrow[T\to\infty]{}\infty}\right)$$

לכל $k$:

$$\sum_{m=k}^{T}P\left(C_m\right)\xrightarrow[T\to\infty]{}\infty$$

כלומר:

$$P\left(\text{לא קרה }C_m\text{ אחרי זמן }k\right)=0$$

ולכן:

$$P\left(\text{קרו מספר סופי של }C_m\text{־ים}\right)=P\left(\bigcup_{k=1}^{\infty}\left\{\text{לא קרה }C_m\text{ עבור }m>k\right\}\right)\leq$$

$$\sum_{k=1}^{\infty}\underbrace{P\left(\text{לא קרה }C_m\text{ אחרי זמן }k\right)}_{=0}=0\Rightarrow P\left(\text{קרו }\infty\text{ פעמים }C_n\right)=1$$

כנדרש. ■

**דוגמה 5.14** כזכור:

$$A_{n,H}=\left\{\text{יצא }H\text{ בהטלה ה־}n\right\}$$

נשים לב כי:

$$P\left(A_{n,H}\right)=p_H>0$$

לכן:

$$\sum_{1}^{\infty}\underbrace{P\left(A_{n,H}\right)}_{p_H>0}=\infty$$

לכן לפי הלמה, בהסתברות 1 קרו אינסוף $A_n$־ים מתקיימים.

**דוגמה 5.15** האם יתנו לקוף מכונת כתיבה, והוא יקליד אקראית לגמרי על המכונה, מה הסיכוי שהוא יקליד את כל מחזות שקספיר?

$S=[\ldots]$ מחרוזת באורך $N$ שכתוב בה כל מחזות שקספיר.

ההסתברות שהוא יקליד את $S$ היא קטנה מאוד אבל קיימת, וכיוון שהקוף מקליד אינסוף תווים היא תקרה לבסוף, ולא רק זה, היא תקרה אינסוף פעמים.

**דוגמה 5.16** $p_T = p_H = \frac{1}{2}$ נאמר שקרה רצף באורך $l$ בזמן $n$. אם בזמנים: $n, \ldots, n+l$ יצא $H$.

**משפט 5.17**

בהסתברות 1, יהיו אינסוף זמנים $n$ כך שעד זמן $n$ נראה רצף $\lceil \log_2 n \rceil$, ואילו לכל $\varepsilon > 0$ יש הסתברות 0 שלאינסוף $n$-ים נראה עד זמן $n$ רצף באורך:

$$\lceil (1+\varepsilon) \log_2 n \rceil$$

**הוכחה:** של חלק ראשון (חלק שני נראה משהו בתרגול):

$$B'_n = \{\text{עד זמן } n \text{ ראינו רצף באורך } \lceil \log_2 (n) \rceil\}$$

אנו רוצים להראות שבהסתברות 1 אינסוף $B'_n$-ים קורים.

אלה לא מאורעות בלתי תלויים, כיוון שאם יש רצף ארוך מוקדם בסדרה זה ישפיע על קיום $B'_n$ ל$n$-ים הבאים.

נגדיר:

$$B_k = \left\{\text{קרה רצף באור } k \text{ בין הזמנים } 2^{k-1} \text{ ו-} 2^{k-1} \text{כולל}\right\}$$

$B_k$-ים שונים נקבעים ע"י הטלות שונות ולכן בלתי תלויים. נראה זאת בשיעורי בית.

כמו כן:

$$B_k \subseteq B'_{2^k}$$

לכן, כדי שיתרחשו אינסוף $B'_n$-ים, די להראות שמתרחשים $\infty$ $B_k$-ים.

כעת מספקי להראות ש:

$$\sum_{k=1}^{\infty} P(B_k) = \infty$$

צריך להעריך $P(B_k)$ מלמטה. נחלק את הזמנים $2^{k-1}$ עד $2^k$ לבלוקים באורך $k$ (ייתכן שתשאר קצת שארית, אבל לא אכפת לנו)

נקבע $k$:

$$s = \left\lfloor \frac{2^k - 2^{k-1}}{k} \right\rfloor = \left\lfloor \frac{2^{k-1}}{k} \right\rfloor$$

נסמן:

$$B_{k,j} = \{\text{קרה רצף באורך } k \text{ בזמן } t_j\}$$

ברור:

$$\bigcup_{j=1}^{s} B_{k,j} \subseteq B_k$$

$$P(B_k) \geq 1 - P\underbrace{\left(\bigcap_{j=1}^{s} B_{k,j}^c\right)}_{\left(\bigcup_{j=1}^{s} B_{k,j}\right)^c}$$

עבור $i \neq j$ $B_{k,i}, B_{k,j}$ תלויים בהטלות שונות ולכן בלתי תלויות. כמו כן:

$$\underbrace{P\left(B_{k,j}^{c}\right)}_{\text{הסתברות שלא יצא רצף באורך } k} = 1 - 2^{-k}$$

לכן:

$$P\left(B_k\right) \geq 1 - \prod_{j=1}^{c}\left(1 - 2^{-k}\right) = 1 - \left(1 - 2^{2-k}\right)^{2^k/2k}$$

נזכור מאינפי: $\left(1 - \frac{1}{x}\right)^x \underset{x\to\infty}{\longrightarrow} \frac{1}{e}$ ולכן נקבל:

$$= 1 - \left(\left(1 - \frac{1}{2^k}\right)^{2^k}\right)^{1/2k}$$

ולכן עבור $k$ גדול מספיק $1 > \underbrace{c}_{\sqrt{\frac{1}{e}}} > \left(1 - \frac{1}{2^k}\right)^{2^k}$ ן

$$\geq 1 - \left(\sqrt{\frac{1}{e}}\right)^{1/2k}$$

נגדיר $f\left(y\right) = \left(\frac{1}{\sqrt{e}}\right)^y$ ולכן למעשה כתוב לנו:

$$= f\left(0\right) - f\left(\frac{1}{2k}\right) \underbrace{=}_{\text{משפט ערך הביניים}} \ln\left(\frac{1}{e}\right)\ln\left(\frac{1}{\sqrt{e}}\right)^{c_k} \cdot \frac{1}{2k}$$

כאשר $c_k \in \left[0, \frac{1}{2k}\right]$. עבור $k \to \infty$ מתקיים $c_k \to 0$ ולכן:

$$\left(\frac{1}{\sqrt{e}}\right)^{c_k} \to 1$$

לכן:

$$P\left(B_k\right) \geq \frac{1}{4k}$$

כלומר:

$$\sum P\left(B_k\right) \geq \frac{1}{4}\sum\frac{1}{k} = \infty$$

■

28/11/20

**מסקנה 5.18**

אם $C_n$ סדרה בלתי תלויה אז:

$$P\left(\infty \text{ מהם קורי}\right) = 0 \vee 1$$

(לא ייתכן חצי)

זהו מקר פרטי של:

**משפט 5.19 0־1 של קולמוגורב**

אם $C_n$ מאורעות בלתי תלויים.
אם $A$ מאורע המוגדר ע"י $C_n$־ים ($A$ נמצא ב $\sigma$־אלגברה הנוצאת מ $\{C_n\}$)
וגם $A$ אסימפטוטי במובן שאיננו תלוי באף רישא של הסדרה (לכל $N$, $A$ גם ניתן לביטוי ע"י $C_N, C_{N+1}, \ldots$)
אז $P(A) = 0 \vee 1$

**דוגמה 5.20 מאורע אסימפטוטי שאיננו $\liminf$ או $\limsup$:**

$$A = \{\omega : \text{קורה } C_n \ \infty \text{ פעמים אבל הפרש הזמנין בין המאורעות שהתרחשו שואף לאינסוף}\}$$

**הערה 5.21** לא הספקנו לדבר על "Gambler's ruin". אבל עשינו משהו בסגנון לכן נמשיך הלאה.

# 6 משתנים מקריים

**הגדרה 6.1 משתנה מקרי:** (זמנית)
משתנה מקרי הוא פונקציה $x : \Omega \to \Delta$ כאשר $(\Omega, \mathcal{F}, P)$ מרחב הסתברון נתון ו$\Delta$ קבוצה.

**דוגמה 6.2** מטילים שתי קוביות: $\Omega = \{1, \ldots, 6\}^2$.
נגדיר: $X : \Omega \to \{1, \ldots, 6\}$ באופן הבא: $X(i,j) = i$. אז $X(\omega)$ הוא הערך של ההטלה הראשונה.
אפשר להגדיר $y(i,j) = i + j$ (כלומר סכום ההטלות)
זה משתנה מקרה, הטווח המובלע הוא $\{2, \ldots, 12\}$ או $\mathbb{N}$.
אפשר להגדיר מאורעות:

$$\{\omega \in \Omega |\ X(\omega) \in \{1,2,3\}\} = \{\omega \in \Omega |\ \omega = (i,j)\ i \in \{1,2,3\}\}$$

אפשר לקצר ולכתוב:

$$\{\omega :\ X(\omega) \in \{1,2,3\}\}$$

או אפילו לקצר יותר ולכתוב:

$$\{1 \le X \le 3\}$$

**דוגמה 6.3** נטיל $N$ מטבעות בלתי תלויות שמקבלות ערך $H$ בהסתברות $p$ כל אחת. $\Omega =^N$. $P$ הסתברות מתאימה.
נגדיר:

$$X_i(\omega) = \begin{cases} 1 & \omega_i = H \\ 0 & \omega_i = T \end{cases}$$

אלה משתנים מקריים עם ערכים ב $\{0,1\}$.
נגדיר: $S = X_1 + X_2 + \ldots + X_N$.
זה משתנה מקרי עם ערכים ב $\{0, \ldots, N\}$.

$$\{\omega :\ \#\{1 \le i \le N :\ X_i = H\} > N^\alpha\}$$

או בכתיב אחר:

$$\{S > N^\alpha\}$$

או בכתיב נוסף:

$$\{\omega :\ S(\omega) > N^\alpha\}$$

רוצים שאם $A$ מאורע ב$\Delta$ אז

$$\{\omega \in X(\omega) \in A\}$$

תהיה מאורע.

**הגדרה 6.4 משתנה מקרי:** (אמיתית)
יהיו $(\Omega,\mathcal{F})$ ו: $(\Delta,\mathcal{G})$ מרחבים מדידים ($\mathcal{F}$ היא $\sigma$-אלגברה על $\Omega$ ו $\mathcal{G}$ היא $\sigma$-אלגברה על $\Delta$)
פונקציה $X:\Omega\to\Delta$ נקראת משתנה מקרי (פונקציה מדידה) אם לכל $A\in\mathcal{G}$ הקבוצה:

$$X^{-1}(A)=\{\omega\in\Omega|\quad X(\omega)\in A\}$$

הוא מאורע ב $\Omega$, כלומר:

$$X^{-1}(A)\in\mathcal{F}$$

**למה 6.5**

תהי $X:\Omega\to\Delta$

1. $X^{-1}(\Delta)=\Omega$
2. אם $A\subseteq\Delta$ אז $X^{-1}(A^c)=\left(X^{-1}(A)\right)^c$
3. אם $A,B\subseteq\Delta$ אז אם $A\cap B=\varnothing$ אז: $X^{-1}(A)\cap\left(X^{-1}(B)\right)=\varnothing$
4. אם $A_1,A_2,\ldots\subseteq\Delta$ אז:

$$X^{-1}\left(\bigcup_{n=1}^{\infty}A_n\right)=\bigcup_{n=1}^{\infty}X^{-1}(A_n)$$
$$X^{-1}\left(\bigcap_{n=1}^{\infty}A_n\right)=\bigcap_{n=1}^{\infty}X^{-1}(A_n)$$

**הוכחה:** נוכיח רק את 1,2,3 את 4 נשאיר כתרגיל

.1

$$X^{-1}(\Delta)\overbrace{=}^{\text{הגדרה}}\{\omega\in\Omega|\,X(\omega)\in\Delta\}\overbrace{=}^{\text{כי } X \text{מקבל ערכים ב } \Delta}\Omega$$

.2

$$\left(X^{-1}(A)\right)^c=\{\omega|\,X(\omega)\in A\}^c=\{\omega|\,X(\omega)\notin A\}$$

כיוון ש $X(\omega)\in\Delta$ נקבל כי:

$$=\{\omega|\,X(\omega)\in\Delta\backslash A\}=X^{-1}(A^c)$$

3. מניחים $A\cap B=\varnothing$. אם $\omega\in X^{-1}(A)$ אז $X(\omega)\in A$ לכן $X(\omega)\notin B$ (כי $A\cap B=\varnothing$) לכן $\omega\notin X^{-1}(B)$.
4. תרגיל

■

**הגדרה 6.6** יהי $(\Omega,\mathcal{F},P)$ מרחב הסתברות. $(\Delta,\mathcal{G})$ מרחב מדיד. $X:\Omega\to\Delta$ משתנה מקרי.
ההתפלגות של $X$ הוא הפונקציה: $P_x:\mathcal{G}\to\mathbb{R}$ המוגדרת ע"י:

$$P_x\left(A\right)=P\left(X^{-1}\left(A\right)\right)$$

**טענה 6.7**

$P_x$ היא פונקציית הסתברות על $(\Delta,\mathcal{G})$.

**הוכחה:** נבדוק את האקסיומות:

1.

$$P_x\left(A\right)=P\left(X^{-1}\left(A\right)\right)\in[0,1]$$

כי $P$ פונקציית הסתברות.

2.

$$P_x\left(A^c\right)=P\left(X^{-1}\left(A^c\right)\right)=P\left(\left[x^{-1}\left(A\right)\right]^c\right)=1-P\left(X^{-1}\left(A\right)\right)=1-P_x\left(A\right)$$

**הערה 6.8** זו בכל ללא הייתה אקסיומה, אבל הוכחנו גם את זה... אורי מאושרת.

3.

$$P_x\left(\Delta\right)=P\underbrace{\left(X^{-1}\left(\Delta\right)\right)}_{\Omega}=1$$

4. אם $A_1,A_2,\ldots$ זרים (כלומר $A_i\cap A_j=\varnothing$ עבור $i\neq j$) אז $X^{-1}\left(A_i\right)\cap X^{-1}\left(A_j\right)=\varnothing$ עבור $i\neq j$. לפי הלמה. ולכן: $X^{-1}\left(A_1\right),X^{-1}\left(A_2\right),\ldots$ זרים ולכן:

$$P_x\left(\bigcup_{n=1}^{\infty}A_n\right)=P\left(X^{-1}\left(\bigcup_{n=1}^{\infty}A_n\right)\right)\overbrace{=}^{\text{למה}}P\left(\bigcup_{n=1}^{\infty}X^{-1}\left(A_n\right)\right)=\sum_{n=1}^{\infty}P\left(X^{-1}\left(A_n\right)\right)=\sum_{n=1}^{\infty}P_x\left(A_n\right)$$

■

**הערה 6.9**

1. מדידות (העובדה ש $X^{-1}\left(A\right)\in\mathcal{F}$ אם $A\in\mathcal{G}$ ) דרוש כדי ש $P_x$ תהייה מוגדרת כי כדי להפעיל את $P$ על $X^{-1}\left(A\right)$ צריך $X^{-1}\left(A\right)$ תהיה מאורע.
2. נשים לב ש: $P_x=P_0X^{-1}$ ($X$ היא פונקציה מתת קבוצה של $\Delta$ לתת קבוצה של $\Omega$)
3. אם $\Omega$ סופית\בת־מנייה ו: $\mathcal{F}=2^{\Omega}$ אז התנאי של מדידות (התנאי: $A\in\mathcal{G}\Rightarrow X^{-1}\left(A\right)\in\mathcal{F}$) מתקיים תמיד. כלומר במקרה זה, כל פונקציה $X:\Omega\to\Delta$ (לכל $\Delta$) הוא משתנה מקרי

04/12/20

**הערה 6.10** התואר "מדיד" כעת יכול להתייחס גם לקבוצה וגם לפונקציה

**הערה 6.11** (למי שלומד טופולוגיה או אינפי מתקדם)
תנאי מדידות דומה לתנאי של רציפות: מקור של פתוחה היא פתוחה.

היום נתרכז במקרה בו: $\Delta$ סופי או בן מנייה ו: $\mathcal{G}=2^{\Delta}$.
בהנתן משתנה מקרה $X:\Omega\to\Delta$, ההתפלגות של $X$ הוא:

$$P_x:\mathcal{G}\to[0,1]$$

$$P_x\left(A\right)=P\left(\left\{\omega:X\left(\omega\right)\in A\right\}\right)$$

ראינו כי זו מידת הסתברות על $(\Delta,\mathcal{G})$.
אם $|\Delta|\le\aleph_0$ אז: $P_x$ מאופיין ע"י ההסתברות של יחידונים. נסמן:

$$p_x\left(a\right)=P_x\left(\left\{\alpha\right\}\right)$$

כך ש $a\in\Delta$ אז:

$$P_x\left(A\right)=\sum_{a\in A}p_x\left(a\right)$$

אם $X,Y$ משתנים מקריים כך ש:

$$P_X=P_Y$$

(בפרט אותו טווח)
אומרים ש $X,Y$ שווי התפלגות. משמעות: אין סיבה להבדיל בין $X$ ו$Y$ עבור רוב המטרות.

## 6.1 משתנה ברנולי

**הגדרה 6.12 ברנולי:** משתנה מקרי $X$ נקרא ברנולי עם פרמטר $p$ אם $X$ מקבל ערכים $0,1$:

$$\begin{aligned}p_x\left(1\right)&=p\\ p_x\left(0\right)&=1-p\end{aligned}$$

אז מסמנים:

$$X\underbrace{\sim}_{\text{מפולג כמו}}B\left(p\right)$$

**דוגמה 6.13** $\Omega=\{0,1\}$ כך ש: $P\left(\{0\}\right)=1-p$, $P\left(\{1\}\right)=p$, $X:\Omega\to\{0,1\}$, $X\left(\omega\right)=\omega$

**דוגמה 6.14** אם $(\Omega,\mathcal{F},P)$ מרחב הסתברות $A\in\mathcal{F}$. יהי:

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

$x$ מקבל ערכים $\{0,1\}$. מדידות:

$$\begin{aligned}X^{-1}\left(\{0,1\}\right)&=\Omega\in\mathcal{F}\\ X^{-1}\left(\varnothing\right)&=\varnothing\in\mathcal{F}\\ X^{-1}\left(0\right)&=A^c\in\mathcal{F}\\ X^{-1}\left(1\right)&=A\in\mathcal{F}\end{aligned}$$

,

$$P_x(1)=P_x(\{1\})=P\left(X^{-1}(A)\right)=P(A)$$

כלומר: $\mathbb{1}_A$ הוא ברנולי עם פרמטר $P(A)$.

**דוגמה 6.15** הטלת קוביה הוגנת:

$$\Omega=\{1,\dots,6\}\ A=\{\text{יצא 1 או 2}\}$$

$$X(\omega)=\mathbb{1}_A(\omega)=\begin{cases}1 & \omega=1\vee\omega=2\\ 0 & \text{אחרת}\end{cases}$$

$$p_x(1)=P\left(X^{-1}(1)\right)=P(\{1,2\})=\frac{2}{6}=\frac{1}{3}$$

## 6.2 משתנה בינומי

**הגדרה 6.16** משתנה מקרה $X$ המקבל ערכים טבעיים (כולל 0) כך ש:

$$p_x(k)=\begin{cases}\binom{n}{k}p^k(1-p)^{n-k} & k\le n\\ 0 & k>n\end{cases}$$

נקרא בינומי עם פרמטרים: $p\in[0,1]$, $n\in\mathbb{N}$. (כלומר, $n,p$ קבועים נתונים).
מסמנים: $X\sim B(n,p)$

**הערה 6.17**

$$\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=(p+(1-p))^n=1$$

(לפי נוסחת הבינום)
ולכן: $p_x$ בהגדרה אכן וקטור התסברות על $\mathbb{N}$.

**דוגמה 6.18** נטיל $n$ מטבעות בלתי תלויים אשר יוצאים $H$ בהסתברות $p$. נבחר:

$$\Omega=\{H,T\}^n$$

$$P\underbrace{(\{\omega\})}_{(\omega_1,\dots,\omega_n)}=p^{\#\{i|\,\omega_i=H\}}\cdot(1-p)^{\#\{i|\,\omega_i=T\}}$$

נגדיר: $X:\Omega\to\mathbb{N}$ כך ש:

$$X(\omega)=\#\{i|\,\omega_i=H\}$$

מדידות מתקיימת אוטומטית כי $X$ מוגדר על: $\left(\Omega,2^{\Omega}\right)$

$$p_x(k)=P_x(\{k\})=P(\omega|\,X(\omega)=k)=\sum_{\omega\in\Omega|\,\omega\text{-יש }k\text{ פעמים }H\text{ ב}}P(\{\omega\})=\binom{n}{k}\cdot p^k(1-p)^{n-k}$$

כאשר $\binom{n}{k}$ הוא מספר האיברים בסכום. (מספר סדרות של $H,T$ באורך $n$ עם $k$ קואורדינטות $H$)

**מסקנה 6.19**

$X$ משתנה מקרי בינומי עם פרמטרים $n,p$.

**טענה 6.20**

אם $X\sim B\left(n,p\right)$ אז $p_x\left(\right)$ עולה בין $0$ ל: $\lfloor\left(n+1\right)p\rfloor-1$ ואח"כ יורד.

**הוכחה:** נחשב

$$\frac{p_x\left(k\right)}{p_x\left(k+1\right)}=\frac{\frac{n!}{k!(n-k)!}p^k\left(1-p\right)^{n-k}}{\frac{n!}{(k+1)!(n-k-1)!}p^{k+1}\left(1-p\right)^{n-k-1}}=\frac{k+1}{n-k}\cdot\frac{1-p}{p}\overset{?}{\lessgtr}1$$

נעביר אגפים ונקבל:

$$k-\cancel{kp}+1-p\overset{?}{\lessgtr}np-\cancel{kp}$$

$$k+1\overset{?}{\lessgtr}\left(n+1\right)p$$

ולכן נקבל:

$$\begin{aligned}k+1 &< \left(n+1\right)p\\ k &\leq \lfloor\left(n+1\right)p\rfloor-1\end{aligned}$$

■

**משפט 6.21 משפט סטירלינג**

$$n!\sim\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

כלומרף

$$\frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}\underset{n\to\infty}{\longrightarrow}1$$

**הוכחה:** לא נוכיח את זה כאן. אבל רמז: כדאי לחשב את הלוגריתם, כלומר:

$$\ln\left(n!\right)=\sum_{k=1}^{n}\ln k\approx\int_1^n\ln x\mathrm{d}x$$

לא נקבל את הקירוב של סטרילנג אלא רק קירוב לסדר של $\left(\frac{n}{e}\right)^n$. יש הוכחות בוויקיפדיה. ■

**דוגמה 6.22** מטילים $2n$ מטבעות הוגנים כך ש $p=\frac{1}{2}$. מה ההסתברות שיוצאים בדיוק $n$ פעמים $H$?

**פתרון:**

יהי $X\sim B\left(2n,\frac{1}{2}\right)$

$$\underbrace{P\left(X=n\right)}_{p_x\left(\frac{1}{2},\,P\left(\omega:\,x(\omega)=n\right)\right)}=\binom{2n}{n}\left(\frac{1}{2}\right)^n\left(1-\frac{1}{2}\right)^n=\frac{(2n)!}{(n!)^2}\left(\frac{1}{2}\right)^{2n}$$

לפי קירוב סטירלינג:

$$\sim \frac{\sqrt{2\pi 2n}\left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}\right)^{2}}\cdot\frac{1}{2^{2n}}=\frac{\sqrt{2\pi 2n}}{\left(\sqrt{2\pi n}\right)^{2}}=\frac{1}{\sqrt{\pi n}}\underset{n\to\infty}{\longrightarrow}0$$

**הערה 6.23** ובשביל הפיסיקאים: קוביה עם $2n$ דיפולים מגנטיים כך שההסתברות של כל דיפול להיות $\uparrow$ היא $p=\frac{1}{2}$ מה ההסתברות שיוצאים בדיוק $n$ פעמים $\uparrow$?

## 6.3 משתנה פואסון

**הגדרה 6.24 משתנה פואסון:** יהי $\alpha>0$. משתנה מקרי $X$ עם ערכים ב $\mathbb{N}$. נקרא משתנה פואסון עם פרמטר $\alpha$, $X\sim\text{Pois}(\alpha)$ אם:

$$p_X(\{k\})=\frac{\alpha^k}{k!}e^{-\alpha}$$

**הערה 6.25**

$$e^{-\alpha}\sum_{k=0}^{\infty}\frac{\alpha^k}{k!}=e^{-\alpha}e^{\alpha}=1$$

ולכן $p_x$ באמת וקטור הסתברות.

**דוגמה 6.26 התפרקות רדיואקטיבית:** נניח שנתון חומר רדיואקטיבי (המון אטומים). בפרק זמן באורך $\Delta t$ מספר החלקיקים שמתפרקים הוא $\alpha\Delta t$.

ההתפרקות בפרקי זמן זרים, בלתי תלויים. יהי $z$ מספר החלקיקים שהתפרקו מזמן 0 עד 1 . מה ההתפלגות של $z$.

נבחר $n\in\mathbb{N}$ ופרק את קטע הזמן לקטעים באורך $\frac{1}{n}$. בכל קטע:

$$I_i=\left[\frac{i}{n},\frac{i+1}{n}\right)$$

בכל קטע כזה נבצע ניסוי ברנולי בהסתברות: $\frac{\alpha}{n}$ (אורך הקטע) : $z:=\#\{\text{ניסויים שצלחו}\}$

$$z\sim B\left(n,\frac{\alpha}{n}\right)$$

נחשב את:

$$p_z(k)=\binom{n}{k}\left(\frac{\alpha}{n}\right)^k\left(1-\frac{\alpha}{n}\right)^{n-k}=\frac{\frac{1}{k!}\left(n(n-1)\cdot\ldots\cdot(n-k+1)\right)\left(1-\frac{\alpha}{n}\right)^n}{\underbrace{\left(\frac{\alpha}{n}\right)^{-k}\left(1-\frac{\alpha}{n}\right)^k}_{\left(\frac{n}{\alpha}-1\right)^k}}=$$

$$\frac{1}{k!}\left(1-\frac{\alpha}{n}\right)^n\cdot\frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{\underbrace{(n-\alpha)(n-\alpha)\cdot\ldots\cdot(n-\alpha)}_{k}}$$

מה קורה כש: $n\to\infty$?

$$\forall 0\le i\le k-1\quad \frac{n-i}{n-\alpha}\to 1$$

$$\left(1-\frac{\alpha}{n}\right)^n \underset{n\to\infty}{\longrightarrow} e^{-\alpha}$$

קבלנו כי לכל $k$:

$$p_{z^{(n)}(k)} \underset{n\to\infty}{\longrightarrow} \frac{\alpha^k}{k!}e^{-\alpha}$$

הסתברות לפי משתנה פואסון עם התפלגות $\alpha$ של $k$.

**דוגמה 6.27** יהיו נתונים $n$ כדורים ממוספרים $1,\ldots,n$ מסדרים אותם בסדר מקרי. יהי $W$ מספר הכדורים שחזרו למקום המקורי.

$$W = \#\left\{i \middle| \text{הכדור ה}i\text{נשאר במקום}\right\}$$

הטענה היא כי לכל $k\in\mathbb{N}$:
ההסתברות של $k$ תחת:

$$p_{W^{(n)}}(k) \underset{n\to\infty}{\longrightarrow} \frac{1}{k!}e^{-1} = \text{Pois}(1)$$

("Inttentive secretary")

## 6.4 משתנה גיאומטרי

**הגדרה 6.28 משתנה גיאומטרי:** יהי $0<p<1$ משתנה מקרי $X$ על $\{1,2,3,\ldots\}$ כך ש:

$$p_x(k) = p(1-p)^{k-1}$$

נקרא משתנה גיאומטרי עם פרמטר $p$:

$$X \sim \text{Geom}(p)$$

**דוגמה 6.29** מטילים $\infty$ מטבעות עם הסתברות $p$ ל$H$ בכל אחת. (בלתי תלויים)

$$\Omega = \{H,T\}^{\mathbb{N}^+}$$

$$x_i\in\{H,T\} \quad A_{i,x_i} = \{\text{בהטלה ה }i\text{ יצא }x_i\}$$

הסתברות כנדרש.
נגדיר:

$$X\underbrace{(\omega)}_{(\omega_1,\omega_2,\ldots)} = \min\left\{i \middle| \omega_i = H\right\}$$

כמה זמן מחכים עד להטלת הראש הראשונה.

**הערה 6.30** נגדיר $X(T,T,T,\ldots)=\infty$ אבל נשים לב כי: $\{(T,T,T,\ldots)\}$הוא בעל הסתברות 0

$$p_x(k) = P\left(\left\{\bigcap_{i<k} A_{i,T}\cap A_{k,H}\right\}\right) = (1-p)^{k-1}\cdot p$$

לכן:$X\sim\text{Geom}(p)$.

**הערה 6.31** קל לראות ש:

$$\sum_{k=1}^{\infty} p\left(1-p\right)^{k-1}=1$$

## 6.5 תכונת חוסר זיכרון

אם $X$ משתנה מקרי $\sim \mathrm{Geom}\left(p\right)$ ו: $k\in\mathbb{N}$ קבוע. עבור $n>k$ יהי

$$\begin{aligned} A_n &= \{x=n\} \\ B_k &= \{x\geq k\} \end{aligned}$$

$$P_X\left(A_n|B_k\right)=\frac{P_X\left(A_n\cap B_k\right)}{P_X\left(B_k\right)}\underbrace{=}_{A_n\subseteq B_k}\frac{P_X\left(A_n\right)}{P_X\left(B_k\right)}=\frac{\not p\left(1-p\right)^{n-1}}{\sum\limits_{m=k}^{\infty}\not p\left(1-p\right)^{m-1}}=p\left(1-p\right)^{(n-m)-1}$$

כיוון ש:

$$\sum_{m=k}^{\infty}\alpha^m=\frac{-\alpha^k}{\alpha-1}\underbrace{=}_{\alpha=1-p}\frac{1}{p}\left(1-p\right)^k$$

05/12/20 ולכן, זו ההסתברות של $n-k$ תחת $\mathrm{Geom}\left(p\right)$

**הגדרה 6.32 משתנה בדיד:** הטווח בן מנייה וה$\sigma$־אלגברה בטווח מכילה את כל התת קבוצות.

יהיו $X,Y$ משתנים מקריים, מה הפירוש של: $P\left(X\in A,Y\in B\right)$?
השאלה אינה מוגדרת היטב. כיוון ש:

$$P\left(\omega\in\Omega:\ X\left(\omega\right)\in A,\ Y\left(\omega\right)\in B\right)$$

כיוון שלא נאמר ש$X,Y$ משתנים מקריים המוגדרים על אותו מרחב הסתברות. לצורך העניין $A$ הוא ניסוי א' ו$B$ הוא ניסוי $B$ לא בהכרח יש קשר בין הניסויים.
נניח:

$$X:\Omega\to\Delta_1,\quad Y:\Omega\to\Delta_2$$

כרגיל $\left(\Omega,\mathcal{F},P\right)$ מרחב הסתברות. $\left(\Delta_1,\mathcal{G}_1\right),\left(\Delta_2,\mathcal{G}_2\right)$ מרחבים מדידים.
כעת:

$$A\in\mathcal{G}_1,\ B\in\mathcal{G}_2,\ P\left(X\in A,Y\in B\right)=P\left(\omega\in\Omega:\ X\left(\omega\right)\in A,\ Y\left(\omega\right)\in B\right)= P\left(X^{-1}\left(A\right)\cap Y^{-1}\left(B\right)\right)$$

נניח שאנו יודעים את ההתפלגויות $P_X,P_Y$, מה אפשר להגיד על:

$$P\left(X\in A,Y\in B\right)=?$$

לפי הגדרה:

$$\begin{cases} P\left(X\in A\right)=P_X\left(A\right) \\ P\left(Y\in B\right)=P_Y\left(B\right) \end{cases}$$

אנו עדיין לא יודעים להגיד הרבה על $P\left(X\in A,Y\in B\right)$ נראה זאת בדוגמאות הבאות.

**דוגמה 6.33** נטיל מטבע עם הסתברות $p$ לראש. $\Omega=\{H,T\}$ וכו'.

$$Y=X=\begin{cases}1 & \omega=H\\ 0 & \omega=T\end{cases}$$

$$P\left(X\in\{1\},Y\in\{1\}\right)=P\left(X=1,Y=1\right)=P\left(\omega\in\Omega:\ X\left(\omega\right)=1,\,Y\left(\omega\right)=1\right)=P\left(\{h\}\right)=1$$

**דוגמה 6.34** נטיל שני מטבעות בלתי תלויים. עם הסתברות $p$ שיצא $H$.

$$\Omega'=\{H,T\}\times\{H,T\}$$

$$X\left(\omega\right)=\begin{cases}1 & \omega_1=H\\ 0 & \omega_1=T\end{cases}\qquad Y\left(\omega\right)=\begin{cases}1 & \omega_2=H\\ 0 & \omega_2=T\end{cases}$$

נשים לב כי:

$$P\left(X=1,Y=1\right)=P\left(\{H,H\}\right)=p^2$$

נשים לב כי בשתי הדוגמאות הנ"ל, למשתנים המקריים $X,Y$ על $\Delta=\{0,1\}$ יש התפלגות זהה, כלומר:

$$p_X\left(0\right)=1-p \qquad p_Y\left(0\right)=1-p$$
$$p_X\left(1\right)=p \qquad p_Y\left(1\right)=p$$

אבל עבור $P\left(X=1,Y=1\right)$ קיבלנו תוצאות שונות.

**הגדרה 6.35** בהינתן: $\left(\Delta_1,\mathcal{G}_1\right),\left(\Delta_2,\mathcal{G}_2\right)$ מרחבים מדידים, $\sigma$־אלגברת המכפלה $\mathcal{G}_1\otimes\mathcal{G}_2$ הוא ה־$\sigma$־אלגברה הקטנה ביותר על $\Delta_1\times\Delta_2$ המכילה כל קבוצה מהצורה $A\times B$ $A\in\mathcal{G}_1,\ B\in\mathcal{G}_2$.

**דוגמה 6.36** $\sigma$־אלגברה בורל ($\mathcal{B}$) על $\mathbb{R}$ היא ה־$\sigma$־אלגברה הקטנה ביותר המכילה את הקטעים $[a,b]$.
לכן: $\mathcal{B}\times\mathcal{B}$ על $\mathbb{R}^2$ היא ה־$\sigma$־אלגברה הקטנה ביותר המכילה את כל המלבנים עם צלעות מקבילות לצירים.

אם $\Delta_1,\Delta_2$ סופיים (או בני מנייה). $\mathcal{G}_1=2^{\Delta_1},\ \mathcal{G}_2=2^{\Delta_2}$ אז ב $\mathcal{G}_1\otimes\mathcal{G}_2$ נמצאים כל היחידונים כיף $\{(i,j)\}=\{i\}\times\{j\}$ וכן $\Delta_1\times\Delta_2$ סופית לכן $\mathcal{G}_1\otimes\mathcal{G}_2$ מכילה כל תת קבוצה סופית של $\Delta_1\times\Delta_2$ (כי קבוצה סופית היא איחוד סופי של יחידונים ו $\mathcal{G}_1\otimes\mathcal{G}_2$ היא $\sigma$־אלגברה) לכן $\mathcal{G}_1\otimes\mathcal{G}_2$ מכיל את כל תתי הקבוצות של $\Delta_1\times\Delta_2$ כלומר:

$$\mathcal{G}_1\otimes\mathcal{G}_2=2^{\Delta_1\times\Delta_2}$$

**דוגמה 6.37** $\Delta_1=\Delta_2=\{0,1\}$

$$\Delta_1\times\Delta_2\supseteq\underbrace{\{(0,0),(1,1)\}}_{=(\{0\}\times\{0\})\cup(\{1\}\times\{1\})}\neq\Delta_1\times\Delta_2$$

**הגדרה 6.38** אם $X:\Omega\to\Delta_1\left(\mathcal{G}_1\right),\,Y:\Omega\to\Delta_2\left(\mathcal{G}_\in\right)$ משתנים מקריים
ההתפלגות המשותפת של $X,Y$ היא התפלגות של המשתנה המקרי $Z:\Omega\to\Delta_1\times\Delta_2$ כאשר $\omega\mapsto\left(X\left(\omega\right),Y\left(\omega\right)\right)$.
כלומר, זו מידת ההסתברות על $\left(\Delta_1\times\Delta_2,\mathcal{G}_1\otimes\mathcal{G}_2\right)$ המסומנת $P_{X,Y}$ ומאופיינת ע"י:

$$P_{X,Y}\left(A\times B\right)=P\left(X^{-1}\left(A\right)\cap Y^{-1}\left(B\right)\right)$$

**הערה 6.39** העובדה שזה מאפיין את $P_{X,Y}$ הוא משפט שלא נוכיח

אם $X,Y$ משתנים מקריים בדידים (כלומר $\Delta_1,\Delta_2$ סופיים או בני מנייה) אז $P_{X,Y}$ מאופיינת על ידי וקטור ההסתברות:

$$(i,j)\in\Delta_1\times\Delta_2,\quad p_{X,Y}(i,j)=P(X=i,Y=j)$$

בדוגמה 6.33 היה לנו:

$$\begin{aligned}p_Y(0)&=p_X(0)=1-p\\p_Y(1)&=p_X(1)=p\end{aligned}$$

$$\Delta_1\times\Delta_2=\{0,1\}\times\{0,1\}=\{(0,0),(0,1),(1,0),(1,1)\}$$

$$\begin{aligned}p_{X,Y}(0,0)&=P(X=0,Y=0)=1-p\\p_{X,Y}(1,1)&=P(X=1,Y=1)=p\end{aligned}$$

$$p_{X,Y}(0,1)=p_{X,Y}(1,0)=0$$

לעומת זו בדוגמה 6.34 נקבל כי $p_x,p_y$ (אשר מוגדרים כמו קודם) הם משתנים מקריים בלתי תלויים (נגדיר בהמשך)

$$\begin{aligned}p_{X,Y}(0,0)&=P\left(\overbrace{X=0}^{X^{-1}(0)},\overbrace{Y=0}^{Y^{-1}(0)}\right)=(1-p)^2\\p_{X,Y}(1,1)&=p^2\\p_{X,Y}(0,1)&=p_{X,Y}(1,0)=p(1-p)\end{aligned}$$

**טענה 6.40**

אם $X,Y$ ($X$ עם ערכים $\Delta_1$ ו: $Y$ עם ערכים $\Delta_2$) משתנים מקריים עם התפלגות משותפת $P_{X,Y}$ אז $P_X$ נתון ע"י:

$$P_X(A)=P_{X,Y}(A\times\Delta_2)$$

**הוכחה:**

$$P_{X,Y}(A\times\Delta_2)\underbrace{=}_{\text{הגדרה}}P\left(X^{-1}(A)\cap\underbrace{Y^{-1}(\Delta_2)}_{\Omega}\right)=P\left(X^{-1}(A)\right)\underbrace{=}_{\text{הגדרה}}P_x(A)$$

■

**מסקנה 6.41**

אם $X,Y$ בדידים אז:

$$p_X(i)=P_{X,Y}(\{1\}\times\Delta_2)=P_{X,Y}\left(\bigcup_{j\in\Delta_2}\{(i,j)\}\right)=\sum_{j\in\Delta_2}p_{X,Y}(i,j)$$

$p_X(i) = \sum_{j\in\Delta_2} p_{X,Y}(i,j)$ באופן סימטרי: $p_Y(j) = \sum_{i\in\Delta_i} p_{X,Y}(i,j)$

11/12/20

**דוגמה 6.42** $\Omega = \{0,1\}^3$ עם התפלגות אחידה ( $\Longleftrightarrow$ הטלות מטבע הוגנות בלתי תלוייות)
נסמן

$$X\underbrace{(i,j,k)}_{\in\Omega} = i+j+k = \#\{\text{1-ים}\}$$

$$Y(i,j,k) = ij = \begin{cases} 1 & \text{יצא 1 בשתי ההטלות הראשונות} \\ 0 & \text{אחרת} \end{cases}$$

$$X \sim B\left(3, \frac{1}{2}\right)$$

$$p_X(0) = p_X(3) = \frac{1}{8}$$

$$p_X(1) = p_x(2) = \frac{3}{8} = \binom{3}{1}\frac{1}{2}\left(\frac{1}{2}\right)^2$$

$$p_Y(1) = P(\{(1,1,0),(1,1,1)\}) = \frac{2}{8} = \frac{1}{4}$$

$$p_Y(0) = 1 - \frac{1}{4} = \frac{3}{4}$$

$(X,Y)$ מקבל ערכים ב:

$$\Gamma = \{0,1,2,3\} \times \{0,1\}$$

$p_{X,Y}$ הוא וקטור הסתברות על $\Gamma$

| $Y\diagdown X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{2}{8}$ | 0 |
| 1 | 0 | 0 | $\frac{1}{8}$ | $\frac{1}{8}$ |

**הערה 6.43** העמודה $i$ מסתכמת ל $p_X(i)$. זו תופעה כללית.
אם $X,Y$ משתנים מקריים בדידים על אותו מרחב אז:

$$p_X(i) = P(\omega : X(\omega) = i) = P\left(\bigcup_j \{\omega : X(\omega) = i, Y(\omega) = j\}\right) = \sum_j P(X=i, Y=j) = \sum_j p_{X,Y}(i,j)$$

כנל:

$$P_Y(j) = \sum_i p_{X,Y}(i,j)$$

כלומר, ההתפלגות המשותפת קובעת את ההתפלגויות של כל משתנה (ראינו שלא להפך).

**הערה 6.44** בדוגמה הקודמת, נגדיר:

$$\tilde{X}=3-X=\#\{\text{אפסים}\}$$

שוב נקבל: $\tilde{X}=B\left(3,\frac{1}{2}\right)$. אבל הפעם: ההתפלגות: $\tilde{X},Y$ היא:

| $Y\diagdown X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 0 | $\frac{2}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |
| 1 | $\frac{1}{8}$ | $\frac{1}{8}$ | 0 | 0 |

כלומר: $X,\tilde{X}$ שווי התפלגות. אבל הזוגות $(X,Y)$, $\left(\tilde{X},Y\right)$ לא שווי התפלגות.

כמו כן $X\neq\tilde{X}$:

$$\begin{aligned}\{X=0\} &= \{0,0,0\}\\ \left\{\tilde{X}=0\right\} &= \{1,1,1\}\end{aligned}$$

**הגדרה 6.45** אם $X_i:\Omega\to\Delta_i$ כאשר $i=1,\ldots,n$ משתנים מקריים. ההתפלגות המשותפת: $P_{X,\ldots,X_n}$היא התפלגות של המשתנים המקריים: $Z=(X_1,\ldots,X_n)$ $\Omega\to\Delta_1\times\ldots\times\Delta_n$.

**הערה 6.46** צריך עקרונית להגדיר $\sigma$־אלגברה על $\Delta_n\times\ldots\times\Delta_1$ לא נעשה זאת. במקרה ש $X_i$ בדידים לוקחים על $\Delta_n\times\ldots\times\Delta_1$ את $\sigma$־אלגברה כל הקבוצות ואז $P_{X_1,\ldots,X_n}$נקבעת ע"י הסתברות של יחידונים.

$$p_{X_1,\ldots,X_n}(i_1,\ldots,i_n),\quad (i_1,\ldots,i_n)\in\Delta_1\times\ldots\times\Delta_n$$

## 6.6 אי תלות

**הגדרה 6.47** משתנים מקריים $X:\Omega\to\Delta_1$, $Y:\Omega\to\Delta_2$ נקראים בלתי תלויים אם לכל $A\subseteq\Delta_1$, $B\subseteq\Delta_2$ מאורעות, המאורעות: $X^{-1}(A),\ Y^{-1}(B)$ בלתי תלויים.
באופן שקול:

$$P(x\in A,Y\in B)=P_{X,Y}(A\times B)=P_X(A)P_Y(B)=P(X\in A)P(Y\in B)$$

בקצרה:

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

**למה 6.48**

אם $X,Y$ בדידים אז $X,Y$ בלתי תלויים $\Longleftrightarrow$ לכל $(i,j)$מקיים:

$$p_{X,Y}(i,j)=p_X(i)p_Y(j)$$

**הוכחה:** $\Leftarrow$ אם $X,Y$ בלתי תלויים אזי הנוסחה נובעת מההגדרה: $P(X\in A,Y\in B)=P(X\in A)P(Y\in B)$ עבור $A=\{i\},\ B=\{j\}$.
$\Rightarrow$נניח ש $p_{x,y}(i,j)=p_x(i)p_y(j)$ מתקיים לכל $(i,j)$. יהיו $A\subseteq\Delta_1$ ו: $B\subseteq\Delta_2$. נוכיח כי $P(x\in A,y\in B)=P(x\in A)P(y\in B)$. נשים לב כי:

$$P(\omega:X(\omega)\in A,Y(\omega)\in B)=P\left(\bigcup_{(i,j)\in A\times B}\{\omega:X(\omega)=i,Y(\omega)=j\}\right)$$

אבל זהו איחוד זר ובן מנייה לכל היותר כי $X,Y$ בדידים.

$$= \sum_{(i,j)\in A\times B} P\left(X=i,Y=j\right) = \sum_{i\in A}\left(\sum_{j\in B} p_X\left(i\right)p_Y\left(j\right)\right)$$

($P\left(X=i,Y=j\right)$ מסומן כ־$p_{X,Y}\left(i,j\right)$)

נוציא מהסכום הפנימי את $p_X\left(i\right)$ שאיננו תלוי ב$j$:

$$= \sum_{i\in A} p_X\left(i\right)\underbrace{\sum_{j\in B} p_Y\left(j\right)}_{P(y\in B)} = P\left(Y\in B\right)\underbrace{\sum_{i\in A} p_X\left(i\right)}_{P(X\in A)} = P\left(Y\in B\right)P\left(X\in A\right)$$

כנדרש. ■

**דוגמה 6.49** מטילים שתי קוביות הוגנות בלתי תלויות:

$$\Omega = \{1,\ldots,6\}^2$$

תחת הסתברות אחידה.

$$X_i = \text{תוצאות ההטלה ה־}i$$

$$Y = \begin{cases} 0 & \text{זוגי } X_2 \\ 1 & \text{אי זוגי } X_2 \end{cases}$$

נבדוק אילו זוגות משתנים מקריים הם בלתי תלויים.

$X_1,X_2$:

$$p_{X_1}\left(i\right) = \frac{1}{6}$$

$$p_{X_2}\left(i\right) = \frac{1}{6}$$

$$p_{X_1,X_2}\left(i,j\right) = \frac{1}{36} = \frac{1}{6}\cdot\frac{1}{6} = p_{X_1}\left(i\right)p_{X_2}\left(j\right)$$

אז ם בלתי תלויים.

$X_1,Y$:

$$p_Y\left(0\right) = \frac{3}{6} = \frac{1}{2} = p_Y\left(1\right)$$

$$p_{X_1,Y}\left(i,j\right) = \frac{1}{36}\left|\{(i,k) \mid j \text{ בעל אותו זוגיות גמו } k\}\right| = \frac{1}{36}\cdot 3 = \frac{1}{6}\cdot\frac{1}{2} = p_X\left(i\right)\cdot p_Y\left(j\right)$$

$X_2,Y$:

$$P\left(X_2=5,Y=0\right) = 0 \neq p_{X_2}\left(5\right)p_Y\left(0\right) = \frac{1}{12}$$

לכן הם לא בלתי תלויים.

**דוגמה 6.50** יהיו $X,Y$ משתנים גיאומטריים בלתי תלויים עם פרמטרים $\alpha,\beta$ חשבו את:

$$P\left(X\subseteq Y\right) = ?$$

נתון:

$$\begin{aligned} p_X(i) &= \alpha(1-\alpha)^{i-1} \\ p_y(j) &= \beta(1-\beta)^{j-1} \end{aligned}$$

כיוון ש $X,Y$ בלתי תלויים:

$$p_{X,Y}(i,j) = \alpha\beta(1-\alpha)^{i-1}(1-\beta)^{j-1} = p_X(i)\,p_Y(j)$$

כעת:

$$P(\{\omega|\ X(\omega)\le Y(\omega)\}) = P\left(\bigcup_{1\le i\le j<\infty}\{\omega|\ X(\omega)=i,\ Y(\omega)=j\}\right)$$

זהו איחוד זר בן מנייה

$$= \sum_{1\le i\le j<\infty} P(X=i,Y=j) = \sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty}\alpha\beta(1-\alpha)^{i-1}(1-\beta)^{j-1}\right) = \frac{\alpha}{\alpha+\beta-\alpha\beta}$$

**הערה 6.51**

1. $X,Y$ בלתי תלויים אומר כי כל אינפורמציה על אחד האחד בלתי תלויה מאינפורמציה על השני. (אינפורמציה על $X$ למשל: $X\in A$)
2. השיוויון $P(X\in A,Y\in B)=P(X\in A)P(Y\in B)$ חייב להתקיים לכל $A,B$.
3. נסמן $\sigma(X)=\left\{X^{-1}A:\ X \text{ של בטווח מאורע } A\right\}$. מסתבר שזו $\sigma$־אלגברה $\mathcal{F}\supseteq$ כי $X$ מדיד. אי תלות של $X,Y \iff$ בלתי תלויים $A,B$ $\forall A\in\sigma(X),\ B\in\sigma(Y)$.

אומרים אז ש $\sigma(X),\sigma(Y)$ בלתי תלויים ־ הגדרה של אי תלות של $\sigma$־אלגבראות. (לא נראה שקילות זו)

### 6.6.1 אי תלות של סדרות של משתנים מקריים

אם $X_i:\Omega\to\Delta_i$ משתנה מקרי, $i=1,\dots,n$. אומרים שהם בלתי תלויים (במשותף) אם לכל $n$־יה של מאורעות $A_i\subseteq\Delta_i$:

$$P\underbrace{(X_i\in A_i,\dots,X_n\in A_n)}_{\substack{\bigcap_{i=1}^{n} X_i^{-1}(A_i) \\ || \\ \{\omega\in\Omega|\ X_i(\omega)\in A_1,\dots,X_n(\omega)\in A_n\}}} = \prod_{i=1}^{n}P(X_i\in A_i)$$

אם $X_1,\dots,X_n$ בדידים זה שקול:

$$p_{x_1,\dots,x_n}(i_1,\dots,i_n) = \prod_{k=1}^{n} p_{x_k}(i_k)$$

**הערה 6.52** נניח $X_1,\ldots,X_n$ בלתי תלויים. $A_1,A_2$מאורעות בטווח ש $X_1,X_2$ בהתאמה

$$P\left(X_1\in A_1,X_2\in A_2\right)=?$$

נגדיר $A_i=\Delta_i$ עבור $n$. $i\geq 3$נשים לב כי:

$$X_i^{-1}\left(\Delta_i\right)=\Omega$$

לכן:

$$\{X_i\in A_i|\ i=1,\ldots,n \text{ לכל}\}=\{X_i\in A_1\}\cap\{X_2\in A_2\}\cap\bigcap_{j=3}^{n}X_j^{-1}A_j=\{X_1\in A_1,\ X_2\in A_2\}$$

ולכן:

$$P\left(X_1\in A_1,X_2\in A_2\right)=P\left(\{X_i\in A_i,\ \forall 1\leq i\leq n\}\right)\underbrace{=}_{\text{מאי תלות}}\prod_{j=1}^{n}\underbrace{P\left(X_j\in A_j\right)}_{j=3\ldots n \text{ עבור כל}=1}=P\left(X_1\in A_1\right)P\left(X_2\in A_2\right)$$

הערה דומה נכונה לכל תת משפחה של המשתנים.

**הערה 6.53** אם $X_1,\ldots,X_n$בלתי תלויים ו $A_1,\ldots,A_n$מאורעות בהתאמה. אז $X_1^{-1}A_1,\ldots,X_n^{-1}A_n$ מאורעות בלתי תלויים. (נובע בהערה הקודמת)

**הגדרה 6.54** סדרה אינסופית של משתנים מקריים נקראת בלתי תלויה אם כל $n$־יה מתוכה בלתי תלויה.

## 6.7 פונקציות של משתנים מקריים

**למה 6.55**

נניח ש $X:\Omega\to\Delta$ משתנה מקרי בדיד. ו: $f:\Delta\to\Gamma$ פונקציה אז:

$$f\circ x=Z:\Omega\to\Gamma$$

היא משתנה מקרי ומתקיים:

$$P\left(Z\in A\right)=\sum_{i\in f^{-1}(A)}P\left(X=i\right)$$

**דוגמה 6.56** מקרה פרטי:

$$P\left(Z=j\right)=\sum_{i\in f^{-1}(j)}P\left(X=i\right)$$

$$p_z\left(j\right)=\sum_{i\in f^{-1}(1)}p_x\left(i\right)$$

**הוכחה:** $Z$ פונקציה מ$\Omega$ ל$\Gamma$.

$$\{\omega|\ Z\left(\omega\right)\in A\}=\bigcup_{i\in f^{-1}(A)}\{\omega|\ X\left(\omega\right)=i\}$$

(איחוד זר).

זהו איחוד לכל היותר בן מנייה. כי $X$ בדיד ($\Delta$ בת מנייה)

כל קבוצה $\{\omega|\ X(\omega)=i\}$ הוא מאורע כי $X$ משתנה מקרי $\Leftarrow$האיחוד הוא מאורע.

זה מראה ש $Z$ משתנה מקרי. הנוסחה להסתברות נובעת מהפעלת $P$ על האיחוד $\{\omega|\ X(\omega)=i\}$ $\bigcup_{i\in f^{-1}(A)}$.■

**מסקנה 6.57**

אם $X_1,X_2$משתנים מקריים בדידים $\Delta_1\to\Omega$ ואם $f:\Delta_1\times\Delta_2\to\Gamma$. אז $Z=f(X_1,X_2)$הוא משתנה מקרי ו:

$$p_Z(k)=\sum_{(i,j)\in f^{-1}(k)} p_{X_1,X_2}(i,j)$$

ובאופן דומה ל $Z=f(X_1,\ldots,X_n)$

**הוכחה:** (לשניים)

נגדיר $W=(X_1,X_2)$ אז מהלמה:

$$p_Z(k)=\sum_{(i,j)\in f^{-1}(k)} p_w(i,j)$$

אבל $p_W=p_{X_1,X_2}$לפי הגדרה. ■

**דוגמה 6.58 עיקרית:** נניח:

$$X,Y:\Omega\to\Delta\subseteq\mathbb{R}$$

אז $Z=X+Y$ הוא מקרה של המצב למעלה.

$$f(x,y):\Delta\times\Delta\to\mathbb{R}$$

$$(x,y)\mapsto x+y$$

אז $Z=f(X,Y)$ לפי הנוסחה:

$$p_Z(k)=\sum_{\substack{\cancel{(i,j)\in f^{-1}(k)}\\ i+j=k\\ (i,j)\in\Delta\times\Delta}} p_{x,y}(i,j)$$

נסכים ש $p_{X,Y}(i,j)=0$ אם $i$ או $j$ לא ב$\Delta$.

$$p_z(k)=\sum_{i\in\Delta} p_{X,Y}(i,k-i)$$

נניח עוד ש: $X,Y$ בלתי תלויים אזי:

$$p_{X+Y}(k)=\sum_i p_X(i)\,p_Y(k-1)$$

ביטוי כזה נקרא קונבולוציה (Convolution) של הוקטורים $p_Y,p_X$.

**דוגמה 6.59** נניח $X\sim\mathrm{Pois}(\alpha)$, $Y\sim\mathrm{Pois}(\beta)$, $X,Y$ בלתי תלויים. נחשב את $p_{X+Y}$:

$$p_{X+Y}(k)=\underbrace{\sum_{i=0}^{\infty}}_{X\text{ ערכים אפשריים של}} p_X(i)\,p_Y\underbrace{(k-i)}_{k-1\geq 0\text{ מתרחש בהסתברות חיובית רק אם }Y=k-i}$$

$$\sum_{i=0}^{k}\overbrace{\binom{k}{i}\alpha^{i}\beta^{k-i}}^{(\alpha+\beta)^{k}}\frac{e^{-(\alpha+\beta)}}{k!}\underbrace{=}_{\text{נוסחת הבינום}}\sum_{i=0}^{k}\frac{\alpha^{i}}{i!}e^{-\alpha}\cdot\frac{\beta^{k-i}}{(k-i)!}e^{-\beta}\frac{k!}{k!}\underbrace{=}_{\text{משתנים פואסוניים}}\sum_{i=0}^{k}p_X(o)\,p_Y(k-i)=$$

כלומר:

$$Z=X+Y\sim\text{Pois}(\alpha+\beta)$$

**טענה 6.60**

יהיו $X,Y$ משתנים מקריים בלתי תלויים בדידים ו $Z=f(X)$ אז: $Z,Y$ בלתי תלויים

**הוכחה:**

$$p_{Z,Y}(i,j)=\sum_{u\in f^{-1}(i)}p_{X,Y}(u,j)$$

**תרגיל:** להצדיק שלבי ביניים בשוויון הנ"ל.

$$=\sum_{u\in f^{-1}(i)}p_X(u)\,p_Y(j)=p_Y(j)\underbrace{\sum_{u\in f^{-1}(i)}p_X(u)}_{p_z(i)}$$

קיבלנו:

$$p_{Z,Y}(i,j)=p_Z(i)\,p_Y(j)$$

כנדרש. ■

**מסקנה 6.61**

הכללה:

אם $X_1,\ldots,X_n$ ו: $Y_1,\ldots,Y_m$ בלתי תלויים ואם:

$$Z=f(X_1,\ldots,X_m)$$

אז:

$$Z,Y_1,\ldots Y_n$$

בלתי תלויים

**הוכחה:** (ראינו עבור $m=n=1$)... ■

**דוגמה 6.62** נניח $X_1,X_2,X_3$ פואסון: $X_i\sim\text{Pois}(\alpha_i)$ בלתי תלויים אז:

$$X_1+X_2+X_3\sim\text{Pois}(\alpha_1+\alpha_2+\alpha_3)$$

**הוכחה:** ראינו כי $X_1+X_2\sim\text{Pois}(\alpha_1+\alpha_2)$. לפי הכללה $X_1+X_2,X_3$ בלתי תלויים. שוב מפעילים את הדוגמה ונקבל:

$$X_1+X_2+X_3\sim\text{Pois}(\alpha_1+\alpha_2+\alpha_3)$$

כנדרש. ■

12/12/20

**תזכורת 6.63** $X_i:\Omega\to\Delta_i$ כאשר $i=1,\dots,n$ נקראים בלתי תלויים אם לכל למאורעות: $A_i\subseteq\Delta_i$:

$$P\left(X_i\in A_i,\dots,X_n\in A_n\right)=\prod_{i=1}^{k}P\left(X_i\in A_i\right)$$

**הערה 6.64** זה גורר שלכל $A_1,\dots,A_n$כאלה:

$$\left\{X_1\in A_i\right\},\dots,\left\{X_n\in A_n\right\}\subseteq\Omega$$

בלתי תלויים.

**הוכחה:** כפי שראינו, לכל $I\subseteq\{1,\dots,n\}$ אי תלות של $X_1,\dots,X_n$ גורר:

$$P\left(X_i\in A_i\middle|\,\forall i\in I\right)=\prod_{i\in I}P\left(X_i\in A_i\right)$$

וזו ההגדרה של אי־תלות של מאורעות למעשה. ■

**דוגמה 6.65** יהי $0<p<1$ ו: $0<c$.
יהי $X_n$ סדרה בלתי תלויה של משתנים מקריים: $n=1,2,3,\dots$ כך ש: $X_n\sim\mathrm{Geom}\left(p\right)$. מה ההסתברות ש: $X_n>c\cdot\log n$ לאינסוף $n$־ים.
**פתרון:**
נסמן $A_n=\left\{X_n>c\log n\right\}$

$$P\left(A_n\right)=P\left(X_n>c\log n\right)=\sum_{j>c\log n}P\left(X_n=j\right)=\sum_{j>c\log n}p\left(1-p\right)^{j-1}=\underbrace{D}_{\text{קבוע שתלוי רק ב } p}\left(1-p\right)^{\lfloor c\log n\rfloor+1}$$

$$\Rightarrow D'\left(1-p\right)^{c\log n}\leq P\left(A_n\right)\leq D''\left(1-p\right)^{c\log n}$$

$$\left(1-p\right)^{c\log n}=n^{c\log n(1-p)}$$

לפי בורל קנטלי I, אם $\sum P\left(A_n\right)<\infty$ אזי $A_n$קורה מספר סופי של פעמים בהסתברות 1.
לפי בורל קנלטי II, אם בנוסף $A_n$בלתי תלויים (וזה המקרה כאן) $\sum P\left(A_n\right)=\infty$ אזי: $A_n$ קורה אינסוף פעמים בהסתברות 1.
כעת:

$$\sum_{n=1}^{\infty}P\left(A_n\right)<\infty\iff\sum_{n=1}^{\infty}n^{c\log(1-p)}<\infty\iff c\log\left(1-p\right)<-1$$

**הערה 6.66 על סוף שעור קודם:** ראינו שאם $X,Y$ בלתי תלויים ו: $Z=f\left(X\right)$ אז: $Z,Y$ בלתי תלויים. נרצה להראות גם את הטענה הבאה

**טענה 6.67**

אם $X_1,\dots,X_m,Y_1,\dots,Y_n$ בלתי תלויים ודיסקרטיים ו: $Z=f\left(X_1,\dots,X_m\right)$אז: $Z,Y_1,\dots,Y_m$ בלתי תלויים

**הוכחה:** (רק סקירה)
נוכיח את הלמה:

**למה 6.68**

נסמן $\tilde{X}=(X_1,\ldots,X_m)$, $\tilde{Y}=(Y_1,\ldots,Y_n)$ אז $\tilde{X},\tilde{Y}$ בלתי תלויים.

**הוכחה:** נבחר $\tilde{a}=(a_1,\ldots,a_m)$, $\tilde{b}=(b_1,\ldots,b_n)$

$$P\left(\tilde{X}=\tilde{a},\,\tilde{Y}=\tilde{b}\right)=\underbrace{\prod_{i=1}^{m}P\left(X_i=A_i\right)}_{P\left(\tilde{X}=\tilde{a}\right)}\cdot\underbrace{\prod_{j=1}^{n}P\left(Y_j=b_j\right)}_{P\left(\tilde{Y}=\tilde{b}\right)}$$

(מכך ש $X_1,\ldots,X_m,Y_1,\ldots,Y_n$ בלתי תלויים).
מאותם שיקולים ומהערה הראשונה.
ולכן $\tilde{X},\tilde{Y}$ בלתי תלויים כנדרש. ■

נשים לב: $Z=f\left(\tilde{X}\right)$ לפי מה קראינו פעם שעברה, $Z,\tilde{Y}$ בלתי תלויים (כי $\tilde{X},\tilde{Y}$בלתי תלויים)
כעת אם $\tilde{b}=(b_1,\ldots,b_n)$אזי:

$$P\left(Z=K,Y_1=b_1,\ldots,Y_n=b_n\right)=P\left(Z=k,\,\tilde{Y}=\tilde{b}\right)=P\left(Z=k\right)P\left(\tilde{Y}=\tilde{b}\right)=P\left(Z=k\right)P\left(Y_1=b_1,\ldots,Y_n=b_n\right)$$
$$P\left(Z=k\right)P\left(Y_1=b_1\right)\ldots P\left(Y_n=b_n\right)$$

כנדרש. ■

## 6.8 התפלגויות מותנות

כבר הגדרונו:

$$A,B\subseteq\Omega\qquad P\left(A|B\right)=\frac{P\left(A\cap B\right)}{P\left(B\right)}$$

זה נותן משמעות ל:

$$A\subseteq X \text{ טווח של },B\subseteq Y\text{ טווח של}\qquad P\left(X\in A|Y\in B\right)=\frac{P\left(\{X\in A\}\cap\{Y\in B\}\right)}{P\left(Y\in B\right)}$$

כאשר $X,Y$ בדידים נסמן:

$$p_{X|Y}\left(i|j\right)=P\left(X=i|Y=j\right)=\frac{P\left(X=i,Y=j\right)}{P\left(Y=j\right)}=\frac{p_{X,Y}\left(i,j\right)}{p_Y\left(j\right)}$$

כל זה בהנחה שבמכנה לא מופיע 0.

**דוגמה 6.69** יהיו $X,Y$, $p_{X,Y}$ נתונה ע"י:

| $Y\diagdown X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{2}{8}$ | 0 |
| 1 | 0 | 0 | $\frac{1}{8}$ | $\frac{1}{8}$ |

מה ההסתברות של:

$$p_{X|Y}\left(2|1\right)=\frac{p_{X,Y}\left(2,1\right)}{p_Y\left(1\right)}=\frac{1/8}{2/8}=\frac{1}{2}$$

**הערה 6.70**

1. $p_{x,y}(i,j) = p_{X|Y}(i|j)\, p_Y(j)$נכון תמיד.
2. נסכום על $j$:

$$p_X(i) = \sum_j p_{X,Y}(i,j) = \sum_j p_{X|Y}(i|J)\, p_Y(j)$$

וזאת נוסחת ההסתברות השלמה תחת הסימון החדש.

3. בהינתן מאורע: $B \subseteq \Omega$ ראינו ש: $Q(F) = P(F|E)$ זו פונקציית הסתברות על $\mathcal{F}$. אפשר להתבונן במשתנה המקרי $X : \Omega \to \Delta$ ביחס להסתברות $Q$

$$\underbrace{Q(X \in A)}_{\text{ההתפלגות של } X \text{ביחס ל } Q} = P(X \in A|E) = P(X \in A|Y \in B)$$

**דוגמה 6.71** יתושה מטילה $X$ ביצים כך ש $X \sim \text{Pois}(\lambda)$ כל ביצה בוקעת באופן בלתי תלוי זה מזה בהסתברות $p$.
יהי $Y$ מספר הצאצאים (= ביצים שבקעו)
מה ההתפלגות של $Y$?
**<u>פתרון:</u>**
יודעים ש:

$$P(Y = K|X = n) = \binom{n}{k} p^k (1-p)^{n-k}$$

נכון אם $n \leq k$ כי אם $k > n \Leftarrow p = 0$.

$$P(Y = k) = \sum_{n=0}^{\infty} P(Y = k|X = n)\, P(X = n) = \sum_{n=k}^{\infty} \left[\binom{n}{k} p^k (1-p)^{n-k}\right] \left[\frac{\lambda^n}{n!} e^{-\lambda}\right] =$$

$$\frac{p^k \lambda^k}{k!} e^{-\lambda} \sum_{n=k}^{\infty} \frac{\lambda^{n-k}}{(n-k)!} (1-p)^{n-k} = \frac{p^k \lambda^k}{k!} e^{-\lambda} e^{\lambda(1-p)} = \frac{(p\lambda)^k}{k!} e^{-\lambda p}$$

18/12/20

# 7 תוחלת

**הגדרה 7.1 תוחל:** בהינתן $(\Omega, 2^{\Omega}, P)$ מרחב הסתברות, $\Omega$ סופית או בת מנייה. יהי $X : \Omega \to \underbrace{\Delta}_{\text{בת מנייה}} \subseteq \mathbb{R}$ משתנה מקרי. התוחלת של $X$ היא:

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} P(\{\omega\})\, X(\omega)$$

בתנאי שהסכום מתכנס בהחלט ($\sum_{\omega \in \Omega} P(\{\omega\}) |X(\omega)| < \infty$) אם זה מתקיים, אומרים שהתוחלת קיימת.

מילים נרדפות: ממוצע, Expectation, mean ולפעמים מסמנים: $\mathbb{E}(X) = \langle X \rangle$.
נשים לב כי $P(\{\omega\})$ הם "משקולות" אי שליליים המסתכמים ל1. לכן הסכום = ממוצע שקולל (עם משקולות) של הערכים $X(\omega)$, $\omega \in \Omega$.

**דוגמה 7.2** נתבונן בהטלת קובייה, $\Omega=\{1,\ldots,6\}$ . עם הסתברות אחידה. נגדיר:

$$X(i)=\begin{cases} i & i=1,2,3 \\ 1 & i=4,5,6 \end{cases}$$

נחשב את התוחלת של $X$:

$$\mathbb{E}(X)=\frac{1}{6}\cdot\underbrace{1}_{X(1)}+\frac{1}{6}\cdot\underbrace{2}_{X(2)}+\frac{1}{6}\cdot\underbrace{3}_{X(3)}+\frac{1}{6}\cdot\underbrace{1}_{X(4)}+\frac{1}{6}\cdot\underbrace{1}_{X(5)}+\frac{1}{6}\cdot\underbrace{1}_{X(6)}=\frac{9}{6}$$

**הערה 7.3** נדמיין שמבצעים ניסוי אופן בלתי תלוי $N$ פעמים. השכיחות של מצב $\omega$ הוא בערך $N\cdot P(\{\omega\})$. אז הערך $X(\omega)$ מופיע $N\cdot P(\{\omega\})$ פעמים. והערך הממוצע (הפשוט) יהיה בערך $\mathbb{E}(X)$.
נוכיח משפט בהמשך ברוח זו.

**טענה 7.4**

אם $X$ כמו בהגדרה, אז:

$$\mathbb{E}(X)=\sum_{a\in\mathbb{R}}a\cdot p_X(a)$$

**הערה 7.5** כאשר הסכימה היא על טווח עם מוסכמה ש $p_x(a)=0$ אם $a$ לא ערך של $X$ ואז מרשים סכומים על $\mathbb{R}$.

**הוכחה:**

$$\sum_a a\cdot p_X(s)=\sum_a a\cdot P(\omega:\ X(\omega)=a)=\sum_a a\cdot\underbrace{\sum_{\omega\in X^{-1}(a)}P(\{\omega\})}_{P(\omega:\ X(\omega)=a)}=\sum_a\sum_{\omega\in X^{-1}(a)}\left(\overset{X(\omega)}{\cancel{a}}\cdot P(\{\omega\})\right)=$$

$$=\sum_{\omega\in\Omega}X(\omega)P(\{\omega\})=\mathbb{E}(x)$$

הסבר למעבר אחרות:

$$\bigcup_{a\in\mathbb{R}}\left(\bigcup_{\omega\in X^{-1}(a)\subseteq\Omega}\{\omega\}\right)=\Omega$$

$\subseteq$ ברור. נרצה לשים לב כי $\supseteq$ נובע מכך שלכל $\omega\in\Omega$ מתקיים: $\omega\in X^{-1}(a)$ עבור $a=X(\omega)$ . כנדרש. ■

**הגדרה 7.6** הגדרה שניה, קצת יותר כללית.
אם $X$ בדיד עם ערכים ב$\mathbb{R}$ מגדירים:

$$\mathbb{E}(x)=\sum_{a\in\mathbb{R}}a\cdot p_x(a)$$

בתנאי שהטור מתכנס בהחלט.

**דוגמה 7.7** מטילים $\infty$ מטבעות. יהי $N$ הזמן הראשון שמקבלים ראש. מרחב המדגם כאן אינו בן מנייה, $\Omega=\{H,T\}^{\mathbb{N}}$. אבל $N$ בדיד (כי מקבל ערכים ב $\mathbb{N}$) לכן $\mathbb{E}(N)$ מוגדרת (צריך לבדוק שהטור סופי)
(נשים לב כי ההגדרה הראשונה אינה מאפשרת לדבר על $\mathbb{E}(N)$)

**דוגמה 7.8** יהי $X$ כמו בדוגמה של הטלת הקובייה, כלומר:

$$\Omega = \{1,\ldots,6\}$$
$$X(i) = \begin{cases} i & i=1,2,3 \\ 1 & i=,4,5,6 \end{cases}$$

נשים לב כי:

$$p_X(k) = \begin{cases} \frac{4}{6} & k=1 \\ \frac{1}{6} & k=2 \\ \frac{1}{6} & k=3 \end{cases}$$

כאשר כמובן $k \in \{1,2,3\}$. כמו כן $\underbrace{\{1,4,5,6\}}_{\{X=1\}} \subseteq \Omega$ לכן $p_X(1) = \frac{4}{6}$

לכן לפי הטענה:

$$\mathbb{E}(X) = 1\cdot p_X(1) + 2\cdot p_X(2) + 3\cdot p_X(3) = \frac{4}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} = \frac{9}{6}$$

## 7.1 תוחלת של התפלגויות קלאסיות

### 7.1.1 משתנה ברנולי

אם $X$ משתנה ברנולי $X \sim \mathrm{Ber}(p)$ כלומר:

$$p_x(k) = \begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$$

אזי:

$$\mathbb{E}(X) = 1\cdot p + 0(1-p) = p$$

כלומר התוחלת היא הפרמטר.

**הערה 7.9** אם $A$ מאורע במרחב הסתברות אז:

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

הוא משתנה ברנולי עם פרמטר:

$$P(\mathbb{1}_A = 1) = P(\omega : \omega \in A) = P(A)$$

לכן:

$$\mathbb{E}(\mathbb{1}_A) = P(A)$$

### 7.1.2 משתנה בינומי

אם $X$ בינומי, כלומר: $X \sim \mathrm{Bin}(n,p)$ אז:

$$\mathbb{E}(X) = \sum_{k=0}^{n} k\binom{n}{k} p^k (1-p)^{n-k} = \ldots = np$$

**הערה 7.10** $k$ זה הערכים ש $X$ מקבל. אנו יודעים שהם טבעיים לכן אין צורך לסכום על כל $\mathbb{R}$.

#### 7.1.3 משתנה פואסוני

אם $X$ משתנה פואסוני, כלומר: $X \sim \text{Pois}(\alpha)$ אז:

$$\mathbb{E}(X) = \sum_{k=0}^{\infty} k \underbrace{\frac{\alpha^k}{k!} e^{-\alpha}}_{p_X(k)} = e^{-\alpha} \sum_{k=1}^{\infty} \frac{\alpha^{k-1}\alpha}{(k-1)!} \underbrace{=}_{\substack{m=k-1 \\ \text{החלפת משתנה סכימה}}} \alpha e^{-\alpha} \underbrace{\sum_{m=0}^{\infty} \frac{\alpha^m}{m!}}_{e^{\alpha}} = \alpha$$

**הערה 7.11** $k$ זה הערכים ש $X$ מקבל. אנו יודעים שהם טבעיים לכן אין צורך לסכום על כל $\mathbb{R}$.

כיוון שהסכום חיובי, התכנסות גוררת התכנסות בהחלט, לכן התוחלת קיימת.

**הערה 7.12** ראינו שאם $\alpha$ נתון: $X_n \sim \text{Bin}\left(n, \frac{\alpha}{n}\right)$ אז לכל $0 \le k$

$$p_X(k) \underset{n\to\infty}{\longrightarrow} p_X(k) = \frac{\alpha^k}{k!} e^{-\alpha}$$

כאשר $X \sim \text{Pois}(\alpha)$ אנו יודעים:

$$\mathbb{E}(X_n) = n\frac{\alpha}{n} = \alpha$$

לכל $n$.
אפשר היה להניח $\mathbb{E}(X) = \alpha$ מהגבול הזה (אבל יש קשיים טכניים)

#### 7.1.4 משתנה גיאומטרי

$X \sim \text{Geom}(p)$ כאשר $0 < p < 1$

$$\mathbb{E}(X) = \underbrace{\sum_{k=1}^{\infty}}_{X=\{1,2,\dots\}} k \cdot \underbrace{p(1-p)^{k-1}}_{p_X(k)} = p \sum_{k=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}t}\left(t^k\right)\bigg|_{t=(1-p)}$$

כיוון ש: $0 < 1-p < 1$ ו $\sum t^k$ מתכנס ב $(-1,1)$. מאינפי:

$$= p\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=1}^{\infty} t^k\right)\bigg|_{t=1-p} = p\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t}{1-t}\right)\bigg|_{t=1-p} = p\frac{1}{(1-t^2)}\bigg|_{t=1-p} = \frac{p}{p^2} = \frac{1}{p}$$

## 7.2 תוחלת של פונקציות על משתנים מקריים

**תזכורת 7.13** אם $X : \Omega \to \Delta$ (דיסקרטי) ו $f : \Delta \to \Gamma$ (נגדיר: $Y = f(x)$) אז :

$$P\left(\underbrace{f(x)}_{Y} = b\right) = \sum_{a \in f^{-1}(b)} p_x(a)$$

**משפט 7.14**

אם $X$ בדיד עם ערכים ב $\Delta$ (קבוצה כללית), $f : \Delta \to \Gamma \subseteq \mathbb{R}$ יהי $Y = f(X)$ משתנה בדיד עם ערכים ממשיים.
אז:

$$\mathbb{E}(f(X)) = \mathbb{E}(Y) = \sum_{a\in\Delta} f(a) p_X(a)$$

**הוכחה:**

$$\mathbb{E}\left(Y\right)=\sum_{b\in\Gamma}p_{Y}\left(b\right)b=\sum_{b\in\Gamma}b\cdot\left(\sum_{a\in f^{-1}(b)}p_{X}\left(a\right)\right)=\sum_{b\in\Gamma}\sum_{a\in f^{-1}(b)}\overset{f(a)}{\cancel{b}}p_{X}\left(a\right)=\sum_{a\in\Delta}f\left(a\right)p_{X}\left(a\right)$$

(מעבר אחרון כמו בטענה קודמת) כנדרש. ■

**דוגמה 7.15** שוב הדוגמה של הטלת הקוביה:

$$\begin{aligned}\Omega &= \{1,\dots,6\}\\ X\left(i\right) &= \begin{cases}1 & i=1,4,5,6\\ 2 & i=2\\ 3 & i=3\end{cases}\end{aligned}$$

נגדיר $f\left(t\right)=t^{2}$ כלומר:

$$Y=X^{2}=f\left(X\right)$$

נקבל:

$$\mathbb{E}\left(Y\right)=1\cdot\underbrace{p_{X}\left(1\right)}_{4/6}+4\cdot\underbrace{p_{X}\left(2\right)}_{1/6}+9\cdot\underbrace{p_{X}\left(3\right)}_{1/6}=\frac{17}{6}$$

## 7.3 תכונות של תוחלת

**טענה 7.16**

יהיו $X,Y$ משתנים מקריים בדידים בעלי תוחלת עם ערכים ב$\mathbb{R}$.

1. אם $X\geq Y$ אז $\mathbb{E}\left(X\right)\geq\mathbb{E}\left(Y\right)$
2. אם $c\in\mathbb{R}$ (נחשוב עליו כפונקציה קבועה על מרחב ההסתברות) אז $\mathbb{E}\left(c\right)=c$. בפרט אם $X\geq c$ אז $\mathbb{E}\left(X\right)\geq c$
3. אם $c\in\mathbb{R}$ אז: $\mathbb{E}\left(cX\right)=c\mathbb{E}\left(X\right)$
4. $\mathbb{E}\left(X+Y\right)=\mathbb{E}\left(X\right)+\mathbb{E}\left(Y\right)$
5. $\mathbb{E}\left(\left|X\right|\right)\geq\left|\mathbb{E}\left(X\right)\right|$
6. $\mathbb{E}\left(\left|X+Y\right|\right)\leq\left|\mathbb{E}\left(\left|X\right|\right)+\mathbb{E}\left(\left|Y\right|\right)\right|$

**הערה 7.17** נשים לב שאם במקום $\mathbb{E}\left(X\right)$ נרשם: $\int_{a}^{b}g\left(t\right)\mathrm{d}t$ התכונות מוקרות.

**הערה 7.18** אם מרחב הסתברות סופי\בן מנייה הטענה מיידית מתכונות של סכומים. למשל:

$$\mathbb{E}\left(\left|X\right|\right)=\sum_{\omega\in\Omega}P\left(\{\omega\}\right)\left|X\left(\omega\right)\right|\geq\left|\sum_{\omega\in\Omega}P\left(\{\omega\}\right)X\left(\omega\right)\right|=\left|\mathbb{E}\left(X\right)\right|$$

**הוכחה:** (נוכיח רק את 4)
נגדיר $Z=(X,Y)$. $Z:\Omega\to\mathbb{R}^2$. נגדיר: $f:\mathbb{R}^2\to\mathbb{R}$ באופן הבא: $f(a,b)=a+b$ אז: $X+Y=f(Z)$.
וגם: $p_Z=p_{X,Y}$ (לפי הגדרה)

$$\mathbb{E}(f(Z))=\sum_{(a,b)\in\mathbb{R}^2}\underbrace{f(a,b)}_{a+b}\cdot p_Z(a,b)=\sum_a\sum_b(a+b)\,p_{X,Y}(a,b)=$$
$$=\sum_a a\underbrace{\sum_b(p_{X,Y}(a,b))}_{p_X(a)}+\sum_b b\underbrace{\sum_a p_{X,Y}(a,b)}_{p_Y(b)}=\mathbb{E}(X)+\mathbb{E}(Y)$$

צריך להצדיק את השינוי סדר סכימה (אשר בוצע בתחילת השורה השניה של הביטוי הנ"ל) וגם להראות ש $X+Y$ בעל תוחלת.
לשם כך נבצע אוו חישוב עם $g(a,b)=|a+b|$ ויצא:

$$\mathbb{E}(|X+Y|)\le\mathbb{E}(|X|)+\mathbb{E}(|Y|)<\infty$$

כי $X,Y$ בעלי תוחלת.

**הערה 7.19** נשים לב כי למעשה הוכחנו את טענה 6

ולכן הטור מתכנס בהחלט, ולכן מותר היה לשנות את סדר הסכימה ■

**הערה 7.20** באינדוקציה נקבל כי:

$$\mathbb{E}(X_1+\ldots+X_n)=\mathbb{E}(X_1)+\mathbb{E}(X_2)+\ldots+\mathbb{E}(X_n)$$

**הערה 7.21** בד"כ לא נכון ש:

$$\mathbb{E}(XY)=\mathbb{E}(X)\,\mathbb{E}(Y)$$

למשל אם $0<p<1$ ו: $X\sim\mathrm{Ber}(p)$ אז $X^2=X$ לכן:

$$\mathbb{E}\left(X^2\right)=\mathbb{E}(X)=p\neq p^2=\mathbb{E}(X)^2$$

**דוגמה 7.22** יהי $X\sim\mathrm{Bin}(u,p)$. ראינו שיש משתנים מקריים $X_1,\ldots,X_n$ כל אחד עם התפלגות $\mathrm{Ber}(p)$ כך ש:

$$X_1+\ldots+X_n\sim\mathrm{Ber}(n,p)$$

לכן, כיוון ש $\mathbb{E}(X)$ תלוייה רק בהתפלגות, נקבל כי:

$$\mathbb{E}(X)=\mathbb{E}(X_1+\ldots+X_n)$$

נדגיש שזה לא בגלל ש $X=X_1+\ldots+X_n$ כי זה לא נכון! הם בכלל לא באותו מרחב. אלא בגלל שההתפלגויות שוות.

$$=\sum_{i=1}^n\mathbb{E}(X_i)=\sum_{i=1}^n p=n\cdot p$$

**דוגמה 7.23** $n$ כדורים ממוספרים $1,\ldots,n$ מושמים באקראי בתאים הממוספרי $1,\ldots,n$ כדור 1 בתא.
יהי $N$ מספר הכדורים שהושמו בתא בעל אותו מספר, מהו $\mathbb{E}\left(N\right)$?
נגדיר קבוצת סידורים = $\Omega$. ברור כי $\left|\Omega\right|=n!$ .

$$X_i\left(\omega\right)=\begin{cases}1 & \text{אם בסדר }\omega\text{כדור }i\text{נמצא בתא }i\\0 & \text{אחרת}\end{cases}$$

אז: $N=\sum_{i=1}^{n}X_i$.
עבור $i$ נתון:

$$\mathbb{E}\left(X_i\right)=\mathbb{E}\left(\mathbb{1}_{\{\text{כדור }i\text{חזר לתא }i\}}\right)=P\left(\text{כדור }i\text{חזר לתא }i\right)=\frac{\#\left\{\text{סידורים בהם }i\text{לא זז}\right\}}{\left|\Omega\right|}=\frac{\left(n-1\right)!}{n!}=\frac{1}{n}$$

לכן:

$$\mathbb{E}\left(N\right)=\mathbb{E}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathbb{E}\left(X_i\right)=n\cdot\frac{1}{n}=1$$

נשים לב שהתשובה בכלל לא תלויה ב $n$.
כמו כן, כאן $X_i$ אינם בלתי תלויים. למש כי אם:

$$X_1=1,\ldots,X_{n-1}=1$$

אז גם בהכרח: $X_n=1$ .

**דוגמה 7.24** נתונים $n$ וקטורי יחידה $v_1,\ldots,v_n\in\mathbb{R}^d$.

**טענה 7.25**

יש בחירה $\varepsilon_i\in\{\pm1\}$ כך ש:

$$\left\|\sum_{i=1}^{n}\varepsilon_iv_i\right\|\geq\sqrt{n}$$

ויש בחירה $\varepsilon_i'\in\{\pm i\}$ כך ש:

$$\left\|\sum_{i=1}^{n}\varepsilon_i'v_i\right\|\leq\sqrt{n}$$

**הוכחה:** נגדיר משתנים מקריים בלתי תלויים $X_1,\ldots,X_n$ שכל אחד מקבל $+1$ בהסתברות $\frac{1}{2}$ ו $-1$ בהסתברות $\frac{1}{2}$.
נתבונן במשתנה המקרי:

$$R=\left\|\sum_{i=1}^{n}X_iv_i\right\|^2=\left(\sum X_iv_i\right)\cdot\left(\sum X_iv_i\right)=\sum_{i,j=1}^{n}X_iX_j\left(v_i\cdot v_j\right)$$

לכן:

$$\mathbb{E}\left(R\right)=\sum_{i,j=1}^{n}v_i\cdot j_j\mathbb{E}\left(X_iX_j\right)$$

נשים לב שאם $i\neq j$ אז:

$$X_iX_j=\begin{cases}1 & \frac{1}{2}\text{ הסתברות}\\-1 & \frac{1}{2}\text{ הסתברות}\end{cases}$$

לבדוק!
לכן:

$$\mathbb{E}\left(X_iX_j\right)=\frac{1}{2}\cdot1+\frac{1}{2}\left(-1\right)=0$$

כלומר, קיבלנו:

$$\mathbb{E}\left(R\right)=\sum_{i=1}^{n}\underbrace{v_i\cdot v_i}_{=1}\qquad\mathbb{E}\left(\underbrace{X_iX_j}_{=1}\right)=\sum_{i}^{n}1=n$$

כי וקטור יחידה

קיבלנו: $\mathbb{E}\left(R\right)=n$. לכן יש $\omega$ במרחב המדגם כך ש: $R\left(\omega\right)\geq n$ ויש $\omega$ כך ש $R\left(\omega'\right)\leq n$. הערכים $\varepsilon_i=X\left(\omega\right)$, $\varepsilon_i'=X_i'\left(\omega\right)$ הם הסימנים המבוקשים. ∎

19/12/20

**הערה 7.26** אם $A$ קבוצה בת מנייה, $t_a\in\mathbb{R}$ כאשר $a\in A$ .
איך מגדרים ? $\sum_{a\in A}t_a=?$

1. נסדר את $A=\{a_1,a_2,\ldots\}$, אם $\sum_{n=1}^{\infty}|t_{a_n}|<\infty$ אז תכונה זו לא תלויה בסדר. וגם אז הערך $\sum_{n=1}^{\infty}t_{z_n}$ לא תלוי בסדר.
2. בהנחה שיש התכנסות בהחלט, נגדיר: $\sum_{n=1}^{\infty}t_{a_n}=\sum_{a\in A}t_a$ אחרת $\sum_{a\in A}t_a$ לא מוגדר.

**הערה 7.27** השתמשנו גם בכך שאם $A=\{a_{i,j}\}_{i,j=1}^{\infty}$ . אם $\sum_{a\in A}t_a$ מוגדר אז הוא שווה: $\sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty}t_{a_{i,j}}\right)$. (זה דורש נימוק נוסף, תרגיל)

**הערה 7.28** ראינו:

$$\mathbb{E}\left(f\left(x\right)\right)=\sum_{a\in A}f\left(a\right)p_X\left(a\right)$$

צריך לסייג אם התוחלת קיימת (אם אגף ימין מתכנס בהחלט). די לבדוק:

$$\mathbb{E}\left(|f\left(x\right)|\right)=\sum_{a}|f\left(a\right)|\,p_X\left(a\right)<\infty$$

(אותו חישוב כמו בהוכחה).

**הערה 7.29** נאמר שמשתנים מקריים $X,Y$ שווים כמעט תמיד אם:

$$P\left(X=Y\right)=P\left(\{\omega:X\left(\omega\right)=Y\left\{\omega\right\}\}\right)=1$$

ואז ברור מההגדרה ש: $\mathbb{E}\left(X\right)=\mathbb{E}\left(Y\right)$ (בהנחה שקיימים). נובע למשל מכך ש: $p_X=p_Y$.

יהי $X$ משתנה מקרי בדיד עם ערכים ב$\mathbb{R}$. נחשוב על $p_X$ כפונקציה ממשית **הוכחה:**

$$\lim_{a\to-\infty}p_X\left(a\right)=\lim_{a\to\infty}p_X\left(a\right)=0$$

יתרה מזאת:

$$\lim_{a\to-\infty}P\left(X\leq a\right)=\lim_{a\to\infty}P\left(X\geq a\right)=0$$

נשים לב כי המשוואה הראשונה נוסעת מהשניה כי למשל:

$$0 \le p_X(0) \le P(X \ge a)$$

כי $P(X \ge a) \supseteq P(X = a) = p_X(0)$.

צריך להוכיח: $\lim_{a\to\infty} P(X \ge a) = 0$. מספיק להראות שלכל סדרה $a_n \to \infty$ עולה, $P(X \ge a_n) \underset{n\to\infty}{\longrightarrow} 0$. נשים לב כי בהינתן $a_n \to \infty$ עולה, המאורעות $A_n = \{X \ge a_n\}$ יורדות, כלומר: $A_{n+1} \subseteq A_n$ כי ($x \ge a_n \Leftarrow x \ge a_{n+1}$).
לכן מרציפות הסתברות:

$$\lim_{n\to\infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P\underbrace{(\forall n\ X > a_n)}_{\varnothing} = 0$$

$\varnothing$ כי $X$ מקבל ערכים ממשיים ו $a_n \to \infty$. ■

אם ל $X$ יש תוחלת אז $\mathbb{E}(x)$ היא "מרכז המסה" של הגרף $p_X$.

**דוגמה 7.30** נניח ש $X$ מפולגת אחיד על $\{1,\dots,n\}$

$$p_X(a) = \begin{cases} \frac{1}{n} & a \in \{1,\dots,n\} \\ 0 & \text{אחרת} \end{cases}$$

אז:

$$\mathbb{E}(X) = \sum_{k=1}^{n} k \overbrace{p_X(k)}^{1/n} = \frac{1}{n}\sum_{k=1}^{n} k = \frac{1}{n}\cdot\frac{n(n+1)}{2} = \frac{n+1}{2}$$

## 7.4 שונות

**הגדרה 7.31** עבור משתנה מקרי בדיד $X$ עם ערכים ממשיים מגדירים שונות (Variance) ע"י:

$$\operatorname{Var}(X) = \mathbb{E}\left((X-\mu)^2\right)$$

כאשר $\mu = \mathbb{E}(X)$ (בהנחה שהתוחלת הזו קיימת).

נשים לב כי $(X-\mu)^2 \ge 0$ וככל ש $X$ רחוק מהתוחלת התרומה ל $\operatorname{Var}(X)$ גדולה יותר לכן $\operatorname{Var}(X)$ קטן.
שקול עקרונית למשתנה שמרוכז סביב התוחלת שלו (כלומר מקבל בהסתברות גבוהה ערכים קרובים לתוחלת).

### 7.4.1 תכונות של שונות

**טענה 7.32**

א. $\operatorname{Var}(X) \ge 0$ יש שיוויות אס"ם $X = \mu$ כמעט תמיד.
ב. $\operatorname{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$
ג. $\operatorname{Var}(aX+b) = a^2 \operatorname{Var} X$ (צד שמאל קיים אם $\operatorname{Var}(X)$ קיים.

**הוכחה:** ■

1. נסמן $Y = (X-\mu)^2$. אם $Y \ge 0$ אז: $\operatorname{Var}(X) = \mathbb{E}(Y) \ge 0$ אם יש הסתברות חיובית ש $X \ne M$ אז יש $a > 0$ כך ש $p_Y(a) > 0$ ואז:

$$\mathbb{E}(Y) = \sum_b b p_Y(b) \underbrace{\ge}_{\text{כל המחוברים אי שליליים}} a p_Y(a) > 0$$

כלומר $\operatorname{Var}(X) > 0$ (מחוברים אי שליליים כי $p_Y(b) = 0$ לכל $b < 0$ כי $Y$ משתנה מקרי חיובי)

2. חישוב:

$$\mathrm{Var}(X)=\mathbb{E}\left((X-\mu)^2\right)=\mathbb{E}\left(X^2-2\mu X+\mu^2\right)=\mathbb{E}\left(X^2\right)-2\mu\mathbb{E}(X)+\mu^2$$

כאשר $\mu=\mathbb{E}(x)$ ולכן:

$$=\mathbb{E}\left(X^2\right)-\mathbb{E}(X)^2$$

3. כתרגיל

**דוגמה 7.33** $X\sim\mathrm{Ber}(p)$ אז $X=X^2$ (כי $X$ בעל ערכים $0,1$ ) ולכן:

$$\mathbb{E}(X)=\mathbb{E}\left(X^2\right)=p\Rightarrow\mathrm{Var}(X)=\mathbb{E}\left(X^2\right)-\mathbb{E}(X)^2=p-p^2=p(1-p)$$

**טענה 7.34**

אם $X\sim\mathrm{Bin}(n,p)$ אז: $\mathrm{Var}\,X=np(1-p)$ . אם $Y\sim\mathrm{Pois}(\alpha)$ אז: $\mathrm{Var}\,Y=\alpha$. אם $Z\sim\mathrm{Geom}(p)$ אז: $\mathrm{Var}\,Z=\frac{1-p}{p^2}$

**הוכחה:** הוכחה כתרגיל. ■

**משפט 7.35 אי שיוויון מרקוב**

אם $X$ בדיד בעל תוחלת ואם $X\geq 0$ אז:

$$P(X\geq t)\leq\frac{\mathbb{E}(X)}{t}$$

כאשר $t>0$.

**הוכחה: הוכחה ראשונה:**

נחשב את התוחלת של $X$:

$$\mathbb{E}(X)=\sum_a ap_X(a)=\overbrace{\sum_{a<t}ap_X(a)}^{\text{מחוברים אי שליליים}}+\sum_{a\geq t}ap_X(a)\geq\sum_{a\geq t}tp_X(a)\geq t\sum_{a\geq t}p_X(a)=tP(X\geq t)$$

מחלקים ב $t$ ומקבלים:

$$P(X\geq t)\leq\frac{\mathbb{E}(X)}{t}$$

כנדרש.

**הוכחה שנייה:**

נסמן $A=\{X\geq t\}$ . נשים לב כי:

$$X\underbrace{\geq}_{0\leq\mathbb{1}_A\leq 1}X\cdot\mathbb{1}_A\underbrace{\geq}_{\text{אם } \mathbb{1}_A=1 \text{ אז } X\geq t}t\mathbb{1}_A$$

ולכן:

$$\mathbb{E}(X)\geq\mathbb{E}(t\mathbb{1}_A)=t\mathbb{E}(\mathbb{1}_A)=tP(A)=tP(X\geq t)$$

ושוב מחלקים ב$t$ ומקבלים:

$$P(X\geq t)\leq\frac{\mathbb{E}(X)}{t}$$

■

26/12/20

בכל השיעור נדון במקרים של $X$ משתנה מקרי בדיד ו $\mu=\mathbb{E}(x)$.
נזכיר כי הגדרנו את השונות באופן הבא:

$$\operatorname{Var}(X)=\mathbb{E}\left((X-\mu)^2\right)$$

כמו כן, שונות קטנה $\longleftrightarrow$נטייה של $X$ להיות קרוב לתוחלת.

**דוגמה 7.36** $X$ משתנ מקרי ממשי בדיד.

$$X=\begin{cases}a & \frac{1}{2}\ \text{הסתברות}\\ a & \frac{1}{2}\ \text{הסתברות}\end{cases}$$

נבחין כי:

$$\begin{aligned}\mathbb{E}(X) &= \frac{1}{2}a+\frac{1}{2}(-1)=0\\ \operatorname{Var}(X) &= \mathbb{E}\left((X-0)^2\right)=a^2\end{aligned}$$

ראינו:

$$P(|X-\mu|>t)\underset{t\to\infty}{\longrightarrow}0$$

וזה כמובן שווה ל:

$$P(X-\mu>t)+P(X-\mu<-t)$$

כמו כן, בשיעור האחרון ראינו את אי־שוויון מרקוב. אם $X\geq 0$, אז: $P(X\geq t)\leq\frac{\mathbb{E}(x)}{t}$.

**משפט 7.37 אי שוויון שבישב - *Chebychev***

אם $X$ בעל שונות, אז ל$t>0$:

$$P(|X-\mu|\geq t)\leq\frac{\operatorname{Var}(x)}{t^2}$$

**הוכחה:** נסמן $Y=(X-\mu)^2$ אז מההגדרה: $\mathbb{E}(Y)=\operatorname{Var}(X)$. נשים לב כי:

$$\{|X-\mu|\geq t\}=\left\{(X-\mu)^2\geq t^2\right\}=\left\{Y\geq t^2\right\}$$

לכן, לפי אי שיוויון מרקוב:

$$P(Y\geq s)\leq\frac{\mathbb{E}(Y)}{s}=\frac{\operatorname{Var}(X)}{s}$$

לכן:

$$P(|X-\mu|\geq t)=P\left(Y\geq t^2\right)\leq\frac{\operatorname{Var}(X)}{t^2}$$

כנדרש. ■

**הערה 7.38** זה נותן לנו מידע כמותי על קצב ההתכנסות ב:

$$P\left(|X-\mu|>t\right)\underset{t\to\infty}{\longrightarrow}0$$

**הגדרה 7.39 סטיית תקן:** סטיית התקן של $X$ היא:

$$\sigma=\sigma\left(X\right)=\sqrt{\mathrm{Var}\left(X\right)}$$

בסימון הנ"ל:

$$P\left(|X-\mu|>n\sigma\right)\leq\frac{\mathrm{Var}\left(X\right)}{\left(n\sigma\right)^{2}}=\frac{1}{n^{2}}$$

**הערה 7.40** נשאלה שאלה בכיתה: למה בכלל לוקחים ריבוע בVar? הסיבה לכך היא שאנחנו רוצים מרחק מהתוחלת, האפשרות השנייה הייתה ערך מוחלט, אבל הוא לא גזיר באפס. באופן כללי במתמטיקה עדיף לעלות בריבוע מתי שניתן.

**הגדרה 7.41 שונות משותפת - Covariance:** השונות המשותפת של משתנים מקריים $X,Y$ היא:

$$\mathrm{Cov}\left(X,Y\right)=\mathbb{E}\left(\left(X-\mathbb{E}\left(X\right)\right)\left(Y-\mathbb{E}\left(Y\right)\right)\right)$$

**טענה 7.42**

1. $\mathrm{Cov}\left(X,X\right)=\mathrm{Var}\left(X\right)$
2. $\mathrm{Cov}\left(XY\right)=\mathbb{E}\left(XY-X\mathbb{E}\left(Y\right)-Y\mathbb{E}\left(X\right)+\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)\right)$
3. 
   (א) קיום $\mathbb{E}\left(\left(X-a\right)^{2}\right)$ לאיזשהו $a$ $\Leftarrow$ קיום $\mathbb{E}$
   (ב) קיום Var לכל אחד מהמשתנים$\Leftarrow$ קיום Cov

**הוכחה:**
ניסם לב כי: $\mathbb{E}\left(cX\right)=c\mathbb{E}\left(X\right)$ כאשר $c$ קבוע.
$\mathbb{E}\left(X\right),\mathbb{E}\left(Y\right)$ קבועים. לכן:

$$\mathbb{E}\left(X\mathbb{E}\left(Y\right)\right)=\mathbb{E}\left(Y\right)\mathbb{E}\left(X\right)$$

ואז נקבל:

$$\mathbb{E}\left(XY\right)-X\mathbb{E}\left(Y\right)-Y\mathbb{E}\left(X\right)=\mathbb{E}\left(XY\right)-\mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)$$

ולכן, מהלינאריות של התוחלת, 2 נכון.
3. ב. נובע מאי שיוויון קושי־שוורץ אם $u,v$ משתנה מקרי אז:

$$\mathbb{E}\left(UV\right)\leq\sqrt{\mathbb{E}\left(U^{2}\right)}\sqrt{\mathbb{E}\left(V^{2}\right)}$$

וצד ימים קיים$\Leftarrow$צד שמאל קיים.
אם לוקחים $U=X-\mathbb{E}\left(X\right)$ ו : $V=Y-\mathbb{E}\left(Y\right)$ ומקבלים את ב.3. כנדרש. ■

**הערה 7.43** אם $\text{Cov}(X,Y)>0$ זה מעיד על נטייה של $X,Y$ להיות בו זמנית באותו צד של התוחלת. כלומר: $X>\mathbb{E}(X)$ ו: $Y>\mathbb{E}(Y)$ ובאותו אופן עם $<$ . הדבר לא אומר שכך זה תמיד אלא רק נטייה.

**הגדרה 7.44 בלתי מתואמים:** $X,Y$ נקראים בלתי־מתואמים (uncorelated) אם $\text{Cov}(X,Y)=0$.
אחרת הם נקראים מתואמים.

**טענה 7.45**

אם $X,Y$ בלתי תלויים אז: $\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y)$)

**הערה 7.46** אינו שזה לא נכון באופן כללי.

**הוכחה:**

$$\mathbb{E}(XY)=\sum_{a,b}abp_{X,Y}(a,b)\underbrace{=}_{\text{אי תלות}}\sum_{a,b}abp_X(a)p_Y(b)=\sum_a p_X(a)\sum_b p_Y(b)=\mathbb{E}(X)\mathbb{E}(Y)$$

כדי להראות שהתוחלת של $XY$ קיימת חוזרים על החישוב עם $|X|$ ו$|Y|$ (עדיין בלתי תלויים כי הם פונקציה של משתנים מקריים בלתי תלויים) מקבלים:

$$\mathbb{E}(|XY|)=\mathbb{E}(|X||Y|)=\mathbb{E}(|X|)\mathbb{E}(|Y|)$$

אבל הם קיימים, לכן התוחלת של $X,Y$ קיימת. ■

**טענה 7.47**

אם $X,Y$ בלתי תלויים, אז הם בלתי מתואמים

**הוכחה:** נסמן:

$$\begin{aligned}\overline{X}&=X-\mathbb{E}(X)\\\overline{Y}&=Y-\mathbb{E}(Y)\end{aligned}$$

אז:

$$\mathbb{E}\left(\overline{X}\right)=\mathbb{E}\left(\overline{Y}\right)=0$$

ומכיין ש $X,Y$ בלתי תלויים, אז גם $\overline{X},\overline{Y}$ בלתי תלויים (פונקציות של $X,Y$ בהתאמה) לכן:

$$\text{Cov}=\mathbb{E}\left(\overline{XY}\right)=\mathbb{E}\left(\overline{X}\right)\mathbb{E}\left(\overline{Y}\right)=0$$

■

**הערה 7.48** $\text{Cov}(X,Y)=0$ לא גורר אי תלות.
זהו מושג חלש יותר, ונראה את הדוגמה הבאה שמדגימה מקרה לא כזה.

**דוגמה 7.49** יהי $p_{X,Y}$ אזר נתון ע"י הטבלה הבאה:

| $X\diagdown Y$ | -1 | 0 | 1 |
|---|---|---|---|
| 0 | $1/3$ | 0 | $1/3$ |
| 1 | 0 | $1/3$ | 0 |

הם לא בלתי תלויים, אבל $\text{Cov}(X,Y)=0$ .

**הערה 7.50** נניח $\Omega=\{1,\ldots,n\}$ סופית. $P\longleftrightarrow(p_i)_{i=1}^n$ . נניח $p_i>0 \quad \forall i$

$$V=\{X:\Omega\to\mathbb{R}\}\longleftrightarrow\mathbb{R}^n$$

$X\longleftrightarrow(X(1),\ldots,X(n))$ אוסף המשתנים המקריים.
נגדיר את המכפלה הפנימית: $u,v\in\mathbb{R}^n$.

$$uv=\sum_{i=1}^n p_iu_iv_i$$

$$\langle X,Y\rangle=\sum_{i=1}^n p_iX(i)Y(i)$$

זה מקיים את כל האקסיומות\תכונות של המכפלה הפנימית הסטנדרטית.
כעת:

$$\mathrm{Var}(X)=\langle X,X\rangle=\|x\|^2$$

(אורך\נורמה ביחס למכלה הזו)

$$\mathrm{Cov}(X,Y)=\left\langle\overline{X},\overline{Y}\right\rangle$$

משמע, $\mathrm{Cov}(X,Y)>0\Leftarrow$ "באותו כיוון".

01/01/20

**הערה 7.51**

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\,\mathrm{Cov}(X,Y)$$

נבחין כי זה דומה ל:

$$\|u+v\|^2=\|v\|^2+\|v\|^2+2u\cdot v$$

## 7.5 תוחלת מותנית

בהינתן $X$ משתנה מקרי ממשי בדיד. ו $Y$ משתנה מקרי בדיד. $b$ ערך של $Y$. $P(Y=b)>0$

**הגדרה 7.52**

$$\mathbb{E}(X\mid Y=b)=\sum_a a\cdot p_{X|Y}(a\mid b)$$

כאשר:

$$p_{X|Y}(a\mid b)=P(X=a\mid Y=b)=\frac{P(X=a,Y=b)}{P(Y=b)}$$

בתנאי שמתכנס בהחלט.

אם $B$ מאורע $P(B)>0$ .

$$\mathbb{E}(X\mid B)=\sum_a a\cdot P(x=a\mid B)$$

כאשר:

$$P\left(x=a\mid B\right)=\frac{P\left(\left\{x=a\right\}\cap B\right)}{P\left(B\right)}$$

אם נגדיר $Y=\mathbb{1}_B$ אז: $\mathbb{E}\left(X\mid B\right)=\mathbb{E}\left(X\mid Y=1\right)$.

**הערה 7.53** אם $\Omega$ סופי:

$$\mathbb{E}\left(X\mid B\right)=\sum_{\omega\in B}x\left(\omega\right)P\left(\left\{\omega\right\}\mid B\right)$$

**הערה 7.54**

$$\mathbb{E}\left(X\mid Y=b\right)=\mathbb{E}_Q\left(X\right)$$

כאשר כאן אנו מסמנים את $Q$ = מידת ההסתברות המותנית על $\{Y=b\}$

$$Q\left(A\right)=\frac{P\left(A\cap B\right)}{P\left(B\right)}$$

כלומר: בהינתן מרחב הסתברות $\left(\Omega,\mathcal{F},Q\right)$ אז $X:\Omega\to\mathbb{R}$. ו $\mathbb{E}_Q$ היא תוחלת ביחס ל$Q$.

**<u>בדיקה:</u>**

$$\mathbb{E}_Q\left(X\right)=\sum_a aQ\left(X=a\right)=\sum a\underbrace{\frac{P\left(\left\{X=a\right\}\cap\left\{Y=b\right\}\right)}{P\left(Y=b\right)}}_{p_{X\mid Y}\left(a\mid b\right)}=\mathbb{E}\left(X\mid Y=b\right)$$

**מסקנה 7.55**

$$\overbrace{\mathbb{E}\left(X_1+X_2\mid Y=b\right)}^{\mathbb{E}_Q\left(X_1+X_2\right)}=\underbrace{\mathbb{E}\left(X_1\mid y=b\right)}_{\mathbb{E}_Q\left(X_1\right)}+\underbrace{\mathbb{E}\left(X_2\mid Y=b\right)}_{\mathbb{E}_Q\left(X_2\right)}$$

**משפט 7.56**

$$\mathbb{E}\left(X\right)=\sum_b\mathbb{E}\left(X\mid Y=b\right)P\left(Y=b\right)$$

וגם:

$$\mathbb{E}\left(X\mid Y=b\right)=\sum_c\mathbb{E}\left(X\mid Y=b,Z=c\right)\cdot P\left(Z=c\mid Y=b\right)$$

(אם משתמשים ב $\mathbb{E}_Q$ אז המשוואה השניה כמעט נובעת מהמשוואה הראשונה).

**הוכחה:**

$$\sum_b\mathbb{E}\left(X\mid X=b\right)P\left(Y=b\right)=\sum_b\sum_a a\underbrace{p_{X\mid Y}\left(a\mid b\right)\cdot p_Y\left(b\right)}_{p_{X,Y}\left(a,b\right)}=\sum_a a\underbrace{\left(\sum_b p_{X,Y}\left(a,b\right)\right)}_{p_X\left(a\right)}=\mathbb{E}\left(X\right)$$

כנדרש. (יש להצדיק שינוי סדר סכימה ע"י שימוש בהנחה של התכנסות בהחלט). ■

**דוגמה 7.57** יהיו $X_1, X_2, X_3, \ldots$ הטלות הוגנות בלתי תלויות של קוביה יהי $N = \min\left\{n \geq 1 \mid X_n = 6,\ X_{n+1} = 5\right\}$, נרצה לשאות מהו $\underbrace{\mathbb{E}(N)}_{\alpha} = ?$ (כלומר כמה הטלות בממוצא עד שנקבל שש־בש.

$$\mathbb{E}(N) = \mathbb{E}(N \mid X_1 = 6)\underbrace{P(X_1 = 6)}_{1/6} + \mathbb{E}(N \mid X_1 \neq 6) \cdot \underbrace{P(X_1 \neq 6)}_{5/6}$$

<u>**טענה:**</u> $\mathbb{E}(N \mid X_1 \neq 6) = \mathbb{E}(N) + 1$

$$\mathbb{E}(N \mid X_1 \neq 6) = \mathbb{E}(N')$$

כאשר $N' = \min\left\{n \geq 2 \mid X_n = 6,\ X_{n+1} = 5\right\}$ וברור כי זה $\sim N+1$.

$$\alpha = \frac{5}{6}(\alpha + 1) + \frac{1}{6}\mathbb{E}(N \mid X_1 = 6)$$

$$\overbrace{\mathbb{E}(N \mid X_1 = 6)}^{\beta} = \mathbb{E}(N \mid X_1 = 6, X_2 = 5)\underbrace{P(X_2 = 5 \mid X_1 = 6)}_{1/6} + \underbrace{\mathbb{E}(N \mid X_1 = 6)}_{1}$$

$$\overbrace{\mathbb{E}(N \mid X_1 = 6, X_2 = 6)}^{\beta+1} \underbrace{P(X_2 = 6 \mid X_1 = 6)}_{1/6} + \underbrace{\mathbb{E}(N \mid X_1 = 6, X_2 \neq 6, 5)}_{\alpha+2} \underbrace{P(X_2 = 6 \mid X_1 \neq 6, 5)}_{4/6}$$

כלומר:

$$\beta = \frac{1}{6} + \frac{1}{6}(\beta + 1) + \frac{4}{6}(\alpha + 2) \Rightarrow \beta = \frac{1}{5}(10 + 4\alpha)$$

$$\alpha = \frac{5}{6}(\alpha + 1) + \frac{1}{6}\left(\frac{1}{5}(10 + 4\alpha)\right) \Rightarrow \alpha = 35$$

**הערה 7.58** ישנו אובייקט המסומן:

$$\mathbb{E}(X \mid Y)$$

וזה משתנה מקרה שבנקודה $\omega \in \Omega$ ערכו:

$$\mathbb{E}(X \mid Y = b)$$

עבור: $b = Y(\omega)$

# 8 משתנים מקריים "רציפים" (לא בדידים)

**תזכורת 8.1** $\mathcal{B}$ היא $\sigma$ אלגברה בורל על $\mathbb{R}$. מוגדרת $\sigma$־אלגברה מינימלית המיכלה את הקטעים:

$$\mathcal{B} = \bigcap_{\mathcal{G}\ \sigma\text{-אלגברה על } \mathbb{R} \text{ המיכלה את כל הקטעים}} \mathcal{G}$$

בהנתן $(\Omega, \mathcal{F}, P)$ מרחב הסתברות. $X : \Omega \to \mathbb{R}$ מדיד אם: $X^{-1}(B) \in \mathcal{F}$ לכל $B \in \mathcal{B}$.

**למה 8.2**

$X$ מדידה $\iff$ $X^{-1}(I)\in\mathcal{F}$ לכל קטע. $P(a\leq X\leq b)$ (או שוויונות חלשים) מוגדר היטב לכל $a\leq b$.

**משפט 8.3**

הערכים $P(a\leq X\leq b)$ מאפיינים את $P_X$ (ידת ההתסברות על $(\mathbb{R},\mathcal{B})$ הנתונה ע"י $P_X(A)=P\left(X^{-1}(A)\right)$).

**הגדרה 8.4** אם $X$ משתנה מקרי ממשי. פונקציית ההתפלגות: $F_X:\mathbb{R}\to\underbrace{\mathbb{R}}_{[0,1]}$ היא הפונקציה: $F_X(t)=P(X\leq t)=P\left(X^{-1}((-\infty,t))\right)$.

**דוגמה 8.5** $X\sim\text{Bin}\left(3,\frac{1}{2}\right)$.

$$p_X(a)=\begin{cases}1/8 & a=0\\ 3/8 & a=1\\ 3/8 & a=2\\ 1/8 & a=3\\ 0 & \text{אחרת}\end{cases}$$

$$F_X(t)=\begin{cases}0 & t<0\\ P(X\leq t)=\frac{1}{8} & 0\leq t<1\\ \frac{1}{8}+\frac{3}{8} & 1\leq t<2\\ \frac{1}{8}+\frac{3}{8}+\frac{3}{8} & 2\leq t<3\\ 1 & 3\leq t\end{cases}$$

**תכונות**

1. $F_X$ לא ירדת (עולה חלש) כי:

$$s\leq t\Rightarrow\{x\leq s\}\subset\{x\leq t\}\Rightarrow F_X(s)P(\{x\leq s\})\leq P(\{X\leq t\})$$

2. $F_X(t)=1-P(X>t)$ (משלים ל $\{x\leq t\}$)

3. $\lim_{t\to-\infty}F_X(t)=0$

**הוכחה:** לכל $t_n\to-\infty$ מונוטונית. המאורעות $\ldots\supseteq\{X\leq t_{n+1}\}\supseteq\{X\leq t_n\}$, $\bigcap\{X\leq t_n\}=\varnothing$. מרציפות של $P$:

$$0=P(\varnothing)=P\left(\bigcap\{X\leq t_n\}\right)=\lim_{n\to\infty}P(X\leq t_n)=\lim_{n\to\infty}F_X(t_b)$$

זה נכון לכל $t_n\to-\infty$ יורדת $\Leftarrow$ כאשר $t\to-\infty$. ∎

4. $\lim_{t\to\infty}F_X(t)=\lim_{t\to\infty}(1-P(X>t))=1$ (הוכחה דומה לקודם)

5. $F_X$ רציפה מימין.

**הוכחה:** בהנתן $t$, לכל סדרה $t_n \searrow t$ (יורדת ל$t$)

$$\lim_{n\to\infty} F_x(t_n) = \lim_{n\to\infty} P(X \le t_n) = P\underbrace{\left(\bigcap_{n=1}^{\infty}\{x \le t_b\}\right)}_{(-\infty,t]} = F_x(t)$$

וזו ההגדרה לרציפות מימין. ■

**הערה 8.6** בהדוגמה ריאים שבאופן כללי $F_X$ אינה רציפה משמאל.

**טענה 8.7**

$F_X$ מאפיין (קובע) את $P_X$

**הוכחה:** די להראות ש $F_X$ קובעת את:$P(a \le X \le b)$ לכל $a \le b$.

ראשית נבחין כי:

$$P(a < X \le b) = P(X \le b) - P(X \le a) = P(\{X \le b\} \setminus \{X \le a\}) = F_X(b) - F_X(a)$$

כעת:

$$P\underbrace{(x \in [a,b])}_{\bigcap_{n=1}^{\infty}\left\{X \in \left(a-\frac{1}{n},b\right]\right\}} = \lim_{n\to\infty} F_X(b) - F_X\left(a - \frac{1}{n}\right)$$

אגף ימין תלוי רק ב $F_X$ ולכן כנדרש. ■

**משפט 8.8**

לכל פונקציה $F : \mathbb{R} \to [0,1]$ שהיא עולה חלש, רציפה מימין

$$\lim_{t\to\infty} F(t) = 0,\ \lim_{t\to\infty} F(t) = 1$$

אזי קיים משתנה מקרי $X$ כך ש $F = F_X$.

## 8.1 פונקציות צפיפות (לא כללי)

**הגדרה 8.9** פונקציה $f : \mathbb{R} \to [0,\infty)$ נקראת פונקציה צפיפות של מתנה מקרי $X$ אם לכל $a \le b$:

$$P(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x$$

(מספיק ש$f$ רציפה למקוטעין)

**הערה 8.10** $f$ אינה יחידה (למשל אפשר לשנות ערכים במספר סופי של נקודות). בכל זאת נסמן $f_X$ ונסכים לא להבחין בין גרסאות של הצפיפות שיש לאן אותו אינטגרל בכל קטע.

**הערה 8.11**

1.

$$1 = P\left(-\infty \le X \le \infty\right) = \lim_{n\to\infty} P\left(-n \le X \le n\right) = \lim_{n\to\infty} \int_{-n}^{n} f_X\left(x\right) \mathrm{d}x = \lim_{n\to\infty} \int_{-\infty}^{\infty} f_X\left(x\right) \mathrm{d}x$$

שיוויון אחרון מהגדרה של אינטגרל לא אמיתי, עבור $f_X \ge 0$ (לא ההגדרה לכל אינטגרל לא אמיתי, תזכורת מאינפי 2: צריך נקודה באמצע ולפרק לשני אינטגרלים)

2. אם ל$X$ עם פונקציות צפיפות $f_X$ אז $P\left(X = x_0\right) = 0$ לכל $x_0$ (בפרט $X$ לא דיסקרטי)

**הוכחה:** נניח לפשטות $f_X$ רציפה, אז היא חסומה. בסביבה של $x_0$ . $M$ חסם.

$$P\left(X = \{x_0\}\right) = \lim_{n\to\infty} P\left(x_0 - \frac{1}{n} \le X \le x_0 + \frac{1}{n}\right) = \lim_{n\to\infty} \int_{x_0-\frac{1}{n}}^{x_0+\frac{1}{n}} f_X\left(x\right) \mathrm{d}x = 0$$

■

**משפט 8.12**

אם $f : \mathbb{R} \to [0,\infty)$ מקיימת: $\int_{-\infty}^{\infty} f = 1$ (ואינטגרבילית) אז $f = f_X$ לאיזשהו משתנה מקרי $X$.

**דוגמה 8.13**

$$f = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{אחרת} \end{cases}$$

$X$ כך ש $f = f_X$ נקרא אחיד ב $[0,1]$, $X \sim U\left[0,1\right]$. אם $0 \le a < b \le 1$ אז:

$$P\left(a \le x \le b\right) = \int_a^b 1 \mathrm{d}x = b - a = [a,b] \text{ אורך}$$

**דוגמה 8.14** יהי $\lambda > 0$:

$$f\left(x\right) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

אז: $\int_{-\infty}^{\infty} f = 1$ אם $X$ משתנה מקרי $f = f_X$ אומרים $X$ אקספוננציאלי עם פרמטר $\lambda$.

$$X \sim \exp\left(\lambda\right)$$

**טענה 8.15**

אם $s < t$ , $X \sim \exp\left(\lambda\right)$ אז:

$$P\left(X > t \mid X > s\right) = P\left(X > t - s\right)$$

(תכונת חוסר הזיכרון)

**דוגמה 8.16** התפלגות נורמלית\גאוסית יהיו $\mu \in \mathbb{R}$, $\sigma \geq 0$ פרמטרים אומרים ש$X$ נורמלים עם פרמטרים $\mu, \sigma^2$ :

$$X \sim N\left(\mu, \sigma^2\right)$$

אם:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

02/01/20

**תזכורת 8.17** $X : \Omega \to \mathbb{R}$ משתנה מקרי. $P_X : \mathcal{B} \to [0,1]$ מידת הסתברות. $F_X(t) = P(X \leq t) = P_X((-\infty, t])$ כאשר $F_X : \mathbb{R} \to [0,1]$. בעזרת הפונקצייה הנ"ל אנו יכולים לחשב את ההסתברות של קטעים לדוגמה: $P_X((a,b]) = P(a < x \leq b) = F_X(b) - F_X(a)$.
במקרה ש $X$ דיסקרטי(אטומי) אנו יודעים כי $P_X$ מאופיין ע"י $p_X$ כאשר $P_X(\{i\}) = P(X = i) = p_X(i)$.
במקרה ש $X$ רציף **בהחלט**, קיימת פונקציית צפיפות $f_X : \mathbb{R} \to [0, \infty)$ אינטגרבילית (רציף למקוטעין) שמקיימת $\int_a^b f_X(x)\,\mathrm{d}x = P(X \in [a,b]) = P_X([a,b])$. במקרה הרציף: $P_X([x, x + \Delta x]) \approx f(x)\Delta x$. במקרה הרציף (לא אטומי, ולא בהחלט הפעם): $P(X = x_0) = 0$ $\forall x_0 \in \mathbb{R}$ אבל לא מתקיימת פונקציית צפיפות.

אם $X$ משתנה מקרי ממשי אז:

$$P_X(A) = \inf\{P_X(U) \mid A \subseteq U \text{ איחוד של קטעים } U\}$$

כאשר $A$ רבוצת בורל. ברור כי $\leq$ כי: $P_X(A) \leq P_X(U)$ ברגע ש $A \subseteq U$. אבל זה שיש שיוויון זה לא כזה ברור.
בהינתן $f_X$ פונקציית צפיפות של $X$. אם $A$ קבוצה מספיק יפה אז:

$$P(X \in A) = P_X(A) = \int_A f_X \underbrace{=}_{A = \uplus[a_i, b_i]} \sum \int_{a_i}^{b_i} f_X(x)\,\mathrm{d}x$$

מה אם $A = \mathbb{Q}$. מה זה $\int_{\mathbb{Q}} f_X$? נגדיר:

$$\int_{\mathbb{Q}} f_X = \int \mathbb{1}_{\mathbb{Q}} f(x)\,\mathrm{d}x$$

$$\begin{cases} 0 & x \notin \mathbb{Q} \\ f_X(x) & x \in \mathbb{Q} \end{cases}$$

בד"כ לא אינטגרבילית (בד"כ לא רציפה בשום מקום) לכן המציאו אינטגרל לבג (lebesgue).

**דוגמה 8.18** $X \sim U[0,1]$:

$$f_X(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{אחרת} \end{cases}$$

$$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(x)\,\mathrm{d}x = \begin{cases} 0 & t \leq 0 \\ \int_0^t \cancel{f_X(x)}^{1}\,\mathrm{d}x = t & 0 \leq t \leq 1 \\ \int_0^1 1\,\mathrm{d}x = 1 & t \geq 1 \end{cases}$$

**דוגמה 8.19** $X\sim\exp\left(\lambda\right)$.

$$f_X\left(x\right)=\begin{cases}0 & x\leq0\\ \lambda e^{-\lambda x} & x>0\end{cases}$$

$$F_X\left(t\right)=\begin{cases}0=\int_{-\infty}^{t}0 & t\leq0\\ \int_0^t\lambda e^{-\lambda x}\mathrm{d}x=1-e^{-\lambda t} & t>0\end{cases}$$

### 8.1.1 הקשר בין $f_X$ ו $F_X$

נניח של$X$ יש צפיפות $f_X$ אז:

$$F_X\left(t\right)=\int\limits_{-\infty}^{t}f_X\left(x\right)\mathrm{d}x$$

אם $t_0$ נקודות רציפות של $f_x$. מהמשפט היסודי של אינפי:

$$F'_X\left(t_0\right)=f_x\left(t_0\right)$$

בהנחה ש $f_X$ רציפה למקוטעין $F'_X$ מוגדר למעט סדרת נקודות $\ldots<t_{-1}<t_0<t_1<\ldots$ (זה בסדר כי ממילא $f_X$ אינה יחידה ומותר לשנות את ערכיה בסדרה כזו).

**הערה 8.20** יש דוגמאות של משתנה מקרי $X$ כך ש $F_X$ גזירה בהמון נקודות (למשל באיחוד של קטעים פתוחים בעלי אורך $b-a$ בכל קטע $[a,b]$). קיום $F'_X$ בהרבה נקודות לא גורר קיום פונקציית צפיפות. (לדוגמה, משתנה דיסקרטי)

# 9 תוחלת של משתנה רציף

**הגדרה 9.1** יהי $X$ משתנה מקרי בעל צפיפות $f_X$ התוחלת של $X$ הוא:

$$\mathbb{E}\left(X\right)=\int\limits_{-\infty}^{\infty}xf_X\left(x\right)\mathrm{d}x$$

(בהנחה שהאינטגרל מתכנס בהחלט, כלומר: $\int\limits_{-\infty}^{\infty}\left|x\right|f_X\left(x\right)\mathrm{d}x<\infty$ )

השוואה עם המקרה הדיסקרטי: יהי $Y$ משתנה דיסקרטי, בהנחה ש $f_X=0$ מחוץ ל$[a,b]$.

$$\mathbb{E}\left(Y\right)=\sum_y y\overbrace{p_Y\left(y\right)}^{P(Y=y)}$$

ואילו:

$$\mathbb{E}\left(X\right)=\int\limits_a^b xf_X\left(x\right)\mathrm{d}x\sim\underbrace{\sum x_i\overbrace{f\left(x_i\right)\Delta x_i}^{\sim P(X\in[x_i,x_i+\Delta x_i])}}_{x_i \text{ ממוצע משוקלל של}}$$

(בהנחה ש $x_i$ חלוקה עדינה של $[a,b]$)

**דוגמה 9.2**

$$f_X(x)=\begin{cases}\frac{1}{b-a} & x\in[a,b]\\ 0 & \text{אחרת}\end{cases}$$

בדיקה מראה: $\int\limits_{-\infty}^{\infty} f_x=1$. אם: $X\sim U[a,b]$ :$f_X=f$ נחשב תוחלת:

$$\mathbb{E}(X)=\int\limits_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x=\int\limits_a^b x\frac{1}{b-a}\mathrm{d}x=\frac{b^2-a^2}{2}\frac{1}{b-a}=\frac{b+a}{2}$$

כלומר אמצע הקטע (לא מפתיע)

**דוגמה 9.3** $X\sim\exp(\lambda)$:

$$f_X(x)=\begin{cases}0 & x\le 0\\ \lambda e^{-\lambda x} & x>0\end{cases}$$

$$\mathbb{E}(X)=\int\limits_{-\infty}^{\infty} xf_X(x)\,\mathrm{d}x=\int\limits_0^{\infty} x\lambda e^{-\lambda x}\mathrm{d}x=\lim_{n\to\infty}\int\limits_0^{n} x\lambda e^{-\lambda x}\mathrm{d}x=\lim_{n\to\infty}\left(xe^{-\lambda x}-\frac{1}{\lambda}e^{-\lambda x}\right)_0^n=\frac{1}{\lambda}$$

**דוגמה 9.4** $X\sim N(\mu,\sigma)$. $f_X(x)=ce^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$$\mathbb{E}(X)=\int\limits_{-\infty}^{\infty} xce^{-\frac{(x-\mu)^2}{2\sigma^2}}=\underbrace{\int\limits_{-\infty}^{\infty}(x-\mu)ce^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x}_{=0}+\int\limits_{-\infty}^{\infty}\mu f_X(x)\,\mathrm{d}x=\mu\int\limits_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x=\mu$$

08/01/20

**טענה 9.5**

$P(X=x_0)>0 \iff F_X$ לא רציפה ב $x_0$ (בתרגיל)

סיטואציה $X$ משתנה מקרי ממשי עם צפיפות $f_X$. $g:\mathbb{R}\to\mathbb{R}$ מונוטונית ממש או מונוטונית ממש למקוטעין + רציפה למקוטעין.

$Y=g(X)$ משתנה מקרי חדש $\Omega\to\mathbb{R}$. מה ההתפלגות צפיפות?

**דוגמה 9.6** $X\sim U[0,1]$ נקבע $n\in\mathbb{R}$, $Y=X^n$

$$F_Y,\ f_Y=?$$

$$F_y(t)=P(Y\le t)=P(X^n\le t)=P\left(X\le t^{1/n}\right)=\int\limits_0^{t^{1/n}}\mathbb{1}_{\{0,1\}}$$

אפשר להניח כי $t\ge 0$ כי $Y\ge 0$ כי $X\ge 0$.

$$=\begin{cases}0 & t\le 0\\ t^{1/n} & 0<t\le 1\\ 1 & t>1\end{cases}$$

וכמו כן:

$$f_Y = F_Y' = \begin{cases} 0 & t \le 0 \\ \frac{1}{n} t^{1/n - 1} & 0 < t \le 1 \\ 0 & t \ge 1 \end{cases}$$

מדוע בכלל ל $Y$ יש צפיפות?
אותם שיקולים מראים שבאמת:

$$\int_a^b f_Y(y)\,\mathrm{d}y = P(Y \in [a,b]) = P(a \le Y \le b) \underbrace{=}_{a\ge 0 \text{ בהנחה}} P\left(a^{1/n} \le X \le b^{1/n}\right)$$

**דוגמה 9.7** $X$ משתנה מקרי עם צפיפות $f_X$ (רציפה) יהי $Y = X^2$ , $F_Y = ?$, $f_Y = ?$

$$P(Y \le t) = P\left(X^2 \le t\right) = P\left(-\sqrt{t} \le X \le \sqrt{t}\right) = F_X\left(\sqrt{t}\right) - F_X\left(-\sqrt{t}\right)$$

($t \ge 0$ אחרת $P = 0$).
לכן (צריך להראות למה ל$Y$ יש צפיפות) מכלל השרשרת:

$$\frac{\mathrm{d}}{\mathrm{d}t} F_Y(t) = \frac{1}{2\sqrt{t}} f_X\left(\sqrt{t}\right) + \frac{1}{2\sqrt{t}} f_X\left(-\sqrt{t}\right)$$

**משפט 9.8**

יהי $X$ משתנה מקרי עם צפיפות רציפה $f_X$. $g : \mathbb{R} \to \mathbb{R}$ מונוטונית ממש גזירה ברציפות (ולא מתאפסת). יהי $Y = g(X)$. ל$Y$ יש צפיפות רציפה.

$$f_Y(y) = \begin{cases} \left|\left(g^{-1}\right)'(y)\right| \cdot f_X\left(g^{-1}(y)\right) & Y \text{ בטווח של } y \\ 0 & \text{אחרת} \end{cases}$$

**הוכחה:**

$$F_Y(t) = P(Y \le t) = \underbrace{P\left(X \le g^{-1} t\right)}_{\ge \text{ אם יורדת}} = \underbrace{F_X\left(g^{-1}(t)\right)}_{1 - F_X\left(g^{-1}(t)\right) \text{ אם יורדת}}$$

גוזרים. ומכלל השרשרת נקבל, אם מונוטונית עולה:

$$\left(g^{-1}\right)'(t)\, f_X\left(g^{-1}(t)\right)$$

ובקרה של יורדת:

$$-\left(g^{-1}\right)'(t)\, f_X\left(g^{-1}(t)\right)$$

בכל מקרה:

$$\left|\left(g^{-1}\right)'(t)\right| f_X\left(g^{-1}(t)\right)$$

ואותה הערה לגבי קיום של $f_Y$. ∎

**משפט 9.9**

אם $X$ משתנה מקרי עם ציפיות $f_X$ רציפה. $g$ מונוטונית ממש למקוטעין וגזירה ברציפות ($g' \neq 0$) אז: $Y = g(x)$ בעל צפיפות:

$$f_Y(y) = \sum_{X \in g^{-1}(y)} \left| \left(g_x^{-1}\right)'(y) \right| f_X(x)$$

(כאשר $g_x^{-1}$ הוא "הופכי לוקאלי" כלומר, ההופכי בקטע בוא אנו נמצאים אשר בו $g$ מונוטונית).

במקרה הבדיד: $X = g(X)$, $p_Y(j) = \sum_{i \in g^{-1}(j)} p_X(i)$ .

הסקנו כי:

$$\mathbb{E}(Y) = \sum_i g(i) p_X(i)$$

**משפט 9.10**

אם $X$ משתנה מקרה בעל צפיפות $f_X$ רציפה, אז $g : \mathbb{R} \to \mathbb{R}$ כמו במשפט הקודם.
יהי $Y = g(X)$, אז:

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} g(x) f_X(x) \,\mathrm{d}x$$

הפסקה קלה, בואו נדבר על אינטגרציה בשני משתנים:
$f : \mathbb{R}^2 \to \mathbb{R}$, נגדיר $D = [a,b] \times [c,d] =$ מלבן (רציפה).

$$\iint_D f(x,y) \,\mathrm{d}x\mathrm{d}y = ?$$

במקרה של משתנה אחד, הגדרנו חלוקה $\{t_i\}_{i=1}^{\infty}$ כך ש:

$$\int_a^b g(t) \,\mathrm{d}t \approx \sum_{i=1}^n g(t_i) \Delta t_i$$

כאשר $n \to \infty$.
במקרה של דו מימד:

$$\iint f \approx \sum_{i,j=1}^n f(x_i, y_j) \overbrace{\Delta x_i \Delta y_j}^{\Delta S_{i,j}} = \sum_j \left( \sum_i f(x,y) \Delta x_i \Delta y_j \right) = \sum_j \Delta y_j \underbrace{\left( \sum_i f(x_i, y_j) \Delta x_i \right)}_{\underbrace{\approx \int f(x, y_j) \,\mathrm{d}x}_{F(y_j)}} =$$

$$\sum_j \Delta y_j F(y_j) \approx \int_c^d F(y) \,\mathrm{d}y$$

וקיבלנו את משפט פוביני:

$$\iint_D f = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x\right)\mathrm{d}y = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y\right)\mathrm{d}x$$

**משפט 9.11**

אם $X$ משתנה מקרי עם תוחלת וציפופת (רציפה למקוטעין)

$$\mathbb{E}(X) = \int_0^\infty \left(\overbrace{1-F_X(t)}^{P(X>t)} - \overbrace{F_X(-t)}^{P(X\le -t)}\right)\mathrm{d}t$$

**הוכחה:**

$$\int_0^\infty P(X>t)\,\mathrm{d}t = \int_0^\infty \left(\int_t^\infty f_X(x)\,\mathrm{d}x\right)\mathrm{d}t = \int_0^\infty\int_0^x f_X(x)\,\mathrm{d}t\mathrm{d}x = \int_0^\infty f_X(x)\int_0^x \mathrm{d}t\mathrm{d}x =$$

$$\int_0^\infty xf_X(x)\,\mathrm{d}x \le \int_{-\infty}^\infty |x|\,f_X(x)\,\mathrm{d}x < \infty$$

כי: $\mathbb{E}(X)$ קיימת.
ובאותו אופן:

$$\int_0^\infty P(X\le -t)\,\mathrm{d}t = -\int_{-\infty}^0 xf_X(x)\,\mathrm{d}x$$

סוכמים וקיבלנו:

$$\int_0^\infty \left(\overbrace{1-F_X(t)}^{P(X>t)} - \overbrace{F_X(-t)}^{P(X\le -t)}\right)\mathrm{d}t = \int_0^\infty xf_X(x)\,\mathrm{d}x + \int_{-\infty}^0 xf_X(x)\,\mathrm{d}x = \int_{-\infty}^\infty xf_X(x)\,\mathrm{d}x = \mathbb{E}(X)$$

כנדרש. ■

נחזור למשפט אשר עצרנו בו, כלומר:

**משפט 9.12**

אם $X$ משתנה מקרה בעל צפיפות $f_X$ רציפה, אז $g:\mathbb{R}\to\mathbb{R}$ כמו במשפט הקודם.
יהי $Y=g(X)$ בעלת צפיפות (למשל אם $g$ גזירה למקוטעין), אז:

$$\mathbb{E}(Y) = \int_{-\infty}^\infty g(x)\,f_X(x)\,\mathrm{d}x$$

**הוכחה:** נחשב, אנו מניחים כי $Y$ בעל צפיפות, לכן הלמה המשפט הקודם רלוונטי, ולכן:

$$\mathbb{E}(Y) = \int_0^\infty P(Y>t)\,\mathrm{d}t - \int_0^\infty P(Y\le -t)\,\mathrm{d}t = \int_0^\infty \int_{\underbrace{\{x:\ g(x)>t\}}_{\text{איחוד של קטעים } g^{-1}((t,\infty))}} f_Y(y)\,\mathrm{d}y\mathrm{d}t - \int_0^\infty\int_{\{x:\ g(x)\le -t\}} f_Y(y)\,\mathrm{d}y\mathrm{d}t$$

אלו הם אינטגרלים כפולים. הראשון על: $\{0 \le t < \infty,\ x \mid\ g(x > t)\}$ וזה שווה ל: $\{x \in \mathbb{R},\ t \mid 0 \le t < g(x)\}$
והשני, אינטגרל כפול על:

$$\{0 \le t < \infty,\ x \mid\ g(x) < -t\} = \{x \in \mathbb{R},\ t \mid\ 0 < t \le -g(x)\}$$

נשתמש בפוביני ונקבל:

$$= \int_{-\infty}^{\infty} \int_{0}^{\max(0,g(x))} f_X(x)\,\mathrm{d}t\mathrm{d}x - \int_{-\infty}^{\infty} \int_{0}^{\max(0,-g(x))} f_X(x)\,\mathrm{d}t\mathrm{d}x =$$

$$\int_{-\infty}^{\infty} f_X(x) \int_{0}^{\max(0,g(x))} \mathrm{d}t\mathrm{d}x - \int_{-\infty}^{\infty} f_X(x) \int_{0}^{\max(0,-g(x))} \mathrm{d}t\mathrm{d}x = \int_{-\infty}^{\infty} f_X(x) \underbrace{[\max(0,g(x)) - \max(0,-g(x))]}_{g(x)} \mathrm{d}x$$

כנדרש. ■

**הערה 9.13** אותו חישוב מבטיח שאם:

$$\int_{-\infty}^{\infty} |g(x)|\, f_X(x)\,\mathrm{d}x < \infty$$

(אינטגרנד אי שלילי) אז $\mathbb{E}(Y)$ קיימת. $Y = g(x)$.

**הערה 9.14** הלמה הקודמת:

$$\mathbb{E}(X) = \int_{0}^{\infty} \left( \overbrace{1 - F_X(t)}^{P(X>t)} - F_X(-t) \right) \mathrm{d}t$$

אנאלוג של אם $Z$ משתנה מקרי עם ערכים ב $\mathbb{N}$ אז:

$$\mathbb{E}(Z) = \sum_{i=1}^{\infty} P(Z \ge 1)$$

## 9.1 שונות

שונות של משתנה מקרי $X$ בעל צפיפות רציפה למקוטעין $f_X$ ובעל תוחלת $\mu = \mathbb{E}(X)$ הוא:

$$\mathrm{Var}(X) = \mathbb{E}\left((X-\mu)^2\right)$$

בהנחה שזה קיים.
מהמשפט:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (X-\mu)^2 f_X(x)\,\mathrm{d}x$$

**הערה 9.15** אפשר להראות ש:

$$\mathrm{Var}(X) = \mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2$$

**דוגמה 9.16** נחשב את $\mathrm{Var}\left(X\right)$ עבור $X\sim N\left(0,1\right)$. $f_X\left(x\right)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. ראינו ש $\mathbb{E}\left(X\right)=0$. נחשב:

$$\mathrm{Var}\left(X\right)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}x^2e^{-x^2/x}\mathrm{d}x$$

אינטגרציה בחלקים מראה:

$$\int x^2e^{-x^2/x}\mathrm{d}x=-xe^{-x^2/x}+\int e^{-x^2/x}\mathrm{d}x$$

לכן:

$$\mathrm{Var}\left(X\right)=\left.-xe^{-x^2/x}\right|_{-\infty}^{\infty}+\int\limits_{-\infty}^{\infty}\frac{e^{-x^2/x}}{\sqrt{2\pi}}\mathrm{d}x=1$$

טענה שנראה בתרגול: אם $X\sim N\left(\mu,\sigma^2\right)$ אז:

$$Y=\frac{X-\mu}{\sigma}\sim N\left(0,1\right)$$

**מסקנה 9.17**

$$\mathrm{Var}\left(Y\right)=1=\frac{\mathrm{Var}\left(X\right)}{\sigma^2}\Rightarrow\sigma^2=\mathrm{Var}\left(X\right)$$

מתכונות של תוחלת.

09/01/20

בגלל שאלות של אנשים, נחזור לשאלה: מתי לפונקציה של משתנה מקרי יש צפיפות?
המשפט היסודי: $f$ רציפה למקוטעין על $[a,b]$ ואינטגרבילית במובן לא־אמיתי. נניח ש $F$ רציפה וגזירה למקוטעין על $[a,b]$. אז: $\int\limits_a^b f=F\left(b\right)-F\left(a\right)$
באינפי אנו מניחים כי $f$ רציפה ו $F$ גזירה בכל $[a,b]$. **הוכחה:** במקרה שיש נקודה שיש נקודה ש<ב<נ יחידה שבה כ לא רציפה (או לא גזירה)ץ נחשב:

$$\int\limits_a^b f=\lim_{s\to c^-}\int\limits_a^s f+\lim_{t\to c^+}\int\limits_t^b f$$

מהמשפט היסודי מאינפי:

$$=\lim_{s\to c^-}\left(F\left(s\right)-F\left(a\right)\right)+\lim_{t\to c^+}\left(F\left(b\right)-F\left(t\right)\right)$$

מרציפות $F$ ב $C$ נקבל:

$$=\left(F\left(c\right)-F\left(a\right)\right)+\left(F\left(b\right)-F\left(c\right)\right)=F\left(b\right)-F\left(a\right)$$

■

**משפט 9.18**

אם $X$ משתנה מקרי עם צפיפות $f_X$ רציפה למקוטעין. $g:\mathbb{R}\to\mathbb{R}$ רציפה חח"ע ועל. וכמו כן, גם מונוטונית עולה. $g^{-1}$ רציפה, גזירה למקוטעין. אז ל: $Y=g\left(X\right)$ יש צפיפות:

$$f_Y\left(g^{-1}\right)'\cdot f_X\circ g^{-1}$$

**הוכחה:**

**הערה 9.19** די למצוא $f_Y$ כך ש $\int_{-\infty}^{t} f_Y = P(Y \le t)$ כי אז לכל $a<b$ נקבל:

$$P(a < Y \le b) = \int_{-\infty}^{b} - \int_{-\infty}^{a} = \int_{a}^{b} f_Y$$

נבחין כי:

$$P(Y \le t) = P(g(X) \le t) = P\left(X \le g^{-1}(t)\right) = F_X\left(g^{-1}(t)\right)$$

$F_X\left(g^{-1}(t)\right)$ היא הרכבה של פונקציות רציפות. ($F_X$ רציפה לפי טענה שהזכרנו אתמול)

$$\frac{\mathrm{d}}{\mathrm{d}t} F \circ g^{-1} = \left(g^{-1}\right)' f_X \circ g^{-1}$$

מוגדר למקוטעין.
אינטגרבילית כי האינטגרל חסום בכל קטע ע"י הסתברות המתאימה והפונקציה אי שלילית.
לכן לפי המשפט היסודי המוכלל $\int_{-\infty}^{t} f_Y = F_X\left(g^{-1}(t)\right)$ כנדרש. ■

**הערה 9.20** (על המקרה הכללי)

$$\int_{a}^{b} f_Y(y) = \int_{a}^{b} \sum_{x \in g^{-1}(y)} \left|\left(g_X^{-1}\right)'\right| f_X \circ g_X^{-1}$$

ייתכנו נקודות בהם הטור לא מתכנס, ייתכן שהיכן שהוא מתכנס בוא לא רציף.

## 9.2 התפלגות משותפת של משתנים מקריים ממשיים

אם $X,Y : \Omega \to \mathbb{R}$, ההתפלגות $p_{X,Y}$ היא מידת הסתברות על $\mathbb{R}$ = התפלגות של $Z = (X,Y) : \Omega \to \mathbb{R}^2$ (אוטומטית מדיד). $p_{X,Y}$ מאופיינת ע"י ההסתברות שהיא נותנת על מלבנים.

**הגדרה 9.21** אם $X,Y$ משתנים מקריים ממשיים, פונקציית ההתפלגות $F_{X,Y}(s,t)$ היא:

$$F_{X,Y}(s,t) = P(X \le s, Y \le t) = P_{X,Y}\left((-\infty, s] \times (-\infty, t]\right)$$

אומרים של $X,Y$ יש צפיפות משותפת $f_{X,Y} : \mathbb{R}^2 \to (0,\infty]$ אם לכל מלבן $R$ במישור:

$$\iint_R f_{X,Y} = P\left((X,Y) \in \mathbb{R}\right) \iff \forall \begin{array}{c} a<b \\ c<d \end{array} \int_{a}^{b}\int_{c}^{d} f_{X,Y}(x,y)\,\mathrm{d}y\mathrm{d}x = P(a \le X \le b, c \le Y \le d)$$

**הערה 9.22** $f_{X,Y}$ מתאר תת הסתברות ליחידת שטח ⇐"צפיפות"

$$P\left((X,Y) \in [x, x+\Delta x] \times [y, y+\Delta y]\right) \approx f_{X,Y}(x,y)\,\Delta x \Delta y$$

**הערה 9.23**

$$P\left(a<X\leq b,c<y\leq d\right)=\underbrace{F\left(b,d\right)-F\left(a,d\right)}_{P(a<X\leq b,Y\leq d)}-\underbrace{\left(F\left(b,c\right)-F\left(a,c\right)\right)}_{P<x\leq b,y\leq c}$$

**מסקנה 9.24**

$F_{X,Y}$ מאפיינת הסתברות לפי $P_{X,Y}$ של מלבנים ("חצי פתוחים"⇐לכל מלבן) לכן $F_{X,Y}$ מאפיין את $P_{X,Y}$.

**הערה 9.25**

$$F_{X}\left(s\right)=P\left(X\leq s\right)=\lim_{t\to\infty}\underbrace{P\left(X\leq s,Y\leq t\right)}_{F_{X,Y}(s,t)}$$

לכן $F_{X,Y}$ קובעת את $F_X$ (וגם את $F_Y$)

**הערה 9.26** בהנחה ש $f_{X,Y}$ קיים ורציף ב $(x,y)$ אז:

$$F_{X,Y}\left(s,t\right)=\int_{-\infty}^{s}\int_{-\infty}^{t}f_{X,Y}\left(x,y\right)\mathrm{d}y\mathrm{d}x$$

נגזור קודם לפי $s$ ואח"כ לפי $t$ ב $(x,y)$:

$$\frac{\partial^{2}}{\partial y\partial x}F_{X,Y}\left(x,y\right)=f_{X,Y}\left(x,y\right)$$

לפי המשפט היסודי.

**הערה 9.27** אם יש צפיפות, אז

$$P\left(X\in I,Y\in J\right)=\int_{I}\int_{J}f_{X,Y}$$

כאשר $I,J$ קטעים. ניקח $J=\mathbb{R}$ ונקבל:

$$P\left(X\in I\right)=\int_{I}\left(\int_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)\mathrm{d}y\right)\mathrm{d}x$$

זו פונקציית צפיפות עבור $X$ (האם היא אינטגרבילית? בד"כ כן)
באופן אופן ל$Y$ יש צפיפות:

$$f_{Y}\left(t\right)=\int_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)\mathrm{d}x$$

**הערה 9.28**

| רציף | בדיד |
|---|---|
| $f_{X,Y}$ (אם קיימת) | $p_{X,Y}$ |
| $f_y\left(t\right)=\int_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)\mathrm{d}x$ | $p_Y\left(j\right)=\sum_i p_{X,Y}\left(i,j\right)$ |

**הערה 9.29**

$$\iint_{\mathbb{R}^2} f_{X,Y} = 1$$

**משפט 9.30**

אם $f : \mathbb{R}^2 \to [0,\infty]$ אינטגרבילית (עם $\iint_{\mathbb{R}^2} f = 1$ ) אז יש משתנים מקריים $X,Y$ כך ש $f$ היא הצפיפות של $X,Y$.

אפשר להכליל ל$n$־יות של משתנים מקריים. לא נוכיח.

**משפט 9.31**

אם $X,Y$ משתנים מקריים עם צפיפות משותפת $f_{X,Y}$. לכל קבוצה "סבירה" $C$ מתקיים:

$$P\left((X,Y)\in C\right) = \iint_C f_{X,Y}$$

(נכון לפי הגדרה כאשר $C$ מלבן)

**דוגמה 9.32** בהינתן משתנים מקריים $X,Y$ אשר מפולגים אחיד על כדור (דיסק) ברדיוס $R$ אם:

$$f_{X,Y}(x,y) = c\mathbb{1}_{\{(u,v)|\ u^2+v^2\le R^2\}}(x,y)$$

1. מהו $c$? $0 \le f_{X,Y}$ יש לבחור $c$ כך ש $\iint_{\mathbb{R}^2} f_{X,Y} = 1$ ולכן: $\pi R^2 c = 1 \Rightarrow c = \frac{1}{\pi R^2}$.
2. מהו $f_X$?

$$f_X(k) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} c\cdot 1\mathrm{d}y = \frac{2\sqrt{R^2-x^2}}{\pi R^2}$$

3. $P\left((X,Y)\in a\right)$ דיסק ברדיוס

$$P\left((X,Y)\in\left\{(u,v)\ |\ \ u^2+v^2\le a^2\right\}\right) = \iint\limits_{\{(u,v)|\ u^2+v^2\le a^2\}} f_{X,Y}$$

אם $a \ge R$ אז כל הדיסק מוכל וכמובן נקבל 1. ואם $a < R$ נקבל:

$$= c\pi a^2 = \frac{a^2}{R^2}$$

15/01/20

**תזכורת 9.33** צפיפות משותפת של $X,Y$, $f_{X,Y}$ מקיימת:

$$P\left(a\leq X\leq b,\ c\leq Y\leq d\right)=\int\limits_{a}^{b}\int\limits_{c}^{d}f_{X,Y}\left(x,y\right)\mathrm{d}x\mathrm{d}y$$

לכל $a\leq b$ , $c\leq d$ .
לכל קבוצה $A\subseteq\mathbb{R}^{2}$ סבירות:

$$P\left(\left(X,Y\right)\in A\right)=\iint_{A}f_{x,y}=\int\iint_{\mathbb{R}^{2}}\mathbb{1}_{A}f_{X,Y}$$

$$p_{X,Y}\longleftrightarrow f_{X,Y}$$

אם $X,Y$ דסיקרטיים.

נתבונן במקרה של משתנה מקרי המוגדר באופן הבא: $Z=g\left(X,Y\right)$ , כאשר $g:\mathbb{R}^{2}\to\mathbb{R}$ ו$X,Y$ משתנים מקריים.

**הערה 9.34** על מנת ש $Z$ יהיה משתנה מקרי, עליו להיות מדיד. כלומר $Z^{-1}\left(A\right)$ קבוצה מדידה לכל $A$ מדיד בטווח של $g$.
מספיק ש$g$ תקיים ש: $g^{-1}\left(\left(-\infty,t\right)\right)$ מדידה לכל $t$. (לכל $g$ סבירה זה נכון).

**משפט 9.35**

(לא נוכיח)
$Z=g\left(X,Y\right)$ הוא בעת תוחלת אס"ס:

$$\iint\left|g\left(x,y\right)\right|f_{X,Y}\left(x,y\right)\mathrm{d}x\mathrm{d}y<\infty$$

אז:

$$\mathbb{E}\left(Z\right)=\mathbb{E}\left(g\left(X,Y\right)\right)=\iint g\left(x,y\right)f_{X,Y}\left(x,y\right)\mathrm{d}x\mathrm{d}y$$

(נכון גם ל $Z\left(x_{1},\ldots,x_{n}\right)$)

**משפט 9.36**

אם ל$X,Y$ יש תוחלת אז:

$$\mathbb{E}\left(X+Y\right)=\mathbb{E}\left(X\right)+\mathbb{E}\left(Y\right)$$

(בפרט ל $X+Y$ יש תוחלת)

**הוכחה:** קיום: נסמן

$$g\left(x,y\right)=x+y$$

$$X+Y=g\left(X,Y\right)$$

על מנת ש $X+Y$ תהא בעלת תוחלת, צריך לבדוק:

$$\iint\left|x+y\right|f_{X,Y}\left(x,y\right)\mathrm{d}x\mathrm{d}y\overset{?}{<}\infty$$

נבחין כי:

$$\iint |x+y| f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y \leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x| f_{X,Y}(x,y) + |y| f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y =$$

$$\int_{-\infty}^{\infty} |x| \overbrace{\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y\right)}^{f_X(x)} \mathrm{d}x + \int_{-\infty}^{\infty} |y| \underbrace{\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\right)}_{f_X(y)} \mathrm{d}y = \underbrace{\int_{-\infty}^{\infty} |x| f_X(x)\,\mathrm{d}x}_{<\infty} + \underbrace{\int_{-\infty}^{\infty} |y| f_Y(y)\,\mathrm{d}y}_{<\infty}$$

כיוון ש$X,Y$ בעלי תוחלת. ו:

$$= \mathbb{E}(|X|) + \mathbb{E}(|Y|)$$

לכן $X+Y$ בעלי תוחלת. אותו חישוב ייתן:

$$\iint g(x,y) f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y = \mathbb{E}(X) + \mathbb{E}(Y)$$

כנדרש. ■

**הגדרה 9.37**

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X-\mu_X)(Y-\mu_Y))$$

כאשר:

$$\begin{aligned} \mu_X &= \mathbb{E}(X) \\ \mu_Y &= \mathbb{E}(Y) \end{aligned}$$

(בהנחה שקיים)

**מסקנה 9.38**

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$

## 9.3 אי תלות

$X,Y$ בלתי תלויים אס"ם

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

לכל מאורעות $A,B\subseteq\mathbb{R}$ .

**עובדה** מספיק שזה יתקיים עבור $A,B$ קטעים (במקרה $X,Y$ ממשיים)

במקרה הבדרד $X,Y$ בלתי תלויים אס"ם $p_{X,Y}(i,j) = p_X(i)\,p_Y(j)$.

**משפט 9.39**

אם $X,Y$ בלתי תלויים ו $(x,y)$ נקודות רציפות של $f_{X,Y}$, $x$ נקודת רציפות של $f_X$ ו $y$ של $f_Y$ אז:

$$f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$$

(אפשר להניח פחות)

**הוכחה:**

$$F_{X,Y}\left(s,t\right)=P\left(X\leq s,\ Y\leq t\right)=P\left(X\leq s\right)P\left(Y\leq t\right)=F_{X}\left(s\right)F_{Y}\left(t\right)$$

נגזור קודם לפי $s$ ואח"כ לפי $t$ נקבל:

$$f_{X,Y}\left(s,t\right)=\frac{\partial^{2}}{\partial t,\partial s}F_{X,Y}\left(s,t\right)=\frac{\partial}{\partial t}\left(F_{Y}\left(t\right)\underbrace{f_{X}\left(s\right)}_{F_{X}'}\right)=f_{X}\left(s\right)f_{Y}\left(t\right)$$

■

**הערה 9.40** אותו משפט נכון לגל $n$־יה של משתנים מקריים לתי תלויים.

$$f_{X_{1},\ldots,X_{n}}=f_{X_{1}}\cdot\ldots\cdot f_{X_{n}}$$

**דוגמה 9.41** יהיו $X,Y,Z\sim U\left[0,1\right]$ בלתי תלויים. נחשב:

$$P\left(X>YZ\right)=?$$

נתון:

$$f_{X}=f_{Y}=f_{Z}=\mathbb{1}_{\left[0,1\right]}$$

מאי תלות:

$$f_{X,Y,Z}\left(x,y,z\right)=\mathbb{1}_{\left[0,1\right]}\left(x\right)\mathbb{1}_{\left[0,1\right]}\left(y\right)\mathbb{1}_{\left[0,1\right]}\left(z\right)=\begin{cases}1 & 0\leq x,y,z\leq1\\0 & \text{אחרת}\end{cases}=\mathbb{1}_{\left[0,1\right]^{3}}\left(x,y,z\right)$$

$$P\left(X>YZ\right)=\iiint\limits_{\left\{(x,y,z):x>yz\right\}}\mathbb{1}_{\left[0,1\right]^{3}}\left(u,v,w\right)\mathrm{d}u\mathrm{d}v\mathrm{d}w=$$

$$\int_{0}^{1}\int_{0}^{1}\int_{vw}^{1}1\mathrm{d}u\mathrm{d}v\mathrm{d}w=\int_{0}^{1}\left(\int_{0}^{1}\left(1-vw\right)\mathrm{d}v\right)\mathrm{d}w=$$

$$\int_{0}^{1}\left(v-\frac{v^{2}}{2}w\right)\Big|_{0}^{1}=\int_{0}^{1}\left(1-\frac{w}{2}\right)\mathrm{d}w=\left(w-\frac{w^{2}}{4}\right)\Big|_{0}^{1}=1-\frac{1}{4}=\frac{3}{4}$$

**משפט 9.42**

יהיו $X,Y$ משתנים מקריים בלתי תלויים אז $Z=X+Y$ מקיים:

$$f_{Z}\left(z\right)=\int_{-\infty}^{\infty}f_{X}\left(z-u\right)f_{Y}\left(u\right)\mathrm{d}u=\int_{-\infty}^{\infty}f_{X}\left(u\right)f_{Y}\left(z-u\right)\mathrm{d}u$$

**הערה 9.43** האינטגרל הנ"ל נקרא קונבולוציה של $f_X, f_Y$.
במקרה הבדיד $p_{X+Y}(i)=\sum_j p_X(i-j)\,p_Y(j)$ וגם אז אמרנו שזוהי קונבולוציה.

**הוכחה:**

$$F_{\underbrace{X+Y}_{Z}}(s)=P(X+Y\le s)=\iint\limits_{\{(x,y):x+y\le s\}} f_{X,Y}(u,v)\,\mathrm{d}u\mathrm{d}v$$

אנו מניחים אי־תלות, לכן: $f_{X,Y}=f_Xf_Y$ (זה לא כל־כך נכון, כי אנחנו מדברים על נקודות רציפות, אבל זה מספיק טוב, כיוון שאנו עושים אינטגרל ואז נקודות אי הרציפות, בהנחה שהן ממידה אפס, לא משפיעות). נקבל:

$$F_{\underbrace{X+Y}_{Z}}(s)=\int_{-\infty}^{\infty}\int_{-\infty}^{s-v} f_X(u)\,f_Y(v)\,\mathrm{d}u\mathrm{d}v=\int_{-\infty}^{\infty} f_Y(c)\underbrace{\left(\int_{-\infty}^{s-v} f_X(u)\,\mathrm{d}u\right)}_{F_X(s-v)}\mathrm{d}v=\int_{-\infty}^{\infty} F_X(s-v)\,f_Y(v)\,\mathrm{d}v$$

נגזור לפי $s$ ונקבל:

$$\underbrace{f_{X+Y}(s)}_{\frac{\mathrm{d}}{\mathrm{d}s}F_{\underbrace{X+Y}_{Z}}(s)}=\int_{-\infty}^{\infty}\frac{\mathrm{d}}{\mathrm{d}s}F_x(s-v)\,f_y(v)\,\mathrm{d}v=\int_{-\infty}^{\infty} f_x(s-v)\,f_Y(v)\,\mathrm{d}v$$

כנדרש. ∎

**משפט 9.44**

(לא נוכיח)
אם $X,Y$ בעלי תוחלת ובלתי תלויים אז:

$$\mathbb{E}(XY)=\mathbb{E}(X)\,\mathbb{E}(Y)$$

**הערה 9.45** ראינו במקרה הבדיד, ונוכיח בתרגיל בית במקרה שיש צפיפות.

**דוגמה 9.46** $X,Y\sim U[0,1]$ בלתי תלויים.

$$f_{X+Y}=?$$

$$f_{X+Y}(x)=\int_{-\infty}^{\infty} f_X(z-u)\,f_Y(u)\,\mathrm{d}u$$

$$f_X=f_Y=\mathbb{1}_{[0,1]}=\begin{cases}1 & x\in[0,1]\\ 0 & \text{אחרת}\end{cases}$$

$$f_X(z-u)\,f_Y(u)=1\iff 0\le z-u\le 1\ \wedge\ 0\le u\le 1$$

אם $z<0$ או $z>2$ אין לזה פתרון.
אם $t<0$, $F_{X+Y}(t)=0 \Leftarrow f_{X+Y}=F'_{X+Y}=0$.

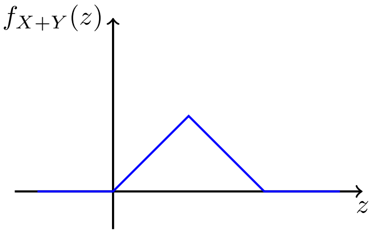
אם $0 \leq z \leq 1$ אזי:

$$\int_{-\infty}^{\infty} f_X(z-u) f_Y(u) \, \mathrm{d}u = \int_0^z \mathrm{d}u = z$$

כי $0 \leq u \leq 1$, אינטגנט מתאפס כאשר $z \geq u$ .
אם $1 \leq z \leq 2$ אז:

$$f_{X+Y}(z) = \int_{z-1}^{1} 1 \mathrm{d}u = 2 - z$$

והמסקנה היא:




**טענה 9.47**

אם $X,Y$ בלתי תלויים עם שונות אז:

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

**הוכחה:** באופן כללי, יש להוסיף לצד ימין

$$2\,\mathrm{Cov}(X,Y) = \mathbb{E}((X-\mu_X)(Y-\mu_Y)) = \mathbb{E}(X-\mu_X)\,\mathbb{E}(Y-\mu_Y) = 0 \cdot 0 = 0$$

■

**הערה 9.48** אם $X,Y$ בלתי תלויים אז לכל:

$$\begin{aligned} X' &= g(X) \\ Y' &= h(Y) \end{aligned}$$

$X',Y'$ בלתי תלויים כי:

$$P(X' \in A,\ Y' \in B) = P\left(X \in g^{-1}(A),\ Y \in h^{-1}(B)\right) = P\left(X \in g^{-1}(A)\right) P\left(Y \in h^{-1}(B)\right) = P(X' \in A)\, P(Y' \in B)$$

**טענה 9.49**

(לא נוכיח)

אם $X_1,\ldots,X_n$ ו: $Y_1,\ldots,Y_m$ בלתי תלויים. לכומר:

$$P\left(\bigcap \{X_i A_1\} \cap \bigcap \{Y_i \in B\}\right) = \prod P(x_i \in A_i) \prod P(Y_i \in B_i)$$

אזל לכל פונקציה $g$, $g\left(X_1,\dots,X_n\right),Y_1,\dots,Y_m$ בלתי תלויים גם כן.

**דוגמה 9.50** יהיו $X\sim N\left(\mu_1,\sigma_1^2\right)$, $Y\sim\left(\mu_2,\sigma_2^2\right)$ בלתי תלויים.

$$f_X=\frac{1}{\sqrt{2\pi\sigma_1^2}}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

אם $Z=X+Y$ אז:

$$Z\sim N\left(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2\right)$$

נוכיח זאת. נשים לב כי:

$$\mathbb{E}\left(Z\right)=\mathbb{E}\left(X+Y\right)=\mathbb{E}\left(X\right)+\mathbb{E}\left(Y\right)=\mu_1+\mu_2$$

וכמו כן:

$$\mathrm{Var}\left(Z\right)=\mathrm{Var}\left(X\right)+\mathrm{Var}\left(Y\right)=\sigma_1^2+\sigma_2^2$$

לכן, מספיק להראות שיש קבועים כך ש:

$$f_Z\left(u\right)=c_1e^{(u-c_2)^2c_3}$$

כי אז נובע ש $Z\sim\left(\mu_3,\sigma_3^2\right)$ וחייב להתקיים:

$$\begin{aligned}\mu_3&=\mu_1+\mu_2\\ \sigma_3^2&=\sigma_1^2+\sigma_2^2\end{aligned}$$

מספיק :

$$f_Z\left(u\right)=c_1e^{c_2u^2+c_3u+c_4}$$

$$f_z\left(u\right)=\int_{-\infty}^{\infty}f_X\left(u-v\right)f_Y\left(v\right)\mathrm{d}v=c_1\int_{-\infty}^{\infty}\exp\left(c_2\left(v-u\right)^2+c_3v^2\right)\mathrm{d}v$$

$$c_2\left(v-u\right)^2+c_3v^2=c_3\left(u-c_4v-c_5\right)^2+c_6u^2+c_7u+c_8$$

כלומר:

$$\exp\left(\dots\right)=\exp\left(u\text{ ריבועי ב }\right)\exp\left(u-cv\text{ ריבועי ב }\right)$$

ולכן האינטגרל הקודם שווה ל:

$$=c_1\exp\left(\text{u ריבועי ב }\right)\int_{-\infty}^{\infty}\exp\left(u-v\text{ ריבועי ב }\right)\mathrm{d}v$$

בתוך האינטגרל אפשר לעשות החלפת משתנה לינארית בלי לשנות את ערך האינטגרל ולהפטר מ $u$:

$$=c_1\exp\left(\dots\right)\int_{-\infty}^{\infty}\exp\left(u\text{ ריבועי ב}v\text{ במקדמים לא תלויים ב }\right)\mathrm{d}v=c_1'\exp\left(u\text{ ריבועי ב }\right)$$

כנדרש.

## 9.4 צפיפות מותנית

**הגדרה 9.51** הצפיפות של $X$ בהנתן $Y=y$ היא:

$$f_{X|Y}\left(x\mid y\right)=\frac{f_{X,Y}\left(x,y\right)}{f_{Y}\left(y\right)}$$

(במקרה הדיסקרטי: $p_{X|Y}=\frac{p_{X,Y}(i,j)}{p_{Y}(j)}$)

למה?

רוצים להגדיר:

$$P\left(X\in A\mid Y=y\right)$$

כאשר $Y=y$ הוא מאורע עם הסתברות אפס?

**מסקנה 9.52**

זו פונקציית צפיפות

**הוכחה:** נחשב בלי לחשוב:

$$\int_{-\infty}^{\infty}f_{X|Y}\left(x\mid y\right)\mathrm{d}x=\frac{1}{f_{Y}\left(y\right)}\underbrace{\int_{-\infty}^{\infty}f_{X,Y}\left(x,y\right)\mathrm{d}x}_{f_{Y}(y)}=1$$

■

אינטואיציה:

$$P\left(X\in A\mid Y=y\right)=\lim_{\delta\to0^{+}}P\left(X\in A\mid y\leq Y\leq y+\delta\right)=\lim_{\delta\to0}\frac{P\left(X\in A,y\leq Y\leq y+\delta\right)}{P\left(y\leq Y\leq y+\delta\right)}=$$

$$\lim_{\delta\to0}\frac{\int_{y}^{y+\delta}f_{X,Y}\left(u,v\right)\mathrm{d}v}{\int_{y}^{y+\delta}f_{Y}\left(v\right)\mathrm{d}v}\underbrace{=}_{\text{להופיטל, והוצאת הגבול החצוה על ידי נפנוף ידיים}}\int_{A}\frac{f_{X,Y}\left(u,y\right)}{f_{Y}\left(y\right)}\mathrm{d}u\underbrace{=}_{\text{לפי הגדרה}}\int_{A}f_{X|Y}\left(u\mid y\right)\mathrm{d}y$$

**דוגמה 9.53** נניח ש:

$$f_{X,Y}\left(u,v\right)=\begin{cases}c\cdot u^{2}v & 0\leq u,v\leq1\\0 & \text{אחרת}\end{cases}$$

מהו $P\left(0\leq X\leq\frac{1}{2}\mid Y=\frac{1}{3}\right)$?

$c$ נקבע ע"י:

$$\iint f_{X,Y}=1$$

$c=6$ אבל לא צריך.

$$P\left(0\leq X\leq\frac{1}{2}\mid Y=\frac{1}{3}\right)=\int_{0}^{\frac{1}{2}}f_{X|Y}\left(u\mid\frac{1}{3}\right)\mathrm{d}u$$

$$f_{X|Y}\left(u \mid \frac{1}{3}\right)=\frac{cu^2\cdot\frac{1}{3}}{f_Y\left(\frac{1}{3}\right)}=\frac{c\frac{u^2}{3}}{\int\limits_{-\infty}^{\infty} f_{X,Y}\left(v,\frac{1}{3}\right)\mathrm{d}v}=\frac{c\frac{u^2}{3}}{c\int\limits_0^1 \frac{v^2}{3}\mathrm{d}v}$$

נציב באינטגרל מהמשוואה הוקדמת ונקבל:

$$\frac{\int\limits_0^{\frac{1}{2}} u^2\mathrm{d}u}{\int\limits_0^1 v^2\mathrm{d}v}=\frac{1}{8}$$

16/01/20

קצת סדר למה שהיה לנו אתמול:

$$z-1\le u\le z \Leftarrow 0\le z-u\le 1$$

$0\le u\le 1 \Rightarrow \max\{0,z-1\}\le u\le \min\{1,z\}$ אם $\max\{0,z-1\}>\min\{1,z\}$ אין $u$ $\Leftarrow$ $z\ge 2$ או $z<0$

במקרים אחרים: אם $0\le z\le 1$ אזי $\max=0$ ו$\min=z$.

אם $1\le z\le 2$ אזי $\max=z-1$ ו $\min=1$ .

# 10 התכנסות של סדרות של משתנים מקריים ממשיים

נתון מרחב הסתברות $(\Omega,\mathcal{F},P)$ ומשתנים מקריים: $X_n:\Omega\to\mathbb{R}$ כאשר $n=1,2,3,\ldots$

**הגדרה 10.1 כמעט תמיד:** $X_n\to X$ כמעט תמיד (almost everywhere, almost surely)

אם יש $\Omega_0\subseteq\Omega$ (מדיד) בעל $P(\Omega_0)=1$ ולכל $\omega\in\Omega_0$ מתקיים $\lim\limits_{n\to\infty}X_n(\omega)=X(\omega)$ . כלומר:

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

**הגדרה 10.2 בהסתברות:** $X_n\to X$ בהסתברות (in probability)

אם לכל $\varepsilon>0$:

$$P\left(|X_n-X|<\varepsilon\right)=P\left(\{\omega \mid |X_n(\omega)-X(\omega)|>\varepsilon\}\right)\underset{n\to\infty}{\longrightarrow}0$$

**דוגמה 10.3** יהי $\Omega=[0,1]$ ו: מידת הסתברות אחידה $=P$.

נסמן:

$$X_n(\omega)=\omega^n$$

נרצה להראות כי $X_n\to 0$ כמעט תמיד ובהסתברות.

לכל $0\le\omega<1$ , $\omega^n\underset{n\to\infty}{\longrightarrow}0$ (סדרת הפונקציות מתכנסת נקודתית ל0 בקטע $[0,1)$)

$$P([0,1))=1$$

כי $P(\{1\})=0$, אבל נבחין כי:

$$P([0,1))=\{\omega \mid X_n(\omega)\to 0\}\Rightarrow$$

עבור $\Omega_0 = [0,1)$ ההגדרה מתקיימת. ($X=0$).
התכנסות בהסתברות:

$$\{|X_n - 0| > \varepsilon\} = \{\omega \in [0,1] \mid |\omega^n - 0| > \varepsilon\} = \{\omega \in [0,1] \mid \omega^n > \varepsilon\} = \left\{\omega \in [0,1] \mid \omega > \varepsilon^{1/n}\right\} = \left[\varepsilon^{1/n}, 1\right]$$

אבל נבחין כי:

$$P(|X_n - 0| > \varepsilon) = P\left(\left[\varepsilon^{1/n}, 1\right]\right) = 1 - \varepsilon^{1/n} \underset{n\to\infty}{\longrightarrow} 0$$

לכן $X_n \to 0$ בהסתברות.

**הערה 10.4** כאן הייתה דוגמה עם ציורים של גרפים אשר מתכנסת בהסתברות אבל לא כמעט תמיד. יותר מידי עבודה לצייר את זה.

**משפט 10.5**

התכנסות כמעט תמיד גוררת התכנסות בהסתברות

**הוכחה:** נתונה סדרה $X_n \to X$ כמעט תמיד, נקבע $\varepsilon > 0$, צריך להוכיח $P(|X_n - X| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0$.
נגדיר:

$$C_n = \{\omega \mid \text{קיים } k \ge n \text{ כך ש } |X_k(\omega) - X(\omega)| > \varepsilon\}$$

כיוון ש $\{|X_n - X| > \varepsilon\} \subseteq C$ כי הגדרת $C_n$ מתקיימת עבור $k=n$.
לכן $0 \le P(|X_n - X| > \varepsilon) \le P(C_n)$ לכן די להראות: $P(C_n) \underset{n\to\infty}{\longrightarrow} 0$. נשים לב כי: $C_{n+1} \subseteq C_n$. לכן:

$$\lim_{n\to\infty} P(C_n) = P\left(\bigcap_{n=1}^{\infty} C_n\right)$$

כיוון ש $X_n \to X$ כמעט תמיד,על מאורע בהסתברות 1:

$$X_n(\omega) \to X(\omega)$$

עברו $\omega$ כזה: $|X_n(\omega) - X(\omega)| < \varepsilon$ לכל $n$ גדול מספיק. ובפרט $\omega \notin C_n$ עבור $n$־ים גדולים. לכן: $\omega \notin \bigcap_{n=1}^{\infty} C_n$.
כלומר, הראינו שיש $\Omega_0$ עם $P(\Omega_0) = 1$ (מהגדרת התכנסות כמעט תמיד) כך ש:

$$\bigcap_{n=1}^{\infty} C_n \subseteq \Omega \backslash \Omega_0$$

(כי $\omega \in \Omega_0 \Leftarrow \omega \notin \bigcap_{n=1}^{\infty} C_n$).
לכן $P\left(\bigcap_{n=1}^{\infty} C_n\right) = 0$, מש"ל. ■

**הערה 10.6** בהנתן $\varepsilon > 0$ אם נסמן את $B_k = \{|X_n - X| > \varepsilon\}$, אז $C_n = \bigcup_{k\ge n} B_k$. וראינו בעבר כי התכנסות כמעט תמיד $\Leftarrow$ $P\left(\underbrace{\bigcap_{n=1}^{\infty} C_n}_{\liminf B_k}\right) = 0$. וראינו כי אם $\liminf B_k = 0$ אז: $P(B_k) \to 0$.

**דוגמה 10.7** יהיו $X_n$ משתנים מקריים בלתי תלויים, עם ערכים $0,1$ כך ש:

$$P(X_n=1) = \frac{1}{n}$$
$$P(X_n=0) = 1-\frac{1}{n}$$

$X_n\to 0$ בהסתברות כי אם $\varepsilon>0$ אז $\{|X_n-0|>\varepsilon\}=\{X_n=1\}$, אבל $P(X_n=1)=\frac{1}{n}\underset{n\to\infty}{\longrightarrow}0$
נראה כי בהסתברות 1, $X_n=1$ אינסוף פעמים ולכן: $X_n\nrightarrow 0$ כמעט תמיד.
לבי בורל קנטלי שני, די להראות: $\sum\limits_{n=1}^{\infty}P(X_n=1)=\infty$ וזה נכון. מש"ל.

22/01/20

## 10.1 חוק המספרים הגדולים

**משפט 10.8**

יהיו $X_n$ ו-$X$ משתנים מקריי סע ל-$(\Omega,\mathcal{F},P)$, אז הפונקציות הבאות מדידות:
1. $Y=\sup\limits_n X_n$
2. $Z=\inf\limits_n X_n$
3. $W=\limsup\limits_{n\to\infty}X_n$
4. $V=\liminf\limits_{n\to\infty}X_n$

ו:

$$\Omega\supset\left\{\omega:\lim_{n\to\infty}X_n(\omega)\ \text{קיים}\right\}\in\mathcal{F}^2$$

הוא מאורע.

**תזכורת 10.9** אם $X\geq 0$ ובעל תוחלת. $t>0$ אז:

$$P(X\geq t)\leq\frac{\mathbb{E}(x)}{t}$$

(זהו אי שיוויון מרקוב, הוכחנו למשתנים בדידים)
אם $Y$ בעל שונות $\mu=\mathbb{E}(Y)$ לכל $t>0$:

$$P(|Y-\mu|>t)\leq\frac{\mathrm{Var}(Y)}{t^2}$$

(אי שיוויון צ'בישב, גם כאן הוכחנו למשתנים בדידים)

**הוכחה:** למשתנים עם צפיפות:
אי שיוויון מרקוב:

$$\mathbb{E}(X)=\int_{-\infty}^{\infty}xf_X(x)\,\mathrm{d}x\underbrace{=}_{x\geq0\Rightarrow x<0,\ f_X(x)=0}\int_0^{\infty}\underbrace{xf_X(x)}_{\geq0}\mathrm{d}x\geq\int_t^{\infty}xf_X(x)\,\mathrm{d}x\geq t\int_t^{\infty}f_X(x)\,\mathrm{d}x=t\cdot\underbrace{P(t\leq x<\infty)}_{P(X\geq t)}$$

נחלק ב-$t$ ונקבל את אי שיוויון מרקוב כנדרש.
נראה את אי שיוויון צ'בישב:

$$P(|Y-\mu|\geq t)=P\left((Y-\mu)^2\geq t^2\right)\leq\frac{\mathbb{E}\left((Y-\mu)^2\right)}{t^2}=\frac{\mathrm{Var}(Y)}{t^2}$$

האי שיוויון לקראת הסו הוא מאי שיוויון מרקוב. ■

נאחד טיפול במשתמש מקרי בעזרת הנחה שיש פעולת תוחלת שמקיימת:

1. לינאריות: $\mathbb{E}(aX+bY)=a\mathbb{E}(X)+b\mathbb{E}(Y)$
2. מונוטוניות $X\geq Y$ בכל נקודה אז: $\mathbb{E}(X)\geq\mathbb{E}(Y)$
3. נירמול: $\mathbb{E}(\mathbb{1}_A)=P(A)$

**הגדרה 10.10** סדרת משתנים מקריים $X_1,X_2,\ldots$ היא סדרה בלתי תלוייה שוות הסתברות ולכולם אותה התפלגות:

$$\forall i,j\quad F_{X_i}=F_{x_j}$$

(independent identically distributed (iid))

**הערה 10.11** $X$ בעל תוחלת $\Longleftrightarrow$ לכולם יש תוחלת
אז אותו תוחלת. כנ"ל לגבי שונות

**משפט 10.12 חוק המספרים הגדולים החלש**

אם $X_1,X_2,\ldots$ משתנים מקריים בלתי תלויים שווי הסתברות, קיים $\mathrm{Var}(X_1)$ונסמן $\mu=\mathbb{E}(X_1)$ וגם:

$$S_n=\frac{1}{n}(X_1+X_2+\ldots+X_n)$$

אז: $S_n\to\mu$ בהסתברות.

**הוכחה:** צריך להוכיח, בהנתן $\varepsilon>0$:

$$P(|S_n-\mu|\geq\varepsilon)\underset{n\to\infty}{\longrightarrow}0$$

נשים לב:

$$\mathbb{E}(S_n)=\mathbb{E}\left(\frac{1}{n}X_1+\frac{1}{n}X_2+\ldots+\frac{1}{n}X_n\right)=\frac{1}{2}\left(\underbrace{\mathbb{E}(X_1)}_{\mu}+\ldots+\underbrace{\mathbb{E}(X_n)}_{\mu}\right)=\frac{1}{n}\cdot n\cdot\mu=\mu$$

לכן, לפי אי שיוויון צ'בישב:

$$P(|S_n-\mu|\geq\varepsilon)\leq\frac{\mathrm{Var}(S_n)}{\varepsilon^2}$$

צריך להוכיח כי זה שואף לאפס.
נחשב:

$$\mathrm{Var}(S_n)=\mathrm{Var}\left(\frac{1}{n}\left(\sum_{i=1}^{n}X_i\right)\right)=\frac{1}{n^2}\mathrm{Var}\underbrace{(X_1+\ldots+X_n)}_{\text{משפחה בלתי תלוייה}}=\frac{1}{n^2}\left(\mathrm{Var}(X_1)+\mathrm{Var}\underbrace{(X_2+\ldots+X_n)}_{\text{משפחה בלתי תלוייה}}\right)$$

השיווין הוא כי שונות של סכום של משתנים מקרים בלתי תלויים הוא סכום השונויות.

$$=\frac{1}{n^2}(\mathrm{Var}(X_1)+\ldots+\mathrm{Var}(X_n))\underbrace{=}_{\forall k\ \ \mathrm{Var}\,X_k=\mathrm{Var}\,X_1}\frac{1}{n^2}(n\,\mathrm{Var}\,X_1)$$

נציב ונקבל:

$$P(|S_n-\mu|\geq\varepsilon)\leq\frac{\mathrm{Var}(X_n)}{\varepsilon^2}=\underbrace{\frac{\mathrm{Var}(X_1)}{\varepsilon^2}}_{\text{קבוע}}\cdot\frac{1}{n}\underset{n\to\infty}{\longrightarrow}0$$

■

**מסקנה 10.13**

אם $X_n$ משתנים $\mathrm{Ber}\left(p\right)$ בלתי תלויים אז:

$$\begin{aligned}\mathbb{E}\left(X_1\right) &= p\\ S_n &= \frac{1}{n}\sum_{i=1}^{n}X_i=\frac{1}{n}\#\left\{1\le i\le n \mid X_i=1\right\}\end{aligned}$$

שזה שווה לממוצע מספר ההצלחות.
לכן מחוק המספרים הגדולים:

(הסתברות של הצלחה בניסוי בודד) $\xrightarrow{\text{בהסתברות}}$ (מספר הצלחות ממוצע)

נסמן $A_n=X_1+\ldots+X_n\sim\mathrm{Bin}\left(n,p\right)$. ארינו כי הגרף של $pA_n\left(i\right)$ עולה מ 0 עד $np$ ויורד חזרה עד $n$. ראינו כי $\mathrm{Var}\,A_n=n\,\mathrm{Var}\left(X_1\right)$ וכלכן סטיית התקן מסדר גודל של $\sqrt{n}$.

**הגדרה 10.14** המומנט ה$k$ של משתנה מקרי $Y$ הוא $\mathbb{E}\left(Y^k\right)$ אם קיים. מסומן $\mu_k$. אומרים של־$Y$ יש $n$ מומנטים אם קיים מומנט $k$ לכל $0\le k\le n$ .

**הערה 10.15**

1. $\mu_1=\mathbb{E}\left(Y\right)$
2. אם יש מומנט $k$ אז יש מומנט $k'$ לכל $k'\le k$ (בתרגיל ראינו עבור $k=2$, $k'=1$)
3. אם $Y$ חסום (כלומר, יש $M\in\mathbb{R}$ כך ש $\left|Y\left(\omega\right)\right|\le M$ לכל $\omega\in\Omega$) אז כל המומנטים קיימים (תרגיל)
4. לכל $n$ יש משתנים מקריים עם $n$ מומנטים אבל לא $n+1$ מומנטים. (תרגיל)

**משפט 10.16 חוק המספרים הגדולים החזק**

אם $X_1,X_2,\ldots$ משתנים מקריים בלתי תלויים שווי הסתברות ואם יש להם 4 מומנטים אז: $S_n\to\mu$ כמעט תמיד ($\mu=\mathbb{E}\left(X_1\right)$)

**למה 10.17**

יהיו $Y_n,Y$ משתנים מקריים. עבור $\varepsilon>0$ נסמן:

$$A_\varepsilon=\left\{\omega:\ \left|Y_n\left(\omega\right)-Y\left(\omega\right)\right|>\varepsilon \text{ לאינסוף } n\text{-ים}\right\}$$

אז $Y_n\to Y$ כמעט תמיד אס"ם לכל $\varepsilon>0$ מתקיים $P\left(A_\varepsilon\right)=0$.

**הוכחה:** נניח $Y_n\to Y$ כמעט תמיד. יהי $\Omega_0\subseteq\Omega$ מאורע, $P\left(\Omega_0\right)=1$ ולכל $\omega\in\Omega_0$:

$$Y_n\left(\omega\right)\to Y\left(\omega\right)$$

בהנתן $\varepsilon>0\Leftarrow$ $\omega\in\Omega_0$ אז $\omega\notin A_\varepsilon\Leftarrow$ $A_\varepsilon\subseteq\Omega\backslash\Omega_0\Leftarrow$ $0\Leftarrow P\left(A_\varepsilon\right)=1-P\left(\Omega_0\right)=0$.

**<u>כיוון שני:</u>**

נניח $P\left(A_\varepsilon\right)=0\quad\forall\varepsilon>0$ נגדיר:

$$A=\bigcup_{k=1}^{\infty}A_{1/k}$$

**הערה 10.18** $A_\varepsilon$ הוא מאורע כי הוא $\limsup$ של $\left\{\left|Y_n-Y\right|>\varepsilon\right\}$, וזה מאורע כי $Y,Y_n$ מדידים. ו $\left|Y_n-Y\right|$ היא פונקציה רציפה של $Y_n,Y$. ולכן $A$ גם מאורע.

נגדיר $\Omega_0 = \Omega \backslash A$ . נבחין כי $P(A) \leq \sum_{k=1}^{\infty} P\left(A_{1/k}\right) = \sum 0 = 0$ ולכן: $P(\Omega_0) = 1 - 0 = 1$.

נרצה להראות כי: $A \subseteq \{\omega \mid Y_n(\omega) \nrightarrow Y(\omega)\}$. אם $Y_n(\omega) \nrightarrow Y(\omega)$ אז יש $\varepsilon > 0$ כך ש $|Y_n(\omega) - Y(\omega)| > \varepsilon$ אינסוף פעמים. אפשר להניח ש $\varepsilon = 1/k$ (נקח $k$ כך ש $\frac{1}{k}$ קטן מה$\varepsilon$ המקורי).

לכן $\omega \in A_{1/k} \subseteq A$ (כי ראינו שלכל $\omega \in \Omega_0$ $Y_n(\omega) \to Y(\omega)$) ■

נוכיח את משפט חוק המספרים הגדולים החזק: **הוכחה:** לפי הלמה יש להראות לכל $\varepsilon > 0$ ש:

$$P\left(\{|S_n - \mu| > \varepsilon \quad \text{עבור אינסוף } n\text{-ים}\}\right) = 0$$

נוח להניח $\mu = 0$. וזה מותר כי אם נגדיר: $X_n' = X_n - \mu$ אז $X_n'$ בלתי תלויים ושווי הסתברות. ומתקיים: $\mathbb{E}(X_n') = 0$.

אם

$$\underbrace{S_n' = \frac{1}{n}\left(X_1' + \ldots + X_n'\right)}_{(S_n - \mu)} \xrightarrow[\text{כמעט תמיד}]{}$$

אז $S_n - \mu \to 0$ כמעט תמיד $\iff$ $S_n \to \mu$ כמעט תמיד.

לכן נניח $\mu = 0$ . צריך להוכיח לכל $\varepsilon > 0$:

$$P\left(\{|S_n| > \varepsilon \quad \text{עבור אינסוף } n\text{-ים}\}\right) = 0$$

לפי בורל קנטלי די להראות:

$$\sum_{n=1}^{\infty} P\left(|S_n| \geq \varepsilon\right) < \infty$$

**למה 10.19**

$$\mathbb{E}\left(S_n^4\right) \leq \frac{c}{n^2}$$

כאשר $c$ קבוע התלוי במומנטים של $X_1$, אבל לא ב$n$.

**הוכחה:**

$$\mathbb{E}\left(S_n^4\right) = \mathbb{E}\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^n\right) = \frac{1}{n^4}\left(\mathbb{E}\left(\sum_{i,j,kl=1}^{n} X_i X_j X_k X_l\right)\right) = \frac{1}{n^4}\sum_{i,j,k,l=1}^{n} \mathbb{E}\left(X_i X_j X_k X_l\right)$$

יש כמה מקרים של רביעיות $i, j, k, l$.

1. כולם שווים אז:

$$\mathbb{E}\left(X_i X_j X_k X_k\right) = \mathbb{E}\left(X_i^4\right) = \mu_4$$

2. יש שני זוגום אינדקסים שווים, אבל לא כולם שווים:

$$\mathbb{E}\underbrace{\left(X_1 X_2 X_1 X_2\right)}_{X_1^2 X_2^2} \underbrace{=}_{\text{מאי תלות}} \mathbb{E}\left(X_1^2\right)\mathbb{E}\left(X_2^2\right) = \left(\mu_2\right)^2$$

3. יש אינקס אחד ששונה מכל האחרים (האחרים יכולים להיות זהים, למשל: $\mathbb{E}\left(\boxed{X_1}X_2X_2X_3\right)\underbrace{=}_{\text{אי תלות}}\overbrace{\mathbb{E}\left(X_1\right)}^{=0}\mathbb{E}\left(\ldots\right)=$
0

ולכן:

$$\mathbb{E}\left(S_n^4\right)=\frac{1}{n^4}\left(\sum_{\substack{i,j,k,l\\ \text{מסוג 1}}}\mu_4+\sum_{\substack{i,j,k,l\\ \text{מסוג 2}}}\left(\mu_2\right)^2+\sum_{\substack{i,j,k,l\\ \text{מסוג 3}}}0\right)$$

נבחין כי יש $n$ רבעיות מסוג 1. $\binom{4}{2}\binom{n}{2}\le cn^2$ מסוג 2 (בחירה של 2 אינקסים מתוך $n$, כפול כמות הסידורים שלהם.

$$=\frac{1}{n^4}\left(\mu_4n+c\left(\mu_2\right)^2n^2\right)\le\frac{c'}{n^2}$$

כנדרש. ■

מהלמה ומאי שיוויון מרקוב נקבל כי מתקיים:

$$P\left(|S_n|\ge\varepsilon\right)=P\left(S_n^3\ge\varepsilon^4\right)\underbrace{\le}_{\text{אי שיוויון מרקוב}}\frac{\mathbb{E}\left(S_n^4\right)}{\varepsilon^4}\underbrace{\le}_{\text{הלמה}}\frac{c}{\varepsilon^4}\cdot\frac{1}{n^2}$$

ולכן:

$$\sum_{n=1}^{\infty}P\left(S^4\ge\varepsilon\right)\le\sum_{n=1}^{\infty}\underbrace{\frac{c}{\varepsilon^4}}_{\text{const}}\frac{1}{n^2}<\infty$$

כנדרש. ■

### 10.1.1 תאור של הסתברות אחידה על $[0,1]$

יהיו $X_1,X_2,X_3\ldots$ משתנים מקריים בלתי תלויים אחידים על $\{0,\ldots,9\}$, עבור $i=0,1,\ldots,9$ מתקיים: $P\left(X_n=i\right)=\frac{1}{10}$. נגדיר:

$$X=\sum_{n=1}^{\infty}\frac{X_n}{10^n}=0.X_1X_2X_3\ldots$$

(כתב עשרוני)

זהו משתנה מקרי (כלומר הוא מדיד) מכיוון ש:

$$=\lim_{N\to\infty}\sum_{n=1}^{N}\frac{X_n}{10^n}$$

**טענה 10.20**

$X\sim U\left[0,1\right]$

**הוכחה:** ראשית:

$$\forall x\in[0,1]\quad P\left(X=x\right)=0$$

כי $\{X_n = x_n\} = \bigcap_{n=1}^{\infty} \{X = x\}$ כאשר $x = 0.x_1x_2x_3\ldots$

$$P(X=x) = \lim_{N\to\infty} P\left(\bigcap_{n=1}^{N} \{X_n = x_n\}\right) = \lim_{N\to\infty} \prod_{n=1}^{N} \underbrace{P(X_n = x_n)}_{1/10} = \lim_{N\to\infty} = \frac{1}{10^N} = 0$$

נבחין כי יש מספרים אשר יש להם שני יצוגים (לדוגמה $0.1 = 0.099999999999\ldots$), במקרים אלה נחזור על החישוב לכל אחד מהם.

**למה 10.21**

אם $0 \le k \le 10^n - 1$ :

$$P\left(\frac{k}{10^n} \le X \le \frac{k+1}{10^n}\right) = \frac{1}{10^n}$$

**הוכחה:** עד כדי נקודות קצה:

$$\{X_1 \ldots X_n = k_1k_2\ldots k_n\}$$

פתוח הערוני של$k$ כשלם.

**דוגמה 10.22**

$$\frac{315}{1000} \le X \le \frac{316}{1000} = \{X_1 = 3, X_2 = 1, X_3 = 5\} \cup \underbrace{\{\text{המאורע ש } X \text{=הקצוות}\}}_{P=0}$$

$$P(X_1 = k_1, \ldots, X_n = k_n) \underbrace{=}_{\text{אי תלות}} P(X_1 = k_1)\cdot\ldots\cdot P(X_n = k_n) = 10^{-n}$$

■

**מסקנה 10.23**

$P(X \in I) = P(U \in I)$ עבור $U \sim U[0,1]$ ו $I$ מהצורה כנ"ל (כלומר של הלמה).

**עובדה** כל קטע $[0,t)$ הוא איחוד בן מנייה של קטעים $I_1, I_2, \ldots$ כנ"ל.
נובע לכל $t$:

$$P(X \in [0,t)) = P\left(X \in \bigcup_{i=1}^{n} I_i\right) = \sum P(X_i \in I_i) = \sum (U \in I_i) = P(U < t)$$

כאשר האיחוד הוא זר, של קטעים כנ"ל.
לכן הם שווי התפלגות

■

**טענה 10.24**

יהי $U \sim U[0,1]$, $U = 0.U_1U_2\ldots$ אזי לכל $K \in \{0,\ldots,9\}$:

$$\lim_{n\to\infty} \frac{1}{n} \# \{1 \le i \le n \mid U_i = k\} = \frac{1}{10}$$

(מספר הפעמים שהספרה $k$ מופיעה ב $n$ ספרות ראשונות)

**משפט 10.25**

$X_n$ משתנים מקריים,

$$Y = \sup_{n\in\mathbb{N}} X_n$$

אז $Y$ מדיד.

**הוכחה:** צריך להראות כי (קטע) $Y^{-1}$ הוא מאורע.
נשתמש בטענה שלא נוכיח (תרגיל):

**טענה 10.26**

$Y^{-1}\left((-\infty,t)\right)$ הוא מאורע לכל $t$.

$Y^{-1}\left((-\infty,t)\right)$ הוא המאורע: $\{Y<t\} = \left\{\sup\limits_n X_n < t\right\}$ אבל נבחין כי: $\sup X_n < t \iff \exists k \quad \forall n \quad X_n <$ $t-\frac{1}{k}$.

$$\{\sup X_n < t\} = \bigcup_{k=1}^{\infty}\left\{\forall n \quad X_n < t-\frac{1}{k}\right\} = \bigcup_{k=1}^{\infty}\left(\bigcap_n \underbrace{\left\{X_n < t-\frac{1}{k}\right\}}_{\text{מאורע כי } X_n \text{ohshsn}}\right)$$

זהו חיתוך בן מנייה של איחוד בן מנייה של מאורעות $\Leftarrow$ מאורע. ■

**טענה 10.27**

$Z = \limsup X_n$ מדיד.

**הוכחה:** נגדיר $Y_n = \sup\limits_{k\geq n} X_k$, ראינו ש $Y_n$ מדידים.

$$\limsup X_n = \inf_n \overbrace{\sup_{k\geq n} X_k}^{Y_n}$$

אותה הוכחה כמו קודם מראה ש $\inf\limits_n Y_n$ מדיד לכן $Z$ מדיד.
באותו אופן מראים:

$$Z' = \liminf X_n$$

מדיד אז:

$$\left\{\lim_{n\to\infty} X_n \text{ קיים}\right\} = \{Z = Z'\}$$

■

שאלה פתוחה: האם עבור $\pi = 3.x_1x_2x_3\ldots$ מתקיים $\frac{1}{n}\#\{1\leq i\leq n \mid x_i = 5\} \to \frac{1}{10}$? 23/01/20
ארדש אמר שנדרש עוד 200 שנה של מתמטיקה כדי לענות על זה.

**משפט 10.28**

בחוק החזק (חלש) מספיק להיות מומנט 1 (שקיים תוחלת).

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mathbb{E}\left(X_1\right)$$

**הערה 10.29** חלש בהסתברות (הוכחנה בהנחת קיום שונות), חזק: כמעט תמיד (הנחנו מומנט 4) קיימת הוכחה בסיכומים של רז.

## 10.2 התכנסות בהתפלגות

**הגדרה 10.30 התכנסות בהתפלגות:** סדרת משתנים מקריים ממשיים $X_1, X_2, \ldots$ מתכנסים בהתפלגות למשתנה מקרי $X$ אם לכל $t$ כך ש $P(X=t)=0$:

$$P(X_n \le t) \underset{n\to\infty}{\longrightarrow} P(X \le t)$$

באופן שקול, לכל נקודת רציפות $t$ של $F_X$ מתקיים:

$$F_{X_n}(t) \underset{n\to\infty}{\longrightarrow} F_X(t)$$

**הערה 10.31** המשתנים אינם צריכים להיות על מרחב משותף.

**הערה 10.32** ההגדרה ברשימות של רז שגויה!

**דוגמה 10.33** $X_n \sim \text{Ber}(p_n)$ כאשר $p_n \to \frac{1}{2}$.
$t=0,1$ נקודות אי רציפות של $X_n$ וגם של $X$. אם $t<0$ אז: $P(X_n \le t) = 0 = P(X \le t)$. אם $t>1$ אז: $P(X_n \le t) = 1 = P(x \le t)$.
אם $0<t<1$ אז:

$$P(x_n \le t) = P(X_n = 0) = 1 - p_n \underset{n\to\infty}{\longrightarrow} 1 - \frac{1}{2} = P(X=0) = P(X \le t)$$

לכן, בכל מקרה, אם $t \ne 0,1$ : $P(X_n \le t) \to P(X \le t)$. לכן בהתפלגות: $X_n \to X$.

**דוגמה 10.34** יהיו $X_n \sim \text{Ber}\left(\frac{1}{2}\right)$ בלתי תלויים. לכולם אותה התפלגות, לכן:

$$X_n \to X \sim \text{Ber}\left(\frac{1}{2}\right)$$

הם בלתי תלויים לכן:

$$P(X_n = X_{n+1}) = \frac{1}{2}$$

תרגיל: לפרט מדוע לא: $X_n \to X$ כמעט תמיד.
מצד שני, יהי $X \to \text{Ber}\left(\frac{1}{2}\right)$, $X_n \equiv X$. גם אז: $X_n \to X$ בהסתברות, כמעט תמיד, בהתפלגות.

**דוגמה 10.35** יהיו:

$$X_n = \begin{cases} 0 & \frac{1}{2} \text{ בהסתברות} \\ 1+\frac{1}{n} & \frac{1}{2} \text{ בהסתברות} \end{cases}$$

$$X \sim \text{Ber}\left(\frac{1}{2}\right) = \begin{cases} 0 & \frac{1}{2} \text{ בהסתברות} \\ 1 & \end{cases}$$

כמו בדוגמה קודמת. אם $0 \neq t < 1$ אז:

$$P\left(X_n \leq t\right) = P\left(X \leq t\right) = \begin{cases} 0 & t < 0 \\ \frac{1}{2} & 0 < t \leq 1 \end{cases}$$

אם $t = 1$:

$$P\left(X_n \leq 1\right) = \frac{1}{2} \neq 1 = P\left(X \leq 1\right)$$

זו לא בעיה! אבל זה מסביר מדוע ההגדרה לא דורשת התכנסות בנקודת אי רציפות.

$t > 1$ אז: $P\left(X_n \leq t\right) = 1$ ל $n$ גדול מספיק⇐$P\left(X_n \leq t\right) \to P\left(X < t\right)$.

לסיכום: $X_n \to X$ בהתפלגות אבל: $F_{X_n}\left(1\right) \nrightarrow F_X\left(1\right)$

**דוגמה 10.36** עבור $\alpha>0$ יהיו: $X_n \sim \mathrm{Bin}\left(n, \frac{\alpha}{n}\right)$, $X \sim \mathrm{Pois}\left(\alpha\right)$. ראינו שלכל $k$: $P\left(X_n = k\right) \to P\left(X = k\right)$.

נראה: $X_n \to X$ בהתפלגות.

לכל $t < 0$:

$$P\left(X_n \leq t\right) = 0 = P\left(X \leq t\right)$$

לכל $0 < t$: $t \notin \mathbb{N}$ (נקודות אי רציפות של $F_X$), יש $k \in \mathbb{N}$ כך ש: $k < t < k+1$:

$$P\left(X_n < t\right) = \sum_{i=0}^{k} P\left(X_n = i\right) \underset{n\to\infty}{\longrightarrow} \sum_{i=0}^{k} P\left(X = i\right) = P\left(X < t\right)$$

**למה 10.37**

אם $X_n \to X$ בהתפלגות אז $P\left(X < t\right) \underset{n\to\infty}{\longrightarrow} P\left(X_n < t\right)$ לכל נקודת רציפות $t$ של $F_X$.

**הוכחה:** יהי $\varepsilon > 0$:

$$\left\{X_n \leq t - \varepsilon\right\} \leq \left\{X_n < t\right\} \leq \left\{X_n \leq t\right\}$$

$$P\left(\ldots\right) \leq P\left(\ldots\right) \leq P\left(\ldots\right)$$

ומכאן נובע:

$$\liminf P\left(X_n < t\right) \leq \limsup P\left(X_n < t\right) \leq \overset{\lim}{\cancel{\limsup}} P\left(X_n \leq t\right) = P\left(X \leq t\right)$$

ומאותם שיקולים:

$$\liminf P\left(X_n < t\right) \geq \overset{\lim}{\cancel{\liminf}} P\left(X_n \leq t\right) = P\left(X \leq t - \varepsilon\right)$$

אם $t - \varepsilon$ נקודת רציפות של $F_X$. ממשפט באינפי שאומר: אם $G$ פונקציה מונוטונים $\mathbb{R} \to \mathbb{R}$ אז נקודת הרציפות צפופות (קבוצת נקודות אי רציפות היא בת־מנייה), אפשר לבחור $\varepsilon$ קרוב כרצוננו ל0 כך ש $t - \varepsilon$ נקודת רציפות.

נניח $t$, $t - \varepsilon$ נקודות רציפות של $F_X$. קיבלנו:

$$\begin{matrix} P\left(X \leq t - \varepsilon\right) \\ \downarrow \\ P\left(X \leq t\right) = P\left(X < t\right) \end{matrix} \qquad \leq \underbrace{\lim \begin{matrix}\inf \\ \sup\end{matrix} P\left(X_n < t\right)}_{\text{לא תלוי ב}\varepsilon} \leq P\left(X \leq t\right)$$

כי $t$ נקודת רציפות של $F_X$. ■

**מסקנה 10.38**

$X_n \to X$ בהתפלגות אז לכל $a<b$ נקודת רציפות של $F_X$:

$$\underbrace{P\left(a\leq X_n\leq b\right)}_{P(X_n\leq b)-P(X_n<a)}\to P\left(a\leq x\leq b\right)$$

29/01/20

**משפט 10.39**

אם $X_n\to X$ בהסתברות אז גם בהתפלגות

ראשית נוכיח את הלמה:

**למה 10.40**

אם $A_n, B_n$ סדרות מאורעות, $P\left(B_n\right)\underset{n\to\infty}{\longrightarrow}1$ אז $P\left(A_n\cap B_n\right)$, $P\left(A_n\right)$ אותה אסימפטוטיקה:

$$\left|P\left(A_n\right)-P\left(A_n\cap B_n\right)\right|\underset{n\to\infty}{\longrightarrow}0$$

**הוכחה:**

$$0\leq P\left(A_n\right)-\underbrace{P\left(A_n\cap B_n\right)}_{A_n\text{ מוכל ב}}=P\left(A_n\backslash\left(A_n\cap B_n\right)\right)=P\left(A_n\backslash B_n\right)\leq P\left(B_n^c\right)=1-P\left(B_n\right)\to 0$$

ולכן הלמה נכונה. ■

כעת נוכיח את המשפט. **הוכחה:** נקבע $t$, נקבע $\varepsilon>0$ (בסוף נשאיף ל0)

לפי הלמה מספיק להוכיח:

$$\underbrace{P\left(X_n\leq t\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right)}_{\alpha_{n,\varepsilon}}\underset{n\to\infty}{\longrightarrow}P\left(X\leq t\right)$$

מכיוון שאנו יודעים כי: $P\left(\left|X_n-X\right|<\varepsilon\right)\underset{n\to\infty}{\longrightarrow}1$ (מהנתון).

נבחין כי:

$$\left\{X\leq t-\varepsilon\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right\}\subseteq\left\{X_n\leq t\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right\}\subseteq\left\{X\leq t+\varepsilon\quad\wedge\quad\left|X_n-X<\varepsilon\right|\right\}$$

ולכן:

$$\liminf_{n\to\infty}P\left(\left\{X\leq t-\varepsilon\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right\}\right)\leq\liminf_{n\to\infty}P\left(\left\{X_n\leq t\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right\}\right)\leq$$
$$\limsup_{n\to\infty}P\left(\left\{X_n\leq t\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right\}\right)\leq\limsup_{n\to\infty}P\left(\left\{X\leq t+\varepsilon\quad\wedge\quad\left|X_n-X<\varepsilon\right|\right\}\right)$$

מכיוון ש $P\left(\left|X_n-X\right|<\varepsilon\right)\underset{n\to\infty}{\longrightarrow}1$, מתקיים:

$$\liminf_{n\to\infty}P\left(X\leq t-\varepsilon\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right)=\liminf_{n\to\infty}P\left(X\leq t-\varepsilon\right)$$

(מהלמה)

כנ"ל גם נקבל:

$$\limsup_{n\to\infty}P\left(X\leq t+\varepsilon\quad\wedge\quad\left|X_n-X\right|<\varepsilon\right)=\limsup_{n\to\infty}P\left(X\leq t+\varepsilon\right)$$

נבחין כי ההתסברויות הנ"ל כבא אינן תלויות ב $n$.
יצא שהסדרה המבוקשת. נסמן: $\beta_n = P(X_n \le t)$ נבחין כי:

$$\liminf \alpha_{n,\varepsilon} = \liminf \beta_n$$

**הערה 10.41**

$$\alpha_{n,\varepsilon} = P(X_n \le t,\ |X_n - X| < \varepsilon)$$

וכנ"ל לגבי $\limsup$. ולכן:

$$P(X \le t - \varepsilon) \le \liminf_{n\to\infty} \beta_n \le \limsup_{n\to\infty} \beta_n \le P(X \le t + \varepsilon)$$

אבל הגורף באמצע כלל לא תלוי ב$\varepsilon$. נשאיף את $\varepsilon$ לאפס. בהנחה ש $t$ נקדות רציפות של $F_X$ נקבל:

$$\lim_{\varepsilon\to 0^+} P(X \le t + \varepsilon) = P(X \le t) = \lim_{\varepsilon\to 0^+} P(X \le t - \varepsilon)$$

כלומר קבלנו:

$$P(X \le t) \le \liminf \beta_n \le \limsup \beta_n \le P(X \le t)$$

אבל הכל שוויונות. ולכן כנדרש. ■

וסיכום: התכנסות כמעט תמיד $\Leftarrow$ התכנסות בהסתברות $\Leftarrow$ התכנסות בהתפלגות אבל לא בהכרח להפך.

**דוגמה 10.42**

$$\Omega = \{\omega_1, \omega_2\}$$

$P$ אחיד

| | $\omega_1$ | $\omega_2$ |
|---|---|---|
| $X_1$ | 0 | 1 |
| $X_2$ | 1 | 0 |

התפלגות זהה (כל אחד מקבל $0$ בהסתברות $\frac{1}{2}$, $1$ בהתסתברות $\frac{1}{2}$)
הסדרה

$$X_1, X_2, X_1, X_2, \dots$$

לא מתכנסת בהסתברות. אם $X$ משתנה מקרי מקבל ערך $x$ בהסתברות חיובית או ש:

$$\left\{|X_1 - X| > \frac{1}{2}\right\} \supseteq \{X = x\}$$

או:

$$\left\{|X_2 - X| > \frac{1}{2}\right\} \supseteq \underbrace{\{X = x\}}_{\text{בהסתברות } 0}$$

לכן:

$$P\left(|X_i - X| > \frac{1}{2}\right) \underset{n\to\infty}{\not\to} 0$$

$$P\left(|X_i - X| > \frac{1}{2}\right) \ge P(X = x)$$

**הגדרה 10.43**

$$C_c(\mathbb{R})=\{f:\mathbb{R}\to\mathbb{R}\mid \text{ רציפה } f \text{ מחוץ לקטע כלשהו } f=0\}$$

**משפט 10.44**

$X_n\to X$ בהתפלגות אס"ס לכל $f\in C_c(\mathbb{R})$ מתקיים:

$$\mathbb{E}\left(f\left(X_n\right)\right)\underset{n\to\infty}{\longrightarrow}\mathbb{E}\left(X\right)$$

**הוכחה:** ראשית נוכיח את הלמה:

**למה 10.45**

אם $X_n,X$ מקיימים:

$$P\left(a\leq X_n\leq b\right)\underset{n\to\infty}{\longrightarrow}P\left(a\leq X\leq b\right)$$

לכל נקודות רציפות $a<b$ של $F_X$ אז $X_n\to X$ בהתפלגות.

**הוכחה:** **בשיעורי הבית.** ■

הרעיון: $a<b$ כמו בלמה אילו $\mathbb{1}_{[a,b]}\in C_c(\mathbb{R})$ אז הנתון היה:

$$\mathbb{E}\left(\mathbb{1}_{[a,b]}\left(X_n\right)\right)\underset{n\to\infty}{\longrightarrow}\mathbb{E}\left(\mathbb{1}_{[a,b]}\left(X\right)\right)$$

כלומר:

$$P\left(a\leq X_n\leq b\right)=\mathbb{E}\left(\mathbb{1}_{\{a\leq X_n\leq b\}}\right)\underset{n\to\infty}{\longrightarrow}\mathbb{E}\left(\mathbb{1}_{\{a\leq X\leq b\}}\right)=P\left(a\leq X\leq b\right)$$

ואז $X_n\to X$ בהתפלגות היה נובע מהלמה.

חזרה למציאות:

יהיו $a<b$ כנ"ל נוכיח:

$$\limsup P\left(a\leq X_n\leq b\right)\leq P*a\leq X\leq b$$

(ההוכחה לגבי $\liminf$ דומה).

נקבי פרמטר $\frac{b-a}{2}>\varepsilon>0$. קיימת $\mathbb{1}_{[a,b]}$ . קיימת $\varphi\in C_c(\mathbb{R})$ כך ש:

$$\mathbb{1}_{[a+\varepsilon,b-\varepsilon]}\leq\varphi\leq\mathbb{1}_{[a,b]}$$

$$P\left(a+\varepsilon\leq X_n\leq b-\varepsilon\right)=\mathbb{E}\left(\mathbb{1}_{[a+\varepsilon,b-\varepsilon]}\right)\leq\mathbb{E}\left(\varphi_\varepsilon\left(X_n\right)\right)\underset{n\to\infty}{\longrightarrow}\mathbb{E}\left(\varphi_\varepsilon\left(X\right)\right)$$

מהנתון

$$P\left(a\leq X\leq b\right)=\mathbb{E}\left(\mathbb{1}_{[a,b]}\left(X\right)\right)\underbrace{\geq}_{\varphi\leq\mathbb{1}_{[a,b]}}\mathbb{E}\left(\varphi_\varepsilon\left(x\right)\right)$$

לכן:

$$\limsup_{n\to\infty}P\left(a+\varepsilon\leq X_n\leq b-\varepsilon\right)\leq P\left(a\leq X\leq b\right)$$

(צירוף האי שיוויון ולקיחת $\limsup$) אבל:

$$P\left(a\leq X_n\leq b\right)=P\left(a+\varepsilon\leq X_n\leq b-\varepsilon\right)+P\left(a\leq X_n\leq a+\varepsilon\right)+P\left(b-\varepsilon\leq X_n\leq b\right)$$

לכן:

$$\limsup_{n\to\infty} P\left(a\le X_n\le b\right)\le$$

$$\le P\left(a\le X\le b\right)$$

$$\limsup_{n\to\infty} P\left(a+\varepsilon\le X_n\le b-\varepsilon\right)+\limsup_{n\to\infty} P\left(a\le X_n\le a+\varepsilon\right)+\limsup_{n\to\infty} P\left(b-\varepsilon\le X_n\le b\right)$$

אנו רוצים כי : $\limsup_{n\to\infty} P\left(a\le X_n\le a+\varepsilon\right),\limsup_{n\to\infty} P\left(b-\varepsilon\le X_n\le b\right)$ יהיו שווים לאפס. תהי $\psi\in C_c\left(\mathbb{R}\right)$ כך ש:

$$\mathbb{1}_{[a,a+\varepsilon]}\le\psi\le\mathbb{1}_{(a-\varepsilon,a+2\varepsilon]}$$

ראינו (אותו נימוק עם $\psi$ במקום $\varphi$) כי:

$$\limsup_{n\to\infty} P\left(a\le X_n\le a+\varepsilon\right)\le P\left(a-\varepsilon\le X\le a+2\varepsilon\right)\to 0$$

באותו אופן:

$$\limsup_{n\to\infty} P\left(b-2\varepsilon\le X_n\le b+\varepsilon\right)\le P\left(b-2\varepsilon\le X\le b+\varepsilon\right)\to 0$$

ולכן קיבלנו את הנדרש (כי $a,b$ נקודות רציפות של $F_X$) ולכן:

$$\limsup_{n\to\infty} P\left(a\le X_n\le b\right)\le P\left(a\le X\le b\right)+P\left(a-\varepsilon\le X\le a+2\varepsilon\right)+P\left(b-2\varepsilon\le X\le b+\varepsilon\right)$$

משאיפים את אפסילון לאפס ונקבל:

$$\limsup_{n\to\infty} P\left(a\le X_n\le b\right)\le P\left(a\le X\le b\right)$$

■

# 11 משפט הגבול המרכזי - Central Limit Theorem

**דוגמה 11.1** לראות את הדוגמה של Galton Board. יש הרבה דוגמאות בYoutube.

$X_1,X_2,\ldots$ משתנים מקריים בלתי תלויים שווי הסתפלגות ממשיים. $\sigma=\sqrt{\mathrm{Var}\left(X_1\right)}$, $\mu=\mathbb{E}\left(X_1\right)$. $\mathrm{Var}\left(X_1+X_2\right)=2\sigma^2$.

נבתונן ב: $X_1+\ldots+X_n$ אז נקבל כי:

$$\begin{aligned}\mathbb{E}\left(X_1+\ldots+X_n\right)&=n\mu\\ \mathrm{Var}\left(X_1+\ldots+X_n\right)&=n\sigma^2\end{aligned}$$

נבחין כי זה מתבדר. מחוק המספרים הגדולים, אם נחלק ב $n$ (כלומר $S_n=\frac{X_1+\ldots+X_n}{n}$ ) אז נקבל כי השונות מתכווצת למשהו נקודתי.

נגדיר:

$$W_n=\frac{1}{\sqrt{n\sigma^2}}\left(X_1+\ldots+X_n-n\mu\right)$$

הנרמול נבחר כך ש $\mathbb{E}\left(W_n\right)=0$ ו: $\mathrm{Var}\left(W_n\right)=1$ . אין בירה אחרת של הנרמול שייתן את זה.

**משפט 11.2**

$X_n$ בלתי תלויים, שוויו התפלגות ובעלי שונות:

$$W_n\to W\sim N\left(0,1\right)$$

בהתפלגות. (משתנה נורמלי עם תוחלת 0 ושונות 1)

**הערה 11.3** אם השונות של $X_n$ היא אפס, אז $X_n$ מקבל ערך $X_0$ בהסתברות 1. $\Leftarrow X_1+\ldots+X_n = nx_0$ בהסתברות 1. ואז נקבל כי $W_n = 0$ כמעט תמיד. ולכן זהו לא משתנה נורמלי.
כלומר יש מקרה מנוון אשר עבורו המשפט לא מתקיים.

**הערה 11.4** ניסוח חלופי למשפט:

$$\frac{1}{\sqrt{n}}\left(X_1+\ldots+X_n-n\mu\right) \underset{n\to\infty}{\longrightarrow} N\left(0,\sigma\right)$$

בהתפלגות.

**הערה 11.5** שלושה תתי משפטים:

- העובדה ש: $W_n$ מתכנס
- הגבול לא תלוי ב $X_n$ (אוניברסליות)
- הגבול הוא נורמלי

אם מאמינים לשניים הראשונים, אז השלישי נובע.
למה משתנה נורמלי? נזכור כי אם $U,V$ בלתי תלויים אז:

$$\begin{aligned} U &\sim N\left(\mu_U,\sigma_U^2\right) \\ V &\sim N\left(\mu_U,\sigma_V^2\right) \end{aligned}$$

גורר:

$$U+V\sim N\left(\mu_U+\mu_V,\sigma_U^2+\sigma_V^2\right)$$

לכן, אם $X_i\sim N\left(0,1\right)$ אז:

$$X_1+\ldots+X_n\sim N\left(0,n\right)\Rightarrow W_n=\frac{1}{\sqrt{n}}\left(X_1+\ldots+X_n\right)\sim N\left(0,1\right)$$

ראיו אם $Y$ נורמלי אז גם $aY+b$ נורמלי

**הוכחה:** תחת ההנחה של$X_n$ יש 3 מומנטים.
עוד עובדה: באופן כללי: $Y_n\to Y$ בהתפלגות) אס"ם $\mathbb{E}\left(f\left(Y_n\right)\right)\to\mathbb{E}\left(f\left(Y\right)\right)$ עבור כל $f$ גזירה 3 פעמים כך ש: $f'''\in C_c\left(\mathbb{R}\right)$
תהי $f$ כנ"ל. נניח לפשטות $\mu=0$ ו $\sigma=1$ (בה"כ)

$$W_n=\frac{1}{\sqrt{n}}\left(X_1+\ldots+X_n\right)$$

נגדיר משתנים מקריים נורמליים בלתי תלויים $Y_1,\ldots,Y_n\sim N\left(0,1\right)$ בלתי תלויים מה-$X_i$-ים.

$$Z_n=\frac{1}{\sqrt{n}}\left(Y_1+\ldots+Y_n\right)\sim N\left(0,1\right)$$

צריך להוכיח:

$$\mathbb{E}\left(f\left(W_n\right)\right)\to\mathbb{E}\left(f\left(W\right)\right)$$

כאשר $W\sim N\left(0,1\right)$. אנו יודעים כי:

$$\mathbb{E}\left(f\left(Z_n\right)\right)=\mathbb{E}\left(W\right)$$

(כל הוגדר).
נגדיר עבור $0 \le k \le n$:

$$\begin{aligned} W_{n,0} &= W_n \\ W_{n,k} &= \frac{1}{\sqrt{n}}\left(Y_1 + \ldots + Y_k + X_{k+1} + \ldots + X_n\right) \\ W_{n,n} &= Z_n \end{aligned}$$

$$\mathbb{E}\left(f\left(W_{n,0}\right)\right) = \mathbb{E}\left(f\left(W_{n,0}\right) - f\left(W_{n,1}\right) + f\left(W_{n,1}\right) - f\left(W_{n,2}\right) + \ldots - f\left(W_{n,n}\right) + f\left(W_{n,n}\right)\right) =$$

$$\mathbb{E}\left(f\underbrace{\left(W_{n,n}\right)}_{Z_n \sim W}\right) + \sum_{k=1}^{n}\mathbb{E}\left(f\left(W_{n,k}\right) - f\left(W_{n,k-1}\right)\right)$$

אנו רוצים להראות כי הסכום שואף לאפס ובכך נסיים את המשפט.
נבונן במחובר ה$k$. ממשפט טיילור על $f$ לסדר 2. לכל $u, v$:

$$f(u) = f(v) + a(u-v) + b(u-v)^2 + cf^{(3)}(w)(u-v)^3$$

כאשר $u \le w \le v$ (תלויה ב $u, v$) מדובר כמובן בשארית לגראנז'. נבחין כי:

$$\mathbb{E}\left(f\overbrace{\left(W_{n,k}\right)}^{u} - f\overbrace{\left(W_{n,k-1}\right)}^{v}\right) = \mathbb{E}\left(a\left(W_{n,k} - W_{n,k-1}\right) + b\left(W_{n,k} - W_{n,k-1}\right)^2 + cf'''(T)\right) =$$

$$a\mathbb{E}\left(\left(W_{n,k} - W_{n,k-1}\right)\right) + b\mathbb{E}\left(\left(W_{n,k} - W_{n,k-1}\right)^2\right) + c\mathbb{E}\left(f'''(T)\left(W_{n,k} - W_{n,k-1}\right)^3\right)$$

כאשר $T$ הוא משתנה מקרי.
נבחין כי: $W_{n,k} - W_{n,k-1} = \frac{X_k - Y_k}{\sqrt{n}}$ . ולכן:

$$\mathbb{E}\left(f\left(W_{n,k} - W_{n,k-1}\right)\right) = a\mathbb{E}\left(\frac{X_k - Y_k}{\sqrt{n}}\right)^{\to 0} + b\mathbb{E}\left(\left(\frac{X_k - Y_k}{\sqrt{n}}\right)^2\right)^{\to 0} + c\mathbb{E}\left(f'''(T)\left(\frac{X_k - Y_k}{\sqrt{n}}\right)^3\right)$$

אבל $f''' \in C_c(\mathbb{R})$ לפי הגדרה. לכן יש $M$ כך ש $|f'''| \le M$ ולכן:

$$\left|\mathbb{E}\left(f\left(W_{n,k}\right) - f\left(W_{n,k-1}\right)\right)\right| \le \frac{cM}{n^{3/2}}\mathbb{E}\left(\left(X_k - Y_k\right)^3\right)$$

אבל זה תלוי רק במומנטים של $X_k, Y_k$ עד 3. לכן:

$$\sum_{k=1}^{n}\mathbb{E}\left(f\left(W_{n,k}\right) - f\left(W_{n,k-1}\right)\right) \le \sum_{k=1}^{n} c \cdot M\frac{1}{n^{3/2}}\underbrace{(\text{לא תלוי ב} k)}_{d} = \frac{cdMn}{n^{3/2}} = \frac{CdM}{\sqrt{n}} \underset{n\to\infty}{\longrightarrow} 0$$

■

30/01/20

המשך טיפה של הפיתוח, נתבונן בחלקים השונים של התוחלת. נבחין כי הראשון (זה שנפל ב$a$) שווה ל:

$$\frac{a}{\sqrt{n\sigma^2}}\left(\mathbb{E}\left(X_k\right) - \mathbb{E}\left(Y_k\right)\right) = 0$$

החלק עם $b$:

$$\frac{c}{n\sigma^2}\mathbb{E}\left(X_k^2 - 2X_kY_k + Y_k^2\right) = \frac{c}{n\sigma^2}\left(\overbrace{\mathbb{E}\left(X_k^2\right)}^{\sigma^2} + \overbrace{\mathbb{E}\left(Y_k^2\right)}^{\sigma^2} - \overbrace{2\mathbb{E}\left(X_kY_k\right)}^{0}\right)$$

כאן יש בעיה, לא ברור למה זה מתאפס. זה לא באמת אמור, כי אנו סוכמים על $n$ איברים כנ"ל. מייק יתן הסבר במייל.

והחלק השלישי:

$$\frac{D}{n^{3/2}} = \underbrace{\sum Y_k, X_k \text{ מומנטים קטנים מ3 של }}_{\text{לא תלוי ב } k \text{ כי שווי התפלגות}} \frac{d\,|M|}{n^{3/2}\sigma^3} \geq \mathbb{E}\left(|X_k - Y_k|^3\right) \frac{d\,|M|}{n^{3/2}\sigma^3} \geq$$

**הערה 11.6**

1. לא צריך ש $X_k$ יהיו שווי התפלגות, מספיק $\mathbb{E}\left(|X_k|^3\right) \leq M$ לאיזשהו $M$ לא תלוי ב $k$

2. בהנחה של משפט גבול מרכזי נניח $U \sim N\left(\mu_u, \sigma_u^2\right)$, $V \sim N\left(\mu_v, \sigma_v^2\right)$ בלתי תלויים.

$$U + V = \lim\left(\frac{1}{\sqrt{n}\sigma_u}\left(X_1 + \ldots + X_n - n\mu_u\right) + \frac{1}{\sqrt{n}\sigma_v}\left(Y_1 + \ldots + Y_n - n\mu_v\right)\right) =$$

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\left(\left(\frac{X_1 - \mu_u}{\sigma_u} + \frac{Y_1 - \mu_v}{\sigma_v}\right) + \left(\frac{X_2 - \mu_u}{\sigma_u} + \frac{Y_2 - \mu_v}{\sigma_v}\right) + \ldots + \left(\frac{X_n - \mu_u}{\sigma_u} + \frac{Y_n - \mu_v}{\sigma_v}\right)\right)$$

וזהו משתנה נורמלי.

3. מסקנה: אוסף התפלגויות נורמליות הוא הוסף התפלגויות הקטן ביותר כך ש:

   (א) סגור תחת שינוי לינארי של קואורדינטות

   (ב) סגור תחת סכום של משתנים בלתי תלויים

   (ג) סגות תחת גבולות התפלגות

4. קשר עם חוק המספרים הגדולים: משפט גבול מרכזי⇐ החוק החלש

   **הוכחה:** יהיו $X_1, X_2, \ldots$ משתנים מקריים שווי התפלגות עם מומנט 3, $\mathbb{E}X_1 = 0$.

   יהי $\varepsilon > 0$, אנו רוצים כי:

$$P\left(|W_n| > \frac{\varepsilon}{\sigma}\sqrt{n}\right) = P\left(\left|\frac{X_1 + \ldots + X_n}{n}\right| > \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 0$$

   נקבע פרמטר $M > 0$ עבור $n$ גדול מספיק: $\frac{\varepsilon}{\sigma}\sqrt{n} > M$ עבור $n$ כזה::

$$P\left(|W_n| > \frac{\varepsilon}{\sigma}\sqrt{n}\right) \leq P\left(|W_n > M|\right) \underset{n\to\infty}{\longrightarrow} P\left(\underbrace{|W|}_{\sim N(0,\sigma^2)} > M\right)$$

   לכן, לכל $M > 0$:

$$0 \leq \limsup P\left(\left|\frac{X_1 + \ldots + X_n}{n}\right| > \varepsilon\right) \leq \underbrace{P\left(|W| > M\right)}_{\text{לא תלוי ב } n} \underset{M\to\infty}{\longrightarrow} 0$$

■

## 11.1 de Moivre-Laplace

נדון עכשיו בדרך חלופית להוכחה, בעזרת de Moivre-Laplace

**äøãâä 11.7** הפונקציה האופיינית (? לא בטוח שזה השם) של משתנה מקרי $X$ היא:

$$\phi_X(t) = \mathbb{E}\left(e^{tx}\right)$$

(לא בהכרח קיים לכל $t$)

**äîâåã 11.8** למשל, אם $X$ בעל צפיפות:

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,\mathrm{d}x$$

**11.9 èôùî**

אם $X, Y$ משתנה מקרי ו $\phi_X, \phi_Y$ מוגדרים בסביבת 0 ו: $\phi_X = \phi_Y$ בסביבת 0 אז $X, Y$ שווי התפלגות.

**תכונות**

1. $\frac{\mathrm{d}^n}{\mathrm{d}t^n} \phi_X\big|_{t=0}$ קיימת אם"ם קיים $n$ מומנטים.

$$\begin{aligned}
\mathbb{E}X &= \phi'_X(0) \\
\mathbb{E}\left(X^3\right) &= \phi''_X(0) \\
\frac{\mathrm{d}}{\mathrm{d}t}\phi_X(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} e^{tx} f_X(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} x e^{tx} f_x(x)\,\mathrm{d}x \underset{t\to 0}{\longrightarrow} \int_{-\infty}^{\infty} x f_x(x)\,\mathrm{d}x = \mathbb{E}(X)
\end{aligned}$$

שורה אחרונה בנפנוף ידיים...

2. אם $U, V$ בלתי תלויים אז: $\phi_{U+V} = \phi_U \phi_V$

**הוכחה:** לכל $t$, $e^{tU}, e^{tV}$ בלתי תלויים לכן:

$$\phi_{U+V}(t) = \mathbb{E}\left(e^{t(U+V)}\right) = \mathbb{E}\left(e^{tU}e^{tV}\right) = \mathbb{E}\left(e^{tU}\right)\mathbb{E}\left(e^{tV}\right) = \phi_U(t)\,\phi_V(t)$$

■

3. אם $W \sim N\left(\mu, \sigma^2\right)$ אז:

$$\phi_W(t) = e^{\left(\mu t + \frac{\sigma^2}{2} t^2\right)}$$

(חשוב!)

**משפט 11.10**

אם $X_n, X$ משתנים מקריים, $\phi_{X_n}, \phi_X$ מוגדרים בסביבה משותפת של 0 ולכל $t$ בסביבה זו $\phi_{X_n}(t) \to \phi_X(t)$ אז: $X_n \to X$ בהתפלגות.

**משפט 11.11 משפט הגבול המרכזי**

$X_n$ בלתי תלויים, שווי התפלגות עם מומנט 3, נניח $\mathbb{E}(X) = 0$ ו: $\operatorname{Var} X = 1$

$$W_n = \frac{1}{\sqrt{n}}\left(X_1 + \ldots + X_n\right)$$

$W \sim N(0,1)$ אז $W_n \to W$ בהתפלגות.

**הוכחה:**

$$\phi_{W_n} = \mathbb{E}\left(e^{t\left(\frac{X_1+\ldots+X_n}{\sqrt{n}}\right)}\right) = \mathbb{E}\left(e^{\frac{t}{\sqrt{n}}(X_1+\ldots+X_n)}\right) = \phi_{X_1+\ldots+X_n}\left(\frac{t}{\sqrt{n}}\right) =$$

$$\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\phi_{X_2}\left(\frac{t}{\sqrt{n}}\right)\cdot\ldots\cdot\phi_{X_n}\left(\frac{t}{\sqrt{n}}\right) = \left(\phi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

3 מומנטים $\Leftarrow$ $\phi_{X_1}$ גזירה 3 פעמים. ממשפט טיילור:

$$\phi_X(t) = \underbrace{\phi_X(0)}_{\substack{\mathbb{E}1=1\\ e^{0x}=1}} + \underbrace{\phi_X'(0)}_{\mathbb{E}x=0} t + \underbrace{\frac{\phi''(0)}{2}}_{\frac{\sigma^2}{2}} t^2 + o\left(t^2\right)$$

ולכן:

$$\phi_{W_n}(t) = \left(1 + \frac{\sigma^2}{2}\frac{t^2}{n} + o\left(\frac{t^2}{n}\right)\right)^n \underset{n\to\infty}{\longrightarrow} e^{\frac{\sigma^2 t^2}{2}} = \phi_W(t)$$

ולפי המשפט השני $W_n \to W$ בהתפלגות. ■

**הערה 11.12** היה חסר דרישה לקיום של הפונקציה האופיינית, אבל הכל היום בנפנוף ידיים וגם כך לא הוכחנו שום דבר שקשור לזה