<u>סיכומי שעורים בהסתברות (1), שנת 2008</u>

<u>מרצה: רז קופרמן</u>

סיכם: שיר פלד

ותודה ל: דינה זיל על האירוח באתר

הערת המקליד: אפשר וכדאי להשתמש בסיכומים אלו בצמוד לרשימות שרז עצמו מפרסם, בהן הטעויות פחותות והסדר רב יותר. היתרון העיקרי של הסיכומים המצורפים כאן הוא שהם בעברית ושלעיתים רז הוסיף הסברים והבהרות בעל פה שאינם רשומים בגוף הרשימות שלו.

בהטלת קוביה הוגנת, מה הסיכוי לקבל 4? מה המשמעות של הטענה הזו?

מרחב הסתברות

מאחורי כל בעיה הסתברותית עומד ניסוי (ואפילו תאורטי), והבעיה היא הנסיון למדל את הניסוי. כשאנחנו אומרים ניסוי הכוונה לאוסף כל התוצאות האפשריות, מכאן ש **ניסוי=קבוצה**, כאשר לקבוצה זו נקרא **מרחב המדגם** (Sample Space).
בד"כ נסמן את מרחב המדגם ב Ω .
דוגמאות:

1. הטלת מטבע פעם אחת – מרחב המדגם הוא עץ ופלי (=Heads or Tails): $\{H, T\}$
2. הטלה חוזרת של מטבע 3 פעמים, מרחב המדגם יהיה: $\{H, T\}^3$ (כלומר אוסף הסדרות באורך 3 של עץ ופלי)
3. הטלה של שתי קוביות שונות (=חשוב מי היא מי, נניח אחת כחולה ואחת אדומה), כאן מרחב המדגם יהיה זוגות סדורים של מספרים מ 1 עד 6, כלומר $\Omega=\{1,\dots,6\}^2$
4. הטלת שתי קוביות זהות – כל תוצאה שמופיעה פעמיים בסעיף הקודם – אפשר למחוק אחת מההישנויות, הדאבלים ישארו, כלומר : $\Omega=\{(i, j): 1\le i\le j\le 6\}$
5. משך חייו של אדם (בשנים), עקרונית יכול להיות $\Omega=\mathbb{R}+$
ניתן לטעון שאפשר לקטום את מרחב המדגם במקום מסויים כי אדם לא יכול לחיות 1000 שנה או משהו בסגנון.
6. יורים חץ למטרה בצורת עיגול יחידה, מרחב המדגם יכול להיות עיגול היחידה או $\mathbb{R}^2$ אם לא סומכים על היורה.
אפשר לסמן:
$$\Omega=\{(r, \theta), 0\le r\le 1, 0\le\theta\le 2\pi\}$$
$$\Omega=\{(x, y): x^2+y^2\le 1\}$$
7. הטלת מטבע אינסוף פעמים: $\Omega=\{H, T\}^{\aleph_0}$
8. מטילים מטבע – אם יוצא ראש – ניגשים למבחן בהסתברות ועוברים/נכשלים, יוצא זנב – הולכים לישון ומודדים את זמן השינה בשעות. $\Omega=\{H\}\times\{Fail, Pass\}\cup\{T\}\times\{\mathbb{R}_+\}$
9. תנועת בראון: $\Omega=C([0,\infty), \mathbb{R}^3)$ (תיאור תנועת חלקיק במרחב כפונקציה של הזמן)

הטלנו קוביה ונרצה לשאול מה ההסתברות שתוצאת ההטלה היא זוגית. כלומר תת קבוצה של מרחב המדגם. אך התוצאה "זוגית" לא נמצאת בקבוצה, כלומר היא תוצאה אפשרית אך אינה באוסף התוצאות שהגדרנו. אז אולי קצת רימינו בהגדרה כי הסתפקנו באוסף התוצאות שהן "אטומיות" או משהו בסגנון הזה.
יש לנו עניין בתתי קבוצות של מרחב המדגם, שכן בגינן על פי רוב תעלה השאלה – מה הסיכוי שיצא ... משהו?

מאורע (event) = תת קבוצה של מרחב המדגם
מספר תת הקבוצות $2^{|\Omega|}$
הבעיה – יש מצבים שבהם לא ניתן לבנות מרחב מדגם עבור כל המצבים האפשריים, קורה בקבוצות שאינן בנות מנייה. לא ניתן לקחת את כל תת הקבוצות של הקטע [0,1].

קבוצת המאורעות $\mathscr{F}\subseteq P(\Omega)$
תנאים ש $\mathscr{F}$ צריכה לקיים:
1. $A\in\mathscr{F}\Leftrightarrow A^c\in\mathscr{F}$
2. $\Omega\in\mathscr{F}$ ולכן $\emptyset\in\mathscr{F}$
3. $A, B\in\mathscr{F}\Rightarrow A\cup B\in\mathscr{F}$

מסקנות:
- אם 3 אז גם $A\cap B\in\mathscr{F}$, הוכחה: $A\cap B=(A^c\cup B^c)^c$
- אם $A_1, \dots, A_n\in\mathscr{F}$ אז איחוד כל הקבוצות הוא גם תת קבוצה של קבוצת המאורעות (קל מאד באינדוקציה)

לאוסף התנאים שהגדרנו קוראים **אלגברה של תת-קבוצות**, ולכן $\mathscr{F}$ היא בדיוק זה.
היינו רוצים ש $\mathscr{F}$ תהיה סגורה גם תחת מספר בן מניה של פעולות (ולא רק סופי שכפי שראינו נובע מהאלגברה) ואז נוכל לקרוא לה " σ **- אלגברה של תת קבוצות**".

הערות:

1. הערה טאוטולוגית לחלוטין: ניתן לרשום $A=\{\omega:\Omega:\omega\in A\}$
2. קיומם של יחידונים (singletones) והיותם שונים מאיברים בודדים.
3. הסיגמה-אלגברה הקטנה ביותר שאפשר לבנות היא $\{\emptyset,\Omega\}$
4. לכל $A\subseteq\Omega$ אפשר לבנות סיגמה-אלגברה $\{\emptyset,A,A^c,\Omega\}$
5. לכל אוסף $B\subseteq P(\Omega)$ אפשר להשלים אותו לסיגמה-אלגברה $\mathscr{F}$ קטנה ביותר המכילה את B. באופן עקרוני לוקחים את כל הסיגמה-אלגבראות המכילות את B וחותכים אותן, מכאן מגיעים למינימליות. קיום נובע מכך ש B מוכלת בקבוצה שהיא סיגמה-אלגברה.
6. תנועת בראון חד-ממדית, דהיינו פונקציה רציפה על $\mathbb{R}+$, כל אחת היא מאורע. נראה איך בונים סיגמה אלגברה. נסתכל על כל הפונקציות שבזמן 5 $(x=5)$ קיבלו ערכים בין 1 ל 3.

מרחב מדיד – Measurable Space $(\Omega,\mathscr{F})$

$A_1,A_2,\ldots$ סדרה אינסופית של מאורעות

נסתכל באיחוד האינסופי שלהן – נקבל קבוצה שמכילה את כל האיברים שקיימים בלפחות אחת מהקבוצות.

החיתוך האינסופי – מכילה רק את האיברים שקיימים בכולן.

$\limsup A_n=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k=$ לכל n קיים $k\geq n$ כך ש $\omega\in A_k$, כל ה ω המופיעות אינסוף פעמים : ω

$\liminf A_n=\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k=$ קיים n כך שלכל $k\geq n$ מתקיים $\omega\in A_k$: ω

במילים אחרות – limsup הוא קבוצת האיברים הקיימים ב A_k באופן שכיח ו liminf – קבוצת האיברים הקיימים ב A_k כמעט תמיד.

__הרצאה 6.11.08__

אם ה $\liminf A_k=\limsup A_k$ אז נאמר ש $\lim_{n\to\infty}A_k$ קיים.

דוגמא: $\Omega=\mathbb{N}$ ואז $\mathscr{F}=P(\mathbb{N})$ (כלומר הסיגמה אלגברה היא קבוצת החזקה)

$$A_k=\begin{cases}\{2n,n\in\mathbb{N}\}\text{ k even}\\ \{2n+1,n\in\mathbb{N}\}\text{ k odd}\end{cases}$$

$$\limsup A_n=\Omega$$

$$\lim inf\ A_n=\emptyset$$

עוד דוגמא:

$$A_k=\{k^0,k^1,k^2,\ldots\}$$

$$\lim_{n\to\infty}A_k=\{1\}$$

__הגדרה:__ סדרה של מאורעות תקרא עולה אם $A_1\subseteq A_2\subseteq\ldots$

והיא תקרא יורדת אם $A_1\supset A_2\supset\ldots$

טענה: אם A_k סדרה עולה של מאורעות אז הגבול קיים, והוא $\lim_{k\to\infty}A_k=\bigcup_{n=1}^{\infty}a_n$

הוכחה – נחשב את הגבול התחתון והעליון ונראה שהם שווים.

עבור סדרה עולה – האיחוד מהמקום ה n $\bigcup_{k=n}^{\infty}A_k$ הוא בעצם האיחוד מ 1, זה נובע מתכונת העליה.

אם כעת נחתוך מ n עד אינסוף, משום כך, נקבל שוב את אותה קבוצה: $\limsup_{n\to\infty}A_n=\bigcup_{k=1}^{\infty}A_k$

חיתוך על כל הקבוצות החל מהמקום ה k יתן את הקבוצה ה A_k (זה נובע מעליית הסדרה) ולכן:

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

הגדרה: יהא מרחב מדיד $(\Omega, \mathscr{F})$

<u>הסתברות</u> או (מידת הסתברות = probability measure)

היא פונקציה ממשית P מעל המאורעות $P: \mathscr{F} \to \mathbb{R}$ המקיימת:

1.לכל $A \subseteq \mathscr{F}$ מתקיים $0 \le P(A) \le 1$

2. $P(\Omega)=1$

3. אם A_n סדרה של מאורעות זרים אז $P(\bigcup_{n=1}^{\infty} A_n)=\sum_{n=1}^{\infty} P(A_n)$ (אדיטיביות בת מניה)

לשלשה $(\Omega, \mathscr{F}, P)$ קוראים <u>**מרחב הסתברות**</u>

טענה:

1. $P(\emptyset)=0$

2. אם A_k סדרה סופית של מאורעות זרים אז $P(\bigcup_{k=1}^{n} A_k)=\sum_{k=1}^{n} P(A_k)$

הוכחה $\Omega=\Omega \cup \emptyset \cup \emptyset \cup \ldots$

$$P(\Omega)=P(\Omega)+\sum_{k=2}^{\infty} P(\emptyset)$$

<u>דוגמאות</u>

.1

הטלת מטבע

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

$\mathscr{F}=\{\emptyset, \Omega, \{H\}, \{T\}\}$

כל בחירה של P ו Q שסכומם אחד היא לגיטימית, למשל $P(\{H\})=P(\{T\})=\frac{1}{2}$

.2

הטלת קוביה

$\Omega=\{1,2,3,4,5,6\}$

$\mathscr{F}=P(\Omega)$

עקרונית היינו צריכים לרשום את 2^6 האיברים ב f ולוודא שלכל זוג מאורעות זרים – הסתברות האיחוד שווה לסכום ההסתברויות. אבל אולי לא צריך לעבוד כל כך קשה.

$P(\{i\})=\frac{1}{6}$

וכאן כל תת קבוצה של Ω היא איחוד של יחידונים וזה קובע את ההסתברות שלה באופן יחיד, לכן לכל איחוד של קבוצות ניתן לעשות רדוקציה לסדרת איחודים של יחידונים ומכאן לקבוע את ההסתברות שלו באופן יחיד וכן הלאה.

הערה: את זה לא ניתן לעשות כאשר אומגה אינה קבוצה בת מניה, כיוון שלא ניתן להסתמך על האקסיומה כדי לאחד את כל היחידונים.

טענה: יהא $(\Omega, \mathscr{F}, P)$ מרחב הסתברות אז:

.1

לכל $A \in \mathscr{F}$ מתקיים $P(A^C)=1-P(A)$

.2

אם $A \subseteq B$ אז $P(A) \le P(B)$

הוכחה:

1.נסתמך על $P(\Omega)=1$ ולהעביר אגפים

2. נזכיר $B = A \cup B \setminus A$ וזה נובע מיד

טענה: יהא (Ω, f, P) מרחב הסתברות אז:
לכל $A, B \in f$ מתקיים $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
הוכחה (נציג הכל כאיחוד זר):

$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$$
$$A = (A \setminus B) \cup (A \cap B)$$
$$B = (B \setminus A) \cup (A \cap B)$$

$$P(A \cup B) = P(A \setminus B) + P(B \setminus A) + P(A \cap B)$$
$$P(A) = P(A \setminus B) + P(A \cap B)$$
$$P(B) = P(B \setminus A) + P(A \cap B)$$

<u>הרצאה 10.11</u>

הערה: כשאנחנו מוכיחים תכונה של מרחב ההסתברות זו תכונה של המבנה שנובעת מההגדרה האקסיומטית שלו, היא לא קשורה לשום מודל מתמטי מסויים, לא ידוע מה הניסוי, אין קוביות ומטבעות ולא מוציאים שום כדורים מתוך כובעים.

טענה: המשך משיעור קודם, יהיו $A, B, C \in f$ אז
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$$
הוכחה:
$$P((A \cup B) \cup C) = P(A \cup B) + P(C) - P((A \cup B) \cap C) = P(A) + P(B) - P(A \cap B) + P((A \cap C) \cup (B \cap C))$$
$$= P(A) + P(B) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$
משפט (עיקרון ההכלה וההדחה)
ובאופן כללי אם יש מאורעות $A_1, \dots, A_n$
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} \cdot P\left(\bigcap_{i=1}^{n} A_i\right)$$

<u>מרחב הסתברות בן מניה</u>
הקבוצה Ω היא בת מניה
נבחר $f = P(\Omega)$ (קבוצת החזקה, בפרט – היחידונים מאורעות)
מכיוון שאומגה בת מניה – נוכל לסמן $\Omega = \{1, 2, \dots\}$
נגדיר $P(\{a_i\}) = p_i$
נטען שזה מגדיר בצורה יחידה את מרחב ההסתברות, כי ניתן לכתוב כל מאורע כאיחוד זר בן מניה $A = \coprod_{a_i \in A} \{a_i\}$
ומתכונת פונקציית ההסתברות $P(A) = \sum_{a \in A} p_i$
זה בוודאות מתכנס כי סכימה על כל ה iים תיתן 1 לפי הגדרת פונקציית ההסתברות.
<u>מקרה פרטי</u>
מרחב הסתברות סופי, $|\Omega| < \infty$ ולכל היחידונים אותה הסתברות.
ההנחות שלנו על שוויון ההסתברות הן לא פועל יוצא של התכונות המתמטיות אלא הנחות כלפי המציאות (לפי זה ההסתברות לזכות בלוטו היא חצי – או שזוכים או שלא זוכים...)
בפרט Ω בת מניה ולכן
$$P(\Omega) = \sum_{i=1}^{n} p_i = p \cdot |\Omega| = 1$$
$$\forall i : p_i = p$$

$$p=\frac{1}{|\Omega|}$$
ומכאן:
$$P(A)=\sum_{a_i\in A}p_i=\frac{1}{|\Omega|}\cdot\sum_{a_i\in A}1=\frac{|A|}{|\Omega|}$$

ועכשיו כמה דוגמאות – רז מבקש לשים לב להבדל בין המודל המתמטי, המתמטיקה, והבעיה המיוצגת.

.1
מטילים שתי קוביות, מה ההסתברות שהסכום הוא שבע?
מרחב המדגם:
אפשר לבחור כמרחב המדגם כאוסף הסכומים, המספרים $2..12$, הנסיון היומיומי מלמד אותנו שלא כדאי לבחור כך, כיוון שאז לא נקבל מרחב שבו כל היחידונים שווי הסתברות.
הנסיון הזה מראה לנו שכדאי לקחת את אוסף הזוגות הסדורים.
$$\Omega=(1,\dots,6)^2=\{(i,j):1\leq i,j\leq 6\}$$
כעת נניח שכל היחידונים שווי הסתברות וש $f=P(\Omega)$
$$|\Omega|=36$$
$$A=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$
$$P(A)=\frac{|A|}{|\Omega|}=\frac{6}{36}=\frac{1}{6}$$
.2
יש קופסה ובה 11 כדורים, 6 מתוכם לבנים ו 5 שחורים.
מוציאים באקראי שני כדורים, מה ההסתברות שצבעם שונה?
נבחר מודל הסתברותי.
מהו מרחב המדגם? נדמיין שהכדורים הם מספרים בין 1 ל 11, ואז מרחב המדגם הוא אוסף הזוגות של מספרים בין 1 ל 11:
$$\Omega=\{i,j:1\leq i<j\leq 11\}$$
$$|\Omega|=\binom{11}{2}$$
$$A=\{(i,j):1\leq i\leq 6,7\leq j\leq 11\}$$
שוב נניח שכל המקרים שווי הסתברויות, נסביר זאת ע"י סימטריה, זו לא תכונה של המתמטיקה המעורבת בעניין אלא של העולם שאנחנו מייצגים. *רז: זו תכונה דתית, לא מתמטית.*

$$P(A)=\frac{\binom{6}{1}\binom{5}{1}}{\binom{11}{2}}=\frac{6}{11}$$
ב. בואו ננסה מודל אחר, נחשוב על הכדורים כעל זוג סדור:
$$\Omega=\{(i,j):i\neq j,1\leq i,j\leq 11\}$$
$$|\Omega|=110$$
כעת אנחנו מבינים שעל כל תוצאה שספרנו קודם יש גם את הסימטריה שלה (לבן שחור = שחור לבן) ולכן מספר המקרים כפול, והפלא ופלא – גם גודל מרחב ההסתברות כפול, מכאן שההסתברות נותרה בעינה.

.3
מחלקים חפיסה בת 52 קלפים ל 4 שחקנים.
מה ההסתברות שאחד השחקנים קיבל את כל התלתנים.
מרחב המדגם – כל הדרכים לחלק 52 קלפים בין 4 שחקנים.
$$|\Omega|=\frac{52!}{(13!)^4}$$

נגדיר את המאורע A_i = השחקן ה i קיבל את כל התלתנים
ו A = אחד השחקנים קיבל את כל התלתנים
ולכן:
$$|A_i|=\frac{39!}{(13!)^3}=\frac{39!\cdot(13!)^4}{(13!)^3\cdot 52!}=\frac{1}{\binom{52}{13}}$$
הסיכוי ל A הוא סכום ה A_i מכיוון ש ה A_i הם מאורעות זרים (=מוציאים זה את זה) , ואז
$$P(A)=\frac{4}{\binom{52}{13}}=6.310^{-12}$$

<u>פרדוקס ימי ההולדת (זה לא פרדוקס)</u>
יש n אנשים בקבוצה אקראית, ונשאל מה ההסתברות שאין שניים שיש להם אותו יום הולדת.
Ω כל ההתאמות של ימי הולדת ל n אנשים, כלומר $|\Omega|=365^n$
A = לכל שני אנשים יום הולדת שונה
$$|A|=365\cdot 364\cdot\ldots\cdot(365-n+1)$$
(נזכיר שוב שכל מה שאנחנו עושים תלוי בזה שלכל היחידונים אותה הסתברות)
$$P(A_n)=\frac{|A_n|}{|\Omega|}=\frac{1\cdot 364}{365}\cdot\ldots\cdot\frac{(365-n+1)}{365}$$
$$P(A_{23})<0.5 \qquad P(A_{50})<0.03 \qquad P(A_{100})<3.3\cdot 10^{-7}$$
<u>הפקיד המפוזר (= המזכירה המבולבלת, המלחים השיכורים וכו')</u>
יש n מכתבים ו n מעטפות, מכניסים כל מכתב למעטפה כלשהי באקראי, נשאל מה ההסתברות שאף נמען לא קיבל את המכתב המיועד לו.
כמו בקורס במתמטיקה דיסקרטית, לפנינו הבעיה של תמורות ונקודות שבת, ולכן כל סידור של מכתבים למעטפות הוא תמורה על המספרים בין 1 ל n. ומכאן $S_n=\Omega$
$$|\Omega|=n!$$
A = אף מכתב לא הגיע לתעודתו
מסתבר שיותר קל לחשב את המשלים ואז $P(A)=1-P(A^C)$
A^C = לפחות מכתב אחד הגיע לתעודתו
B_i = המכתב ה i הגיע לתעודתו
$$A^C=\bigcup_{i=1}^{n} B_i$$
$$P(A^C)=\sum_i P(B_i)-\sum_{i<j}P(B_i\cap B_j)+\sum_{i<j<k}P(B_i\cap B_j\cap B_k)-\ldots+(-1)^{n+1}P(\bigcap_{i=1}^{n}B_i)$$
$$=n\,P(B_1)-\binom{n}{2}P(B_1\cap B_2)+\binom{n}{3}P(B_1\cap B_2\cap B_3)-\ldots+(-1)^{n+1}P(\bigcap_i B_i)$$
$$|B_i|=(n-1)!$$
$$P(B_1)=\frac{(n-1)!}{n!}=\frac{1}{n}$$
$$P(B_1\cap B_2)=\frac{(n-2)!}{n!}=\frac{1}{n\cdot(n+1)}$$
נציב זאת בתוך ההכלה וההדחה ונקבל:
$$P(A^C)=n\cdot\frac{1}{n}-\binom{n}{2}\cdot\frac{(n-2)!}{n!}+\binom{n}{3}\cdot\frac{(n-3)!}{n!}+\ldots+(-1)^{n+1}\cdot\frac{1}{n!}$$
$$=1-\frac{1}{2!}+\frac{1}{3!}-\ldots=\sum_{i=1}^{n}\frac{(-1)^i}{i!}$$

$$P(A)=1-P(A^C)=\sum_{i=0}^{n}\frac{(-1)^i}{i!}\ \underset{n\to\infty}{\to}\ \frac{1}{e}\simeq 0.36$$
$$|A|=|\Omega|P(A)=n!\cdot\sum_{i=0}^{n}\frac{(-1)^i}{i!}$$
מהי ההסתברות שבדיוק k נמענים קיבלו את המכתב המיועד.

C = המאורע ש k הנמענים הראשונים קיבלו את המכתב המיועד.

נתעלם מ k הנמענים הראשונים ונשחזר את החישוב הקודם עבור שאר הנמענים שמספרם n-k, כלומר:
$$|C|=(n-k)!\cdot\sum_{i=0}^{n-k}\frac{(-1)^i}{i!}$$
D = בדיוק k נמענים קיבלו מכתב מיועד

נטען ש $|D|=\binom{n}{k}(n-k)!\sum_{i=0}^{n-k}\frac{(-1)^i}{i!}$

כי כל ה C זרים (=מוציאים זה את זה)
$$P(D_k)=\frac{1}{n!}\cdot\binom{n}{k}(n-k)!\sum_{i=0}^{n-k}\frac{(-1)^i}{i!}=\frac{1}{k!}\cdot\sum_{i=0}^{n-k}\frac{(-1)^i}{i!}\ \underset{n\to\infty}{\to}\ \frac{e^{-1}}{k!}$$

<u>הרצאה 13.10</u>

מרחב הסתברות $(\Omega,\mathscr{F},P)$

<u>"האינפי של מרחבי הסתברות"</u>

משפט: פונקציית ההסתברות היא רציפה

באפיון לפי סדרות - אם $\lim_{n\to\infty}A_n$ קיים אז $P(\lim_{n\to\infty}A_n)=\lim_{n\to\infty}P(A_n)$

משפט: אם A_n סדרה עולה של מאורעות אז $P(\lim A_n)=\lim P(A_n)$

הוכחה:

ידוע ש $\lim_{n\to\infty}A_n=\bigcup_{n=1}^{\infty}A_n$

נגדיר:
$$B_1=A_1$$
$$B_2=A_2\setminus A_1$$
.

.
$$B_n=A_n\setminus A_{n-1}$$
נטען ש B_n סדרה של מאורעות זרים.

לכל n:
$$\coprod_{k=1}^{n}B_k=\bigcup_{k=1}^{n}A_n$$
וגם:
$$\coprod_{k=1}^{\infty}B_k=\bigcup_{k=1}^{\infty}A_n$$
ואז:
$$P(\lim A_n)=P(\bigcup_{k=1}^{\infty}A_k)=P(\coprod_{k=1}^{\infty}B_k)=\sum_{k=1}^{\infty}P(B_k)=\lim_{n\to\infty}\sum_{k=1}^{n}P(B_k)=\lim_{n\to\infty}P(\coprod_{k=1}^{n}B_k)=\lim_{n\to\infty}P(A_n)$$
תרגיל: כנ"ל לגבי סדרה יורדת.

כרגע הסתמכנו מאד על העליה של הקבוצה (או הירידה שלה), אבל אנחנו מחפשים את המקרה הכללי.

למה (Fatou)

לכל סדרת מאורעות A_n

$$P(\lim inf\ A_n) \le \lim inf\ P(A_n)$$

הוכחה:

ניזכר $\lim inf$ הוא קבוצת האיברים ש"כמעט תמיד" מופיעים בתוך איברי הסדרה (שהם קבוצות).

נסמן H_n על ידי: $H_n = \bigcap_{k \ge n} A_k$

ואז:

$$\lim inf\ A_n = \bigcup_{n=1}^{\infty} \bigcap_{k \ge n} A_k = \bigcup_{n=1}^{\infty} H_n$$

H_n בוודאי סדרה עולה של מאורעות ולכן:

$$P(\lim inf\ A_n) = P(\lim H_n)$$

$$= \lim P(H_n) \le \lim inf\ P(A_n)$$

$$H_n \subseteq A_n \ \Rightarrow\ P(H_n) \le P(A_n)$$

למה (reverse Fatou)

לכל סדרת מאורעות A_n

$$P(\lim\sup\ A_n) \ge \lim\sup\ P(A_n)$$

הוכחה:

$$\lim\sup\ A_n = \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{k \ge n} A_k}_{G_n}$$

G_n סדרה יורדת ולכן (לפי ה"תרגיל לבית"):

$$P(\lim\sup\ A_n) = P(\lim G_n) = \lim P(G_n)$$

(קיים גבול לסדרה G_n והוא החיתוך של איברי הסדרה)

לכל n מתקיים $A_n \subseteq G_n$ ומכאן $P(G_n) \ge P(A_n)$

$$\lim\ P(G_n) \ge \lim\sup\ P(A_n)$$

<u>נסכם את התוצאות מהלמה (והלמה ההפוכה):</u>

$$\lim\sup\ P(A_n) \le P(\lim\sup\ A_n) \qquad P(\lim inf\ A_n) \le \lim inf\ P(A_n)$$

נשים לב שהאגף הימני ביותר קטן/שווה מהשמאלי ביותר, ולכן אם נצליח ליצור שוויון בין שני האמצעיים – נקבל שכולם שווים (מטרנזיטיביות).

אם קיים גבול ל A_n הם בוודאי שווים (כי הגבול התחתון והעליון שווים, ובוודאי ההסתברויות שלהם) ומכאן: פונקציית ההסתברות רציפה.

הלמה ה"קלה" (Borel – Cantelli)

אם A_n סדרת מאורעות המקיימים $\sum_{n=1}^{\infty} P(A_n) < \infty$ אז $P(\lim\sup\ A_n) = 0$

הוכחה:

(בעיקרון אם ההסתברות של הגבול העליון אינה 0 – נובע שיש מאורעות שחוזרים אינסוף פעמים ב P שהסתברותם אינה 0, ולכן הסכום של ההסתברויות, שהן אי שליליות, לא יהיה סופי)

$$\lim\sup\ A_n = \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{k \ge n} A_k}_{=G_n}$$

כאמור G_n סדרה יורדת ולכן

$$P(\limsup A_n) = \limsup P(G_n)$$

הערה:

$$P(A\cup B)\le P(A)+P(B) \quad , \quad P(\cup_n A_n)\le \sum_n P(A_n)$$

נחזור להוכחה:

$$P(G_n)\le \sum_{k=n}^{\infty} P(A_k)=0$$

(אגף ימין הוא זנב של טור מתכנס לפי ההנחה)

❑

נשאל את השאלה: בדקה ל 12 יש קופסה ריקה, נכניס לתוכה את הכדורים 1-10 ואז נוציא באקראי אחד מהכדורים בקבוצה.
בחצי דקה ל 12 נכניס את הכדורים 10-20, ונוציא באקראי אחד מהכדורים בקופסה.
ברבע דקה ל 12 נכניס את הכדורים 20-30 ונוציא באקראי כדור מהקופסה.
נשאל: מה ההסתברות שכדור מספר 1 נשאר בקופסה בשעה 12?

E_n = קבוצת המאורעות כך ש 1 נשאר בקופסה עד וכולל השלב ה-n
(1 אינו אחד מ n המספרים הראשונים בסדרה)
נסתכל בסדרה E_n ונראה שהיא סדרה יורדת
$\bigcap_{n=1}^{\infty} E_n$ (1 אינו מופיע בסדרה)

$$P(\lim E_n)=\lim P(E_n)=\lim_n \frac{9}{10}\times\frac{18}{19}\times\frac{27}{28}\times\ldots\times\frac{9n}{9n+1}=\lim\prod_{k=1}^{n}\frac{9k}{9k+1}$$

$$=\lim_n \frac{1}{\prod \frac{9k+1}{9k}}=\frac{1}{\lim\prod_{k=1}^{n}\left(1+\frac{1}{9k}\right)}=0$$

<u>שיעור חמישי – 17.11</u>

מטילים שתי קוביות, מה ההסתברות שהסכום הוא 8?
(אמור להיות די פשוט בשלב זה – נבנה מרחב מדגם של הזוגות הסדורים, נראה מה החלק היחסי של הזוגות שסכומם 8)

$$\Omega=\{1,2,3,4,5,6\}^2$$

$$A=\{(i,j)\in\Omega : i+j=8\}$$

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

$$\mathscr{F}=P(\Omega) \;\Rightarrow\; P(\{(i,j)\})=\frac{1}{36} \;\Rightarrow\; P(A)=\frac{|A|}{|\Omega|}=\frac{5}{36}$$

מה יקרה אם רז יסתיר את התוצאה אחרי ההטלה? זה לא ישנה את ההסתברות מבחינתנו.
מה יקרה אם הוא יספר לנו שבקוביה אחת קיבלנו 3?
זה ישנה את "מידת האמונה" שלנו לגבי התוצאה וגם את המודל המתמטי.

אם כן – מה ההסתברות שהסכום יצא 8 אם נתון שהקוביה הראשונה יצאה 3?

$$F=\{(3,j)\in\Omega : i\le j\le 6\}$$

באופן אינטואיטיבי יש רק 6 תוצאות שמקיימות את F ורק אחת מהן היא גם ב A, ולכן ההסתברות תהיה $\frac{1}{6}$

נכתוב $\frac{|A\cap F|}{|F|}=\frac{|A\cap F|/|\Omega|}{|F|/|\Omega|}=\frac{P(A\cap F)}{P(F)}$

נעזוב את המציאות ונגדיר משהו ב"עולם המודלים" שלנו:

<u>הגדרה: הסתברות מותנית</u> - יהיו $(\Omega, \mathscr{F}, P)$ מרחב הסתברות. יהיו A,B מאורעות, כאשר $P(B)\geq 0$

אזי $P(A|B)=\frac{P(A\cap B)}{P(B)}$ (באגף שמאל זה הסימון ל A בהנתן B)

דוגמה:

בקופסה 10 כדורים לבנים, 5 צהובים ו 10 שחורים.

מוציאים כדור באקראי, מה ההסתברות שהוא צהוב?

$\Omega=\{1,\dots,25\}$

$P(\{i\})=\frac{1}{25}$

$A=\{11,\dots,15\}$ = יצא כדור צהוב

ולכן:

$P(A)=\frac{5}{25}=\frac{1}{5}$

נשאל – מה ההסתברות שיצא צהוב בהנתן שיצא כדור שאינו שחור?

$B=\{1,\dots,15\}$ = יצא כדור שאינו שחור

$B\cap A=A$

ולכן על פי ההגדרה שלנו:

$$P(A|B)=\frac{P(A)}{P(B)}=\frac{\frac{1}{5}}{\frac{15}{25}}=\frac{1}{3}$$

דוגמה:

יש מישהו (*רז: בלי שמות*) שצריך לגשת לבחינה, ושתי בחינות מתנגשות לו באותו יום – בחינה בהסתברות ובחינה בספרות.

(רז: הבחירה בהסתברות היא כדי שלכל המשפטים יהיה כפל משמעות כרגע)

ספרות – יעבור בהסתברות חצי

הסתברות – יעבור בהסתברות שליש (*רז: זה בערך נכון*)

הוא קיבל החלטה ע"י הטלת מטבע.

השאלה: מה ההסתברות שהוא ניגש לבחינה בהסתברות ועבר אותה?

הרעיון הוא שאנחנו שואלים כעת שאלות שאפשר לפתור גם בלי לבוא לקורס, אבל בהמשך נראה שאלות שבהן לא ניתן להמשיך את האינטואיציה היומיומית שלנו הלאה, ולכן נפרמל את הרעיונות האלו.

התוצאה של הניסוי שלנו היא זוג סדור $\Omega=\{\text{ספרות}, \text{הסתברות}\}\times\{0,1\}$

נסמן ספרות ב lit והסתברות ב prob

$A=\{lit\}\times\{0,1\}$ ספרות

$B=A^C=\{prob\}\times\{0,1\}$ הסתברות

$C=\{lit, prob\}time\{1\}$ עבר

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

$P(C|A)=\frac{1}{2} \quad P(C|B)=\frac{1}{3}$

$P(\{(lit,1)\})=P(A\cap C)=P(C|A)P(A)=\frac{1}{2}\times\frac{1}{2}$

$$P(\{(prob,1)\})=P(B\cap C)=P(C|B)P(B)=\frac{1}{3}\times\frac{1}{2}$$

מה ההסתברות, למשל של האפשרות שניגשתי לבחינה בספרות ונכשלתי?

$$P(\{(lit,0)\})=P(A\cap C^C)=P((A^C\cup C)^C)=1-P(A^C\cup C)=1-\underbrace{P(B\cup C)}_{A^C=B}$$

$$=1-P(B)-\underbrace{P(C)}_{P(C)=P(C\cap(A\cup B))}+P(B\cap C)=1-P(B)-P(C\cap A)-\cancel{P(C\cap B)}+\cancel{P(B\cap C)}=\frac{1}{4}$$

דוגמה אחרונה:

כרגיל יש לנו כדורים בתוך קופסה – 8 אדומים, 4 לבנים.

מוציאים שני כדורים בזה אחר זה, מה ההסתברות שהשני אדום בהנתן שהראשון אדום?

מרחב המדגם הפעם יהיה כל הזוגות הסדורים בלי חזרות באיברים $1,...,12$ כאשר 8 הראשונים אדומים והשאר לבנים.

$$\Omega=\{(i,j):1\le i,j\le 12\ \ ,i\neq j\}\quad ,\quad |\Omega|=12\times 11$$

A = הכדור הראשון אדום

B = הכדור השני אדום

לכן $|A|=8\times 11$ וגם $|B|=8\times 11$

יש לנו את המאורע שהוא החיתוך: $|A\cap B|=8\times 7$

ההסתברות של B בהנתן A עפ"י ההגדרה היא:

$$P(B|A)=\frac{P(B\cap A)}{P(A)}=\frac{8\times 7}{8\times 11}=\frac{7}{11}$$

אפשר גם לפתור את זה בדרך האלמנטרית – לקחנו כדור אדום? יופי, נשארנו עם 11 כדורים ש 7 מתוכם אדומים, מה הסיכוי לקחת אדום? זה קל $\frac{7}{11}$

<u>משפט:</u> יהא $(\Omega,\mathscr{F},P)$ מרחב הסתברות

F מאורע בעל הסתברות חיובית

לכל מאורע A נגדיר את $Q(A)=P(A|F)$ (ההסתברות של A בהנתן F)

Q היא פונקציית הסתברות על $(\Omega,\mathscr{F})$

הוכחה:

- Q מוגדרת על $\mathscr{F}$ אבל זה מוגדר היטב על פי הגדרתנו את $P(A|F)$ מקודם באמצעות P וכו'.
- $Q(A)\ge 0$

$$Q(\Omega)=P(\Omega|F)=\frac{P(\Omega\cap F)}{P(F)}=1$$

- A_n סדרה של מאורעות זרים, $(A_n\cap F)$ גם כן סדרה של מאורעות זרים ולכן:

$$Q(\coprod_{n=1}^{\infty}A_n)=\frac{P((\coprod A_n)\cap F)}{P(F)}=\frac{\sum_{n=1}^{\infty}P(A_n\cap F)}{P(F)}=\sum_{n=1}^{\infty}Q(A_n)$$

הערה:

$$Q(F)=\frac{P(F\cap F)}{P(F)}=1$$

כלומר ניתן להגדיר את Q על המרחב המדיד $(F,\mathscr{F}\cap F)$

נחדד הבדל חשוב : כשאני מגדיר משהו כהסתברות מותנה – אני מתכוון להגדרה המתמטית, להחליט שבעיה מילולית כלשהי ניתנת לייצוג באופן זה היא החלטה של מידול המציאות, ולא תכונה של ההגדרה המתמטית.

חוק ההסתברות השלמה (Law of of complete probability)

תהי Ω

נניח שיש חלוקה של Ω , כלומר איחוד זר וממצה:

$$A_1, \ldots, A_n \in \mathscr{F} \quad \text{כאשר} \quad \coprod_{i=1}^{n} A_i = \Omega$$

$$B = \coprod_{i=1}^{n} (B \cap A_i)$$

נובע מכך:

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i)$$

אם לכל i נדרוש $P(A_i) > 0$ נקבל

$$P(B) = \sum_{i=1}^{n} P(B|A_i) P(A_i) \quad \text{(זהו חוק ההסתברות השלמה)}$$

חוק Bayes

A, B מאורעות בעלי הסתברות חיובית אז

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \quad \text{(פשוט מניפולציה על הגדרת ההסתברות המותנה)}$$

ולכן חוק בייס נובע מיידית:

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

הכללה:

יהי B מאורע ו $(A_i)_{i=1}^{\infty}$ סדרת מאורעות זרים המכסים את Ω

$$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{P(B)} = \frac{P(B|A_i) P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j) P(A_j)}$$

<u>דוגמה</u>

מעבדה לבדיקת HIV, נשא יוצא חיובי בהסתברות 95% ואדם לא-נשא יצא חיובי בהסתברות 1%.

0.5% מהאוכלוסיה נשאים.

(הכל כמובן שקרים וסטטיסטיקה)

מה ההסתברות שאדם שיצא חיובי הוא נשא?

נצטרך לבנות מרחב הסתברות, חשוב לשים לב שמרחב המדגם יכול להיות ארבע אפשרויות – נשא = C, לא נשא = H וחיובי p ושלילי n.

$$\Omega = \{n, p\} \times \{C, H\}$$

A המאורע שאדם נשא כלומר $A = \{C\} \times \{n, p\}$

B המאורע שאדם מקבל חיובי $B = \{C, H\} \times \{p\}$

נתון:

$$P(B|A) = 0.95 \quad P(B|A^C) = 0.01 \quad P(A|B) = ?$$

סיימנו לתרגם את הבעיה למודל הסתברותי, אפשר לשכוח מהסיפור ולענות על השאלה בתוך הפורמליזם ההסתברותי.

מחוק בייס נקבל:

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

חסרה לנו רק ההסתברות של B, ולפי חוק ההסתברות השלמה נקבל:

$$\frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^C)P(A^C)}=\frac{0.95\cdot 0.005}{0.95\cdot 0.005+0.001\cdot 0.995}\simeq\frac{1}{3}$$

ובאופן מפתיע נקבל שאם קיבלנו תוצאה חיובית – ההסתברות שאנחנו אכן נשאים היא קטנה מחצי.

<u>ניסויים מורכבים</u>

דוגמה:

ההסתברות שבמשפחה k ילדים היא p_k כך שהטור של p_k מתכנס ל 1.

בהנתן שיש k ילדים – יש הסתברות שווה לכל צירוף של בנים ובנות.

מה מרחב המדגם? כל הסדרות הסופיות של בנים ובנות:

$$\Omega=\{a_1,\dots,a_n : a_i\in\{b,g\}, n\in\mathbb{N}\}$$

כדי להגדיר מרחב הסתברות – אנחנו צריכים לתת הסתברות לכל יחידון.

לדוגמה מה ההסתברות של בן,בן,בת ?

$$P(\{b,b,g\})=P(\{b,b,g\}|\text{“three kids”})\cdot P(\text{“three kids”})=\frac{1}{8}\cdot p_3$$

ולכן:

$$P(\{a_1,\dots,a_n\})=\frac{1}{2^n}\cdot p_n$$

שאלה: מה ההסתברות שלמשפחה בדיוק ילד אחד בהנתן שאין בה בנות? (כתוצאה מכך ברור שהילד הוא בן)

A_k = בדיוק k ילדים (=צאצאים)

B = אין בנות

ולפי חוק בייס:

$$P(A_1|B)=P(B|A_1)\cdot\frac{P(A_1)}{P(B)}=\frac{P(B|A_1)P(A_1)}{\sum_{i=1}^{\infty}P(B|A_i)P(A_i)}=\frac{\frac{1}{2}\cdot p_1}{\sum_{i=1}^{\infty}\frac{1}{2^i}\cdot p_i}$$

נסתכל על משפחות בנות שני ילדים:

$$\Omega=\{(b,b),(b,g),(g,b),(g,g)\}$$

נשאל – בהנתן שיש לפחות בן אחד – מה ההסתברות שיש לו אחות? (נניח שהאפשרויות שוות הסתברות)

התשובה - שליש

סיפור שתכליתו לשגע אותנו עד השיעור הבא:

משרד הפנים הוציא הוראה לצבוע את משקופי הדלתות של משפחות שבהן 2 ילדים בדיוק בורוד.

תלמידה: מי צובע את המשקופים?

רז: המשפחה

אני מגיע לדלת ורודה – פותח לי ילד ואומר לי "שלום, אני הילד הצעיר במשפחה", מה ההסתברות שיש לו אחות?

אני מגיע לדלת ורודה – פותח לי ילד ואומר לי "שלום, אני הילד המבוגר במשפחה", מה ההסתברות שיש לו אחות?

אני מגיע לדלת ורודה – פותח לי ילד מבוהל ולא אומר מילה, מה ההסתברות שיש לו אחות? עכשיו הוא אומר לי שהוא הצעיר, מה ההסתברות שיש לו אחות כעת?

<u>שיעור שישי – 20.11</u>

נזכיר הבדל פילוסופי בהסתברות מותנית $P(A|B)$

הבייסיאנים (Bayesian): זוהי מידת האמונה שלנו שהתרחש A בהנתן ש B התרחש.

התדירותניים (Frequentists): נספור את חלקם היחסי של המאורעות A בכלל המאורעות B.

נחזור לפרדוקס הילדים

$$A=\{bb,bg,gb\}$$
$$B=\{bg,gb,gg\}$$

$$P(A|B)=\frac{P(A\cap B)}{P(A)}=\frac{2}{3}$$

הבכור הוא בן = $C=\{bb,bg\}$

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

השאלה היא, במובן מסויים, האם בבית הזה תמיד מטילים מטבע מי פותח את הדלת (בן או בת) או אם-יש-בן-אז-הוא-פותח-את-הדלת?

לכן מרחב המדגם הוא מורכב ובנוי מ – מי פתח את הדלת וגם – מי הילדים בבית?

$$\{bb,bg,gb,gg\}\times\{young,old\}$$

נסמן:

A = בן פתח את הדלת

B = יש לפחות בת אחת במשפחה

$$A=\{(bb,young),(bb,old),(bg,old),(gb\ young)\}$$

$$B=\{bg,gb,gg\}\times\{young,old\}$$

$$P(B|A)=\frac{2/8}{4/8}=\frac{1}{2}$$

ניזכר בחוק Bayes

$$P(A|B)=P(B|A)\cdot\frac{P(A)}{P(B)}\ \Rightarrow\ \underbrace{P(A_i|B)}_{\text{posterior belief}}\propto\underbrace{P(A_i)}_{\text{prior belief}}\cdot\underbrace{P(B|A_i)}_{\text{likelihood of data}}$$

זו הסתכלות בייסיאנית על הנוסחה.

בקורס שלנו זה לא כל כך חשוב, זה יותר משמעותי בסטטיסטיקה.

<u>אי תלות</u>

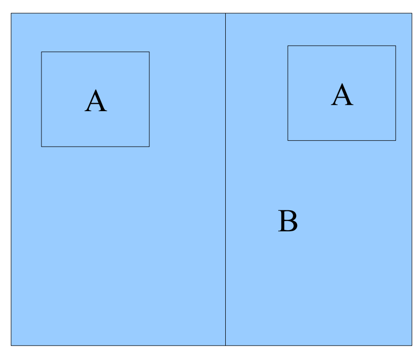
<u>הגדרה</u>: נאמר שמאורע A **בלתי תלוי** במאורע B אם הידיעה ש B אירע "אינה משנה" את ההסתברות ש A אירע.

כלומר: $P(A|B)=P(A)$

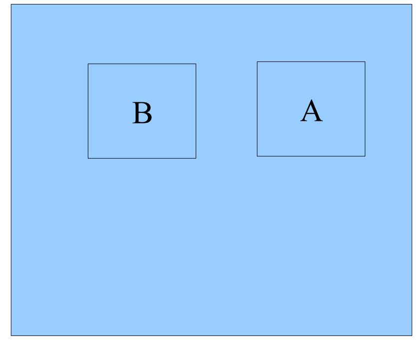
הערה חשובה: זרות <u>אינה</u> אי תלות, כיוון שמאורעות מוציאים הם תלויים ($A\rightarrow not\quad B$)

איורים:

אי תלות בין A ו B



זרות בין A ו B



זרות $\Leftarrow$ "לחבר"

אי-תלות $\Leftarrow$ "להכפיל"

דרך אחרת – אפשר להגדיר ששני מאורעות הם בלתי תלויים אם $P(A\cap B)=P(A)P(B)$, זו בעצם העברת אגפים מההגדרה הקודמת, אלא שאנחנו לא דורשים כעת ש $P(B)\neq 0$ שזה במובן מסויים יותר נוח, בנוסף – כמו שהגדרנו אותה כרגע – מתברר שההגדרה היא סימטרית, מקומוטטיביות של חיתוך וכפל.

<u>דוגמה לאי-תלות:</u>

בוחרים באקראי (=יחידונים שווי הסתברות) קלף מחפיסת קלפים סטנדרטית בת 52 קלפים.

A= הקלף הוא תלתן

B = הקלף הוא נסיך

האם A ו B בלתי תלויים?

$A\cap B$ = הקלף הוא נסיך תלתן

$$P(A)=\frac{13}{52}$$

$$P(B)=\frac{4}{52}$$

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

נשים לב שמתקיים כאן $P(A)\cdot P(B)=\frac{1}{52}=P(A\cap B)$ ולכן המאורעות בלתי תלויים

מסקנה:

רז: הנסיכיות אינה תלויה בתלתניות

<u>דוגמה לתלות:</u>

מטילים שתי קוביות $\Omega=\{1,\ldots,6\}^2$

A = קוביה ראשונה יצאה 4

B = הסכום הוא 6

$$P(A)=\frac{1}{6}$$

$$P(B)=\frac{5}{36}$$

$P(A\cap B)=\frac{1}{36}$ (כי הקוביה השניה יכול להיות רק 2)

מיד רואים ש $P(A\cap B)\neq P(B)\cdot P(A)$

$$P(B|A)=\frac{1}{6}$$

$$P(A|B)=\frac{1}{5}$$

<u>טענה:</u> כל מאורע הוא בלתי תלוי ב Ω כי לכל A מתקבל $P(\Omega\cap A)=P(A)=P(A)\cdot P(\Omega)$
אם משתמשים בהגדרה הכפלית – כל מאורע הוא גם בלתי תלוי בקבוצה הריקה)

<u>טענה:</u> אם B בלתי תלוי ב A אז הוא גם בלתי תלוי ב A^C
<u>הוכחה:</u>
$$P(A^C\cap B)=P(B)-P(A\cap B)=P(B)-P(A)P(B)=P(B)[1-P(A)]=P(B)P(A^C)$$
המעבר הראשון הוא כי B הוא איחוד של חיתוכו עם A ועם המשלים של A.

<u>מסקנה:</u> אם B אינו תלוי ב A אז הוא אינו תלוי בקבוצת המאורעות הבאים:
$$\{A, A^C, \Omega, \emptyset\}=\sigma(A)$$
זו תהיה הסיגמה-אלגברה שנוצרת ע"י A.
כלומר אפשר לומר (בדרך מסובכת) ש B בלתי תלוי ב A אם"ם הוא בלתי תלוי בסיגמה אלגברה ש A יוצר.

למה זה טוב?
A,B,C מאורעות
מהי המשמעות של "A אינו תלוי ב B וב C".
אפשרות: A אינו תלוי ב B וגם A אינו תלוי ב C.
נשתכנע שזו דרך לא מוצלחת להגדיר.
<u>אנטי-דוגמה</u>
מטילים שתי קוביות
A = הסכום הוא 7
B = קוביה ראשונה היא 1
C = קוביה שניה היא 1

לפי הגדרתנו A אינו תלוי ב B וב C, אבל זה טיפשי, כי ברור שאם גם B וגם C אז זה מוציא את A מכלל אפשרות.
בקיצור – לא רעיון טוב.

<u>הגדרה</u> נאמר ש **A אינו תלוי ב B וב C** אם A בלתי תלויה ב:
$$\sigma(B,C)=\{\Omega, \emptyset, B, C, B^C, C^C, B\cup C, B\cap C, B^C\cup C \ldots\}$$
משפט: A אינו תלוי ב B,C אם"ם A אינו תלוי ב B, C וב $B\cap C$

<u>שיעור שביעי – 24.11</u>

הוכחה: אנו כבר יודעים ש A אינו תלוי ב $\{\Omega, \emptyset, B, C, B^C, C^C, B\cap C, B^C\cup C^C\}$
נתבונן לדוגמה ב $B\cup C$, כלומר להראות ש $P(A\cap(B\cup C))=P(A)\cdot P(B\cup C)$ ואכן:
$$P(A\cap(B\cup C))=P((A\cap B)\cup(A\cap C))=P(A\cap B)+P(A\cap C)-P((A\cap B)\cap(A\cap C))$$
(המעבר האחרון מהכלה והדחה)
$$=P(A)P(B)+P(A)P(C)-P(A)P(B\cap C)=P(A)(P(B)+P(C)-P(B\cap C))=P(A)\cdot P(B\cap C)$$
והשאר באופן דומה.

מסקנה מיידית: באיזה מצב נוכל לומר ש A, B ו C בת"ל בזה בזה במובן ש
A בת"ל ב $\sigma(B,C)$
B בת"ל ב $\sigma(A,C)$
C בת"ל ב $\sigma(A,B)$
?
זה קורה כאשר כמובן

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

באופן כללי יותר – אם יש סדרה של מאורעות $A_1,\dots,A_n$ נאמר שכל אחד בת"ל בכל האחרים אם לכל אוסף $A_{j_1},\dots,A_{j_k}$ מתקיים $P(\bigcap_{i=1}^{k}A_{j_i})=\prod_{i=1}^{k}P(A_{j_i})$

<u>ניסויים חוזרים</u>

נתון מרחב הסתברות $(\Omega_0,\mathscr{F}_0,P_0)$ למען הפשטות, נניח שמרחב ההסתברות הוא בן מניה, כלומר $\Omega_0=\{a_1,a_2,\dots\}$

$P_0(\{a_i\})=p_i$

הניסוי "חזרנו על $(\Omega_0,\mathscr{F}_0,P_0)$ n פעמים" מתאים למרחב המדגם $\Omega=\Omega_0^n$ (אז $\mathscr{F}=P(\Omega)$)

$$P(\{a_{j_1},a_{j_2},\dots,a_{j_n}\})=p_{j_1}p_{j_2}\dots p_{j_n}$$

$$\sum_{a_{j_1},\dots,a_{j_n}\in\Omega_0}p_{j_1}\cdot\dots\cdot p_{j_n}=(\sum_{a_{j_1}\in\Omega_0}p_{j_1})(\sum_{a_{j_2}\in\Omega_0}p_{j_2})\dots(\sum_{a_{j_n}\in\Omega_0}p_{j_n})=1$$

איך נראה מאורע שמאפיין אך ורק את תוצאת הניסוי ב k?

$$\underbrace{\Omega_0\times\Omega_0\times\dots\Omega_0}_{k-1}\times\underbrace{A}_{A\in\mathscr{F}_0}\times\underbrace{\Omega_0\times\dots\times\Omega_0}_{n-k}$$

נסמן B_i = מאורע שמאפיין את תוצאת הניסוי בחזרה ה i־ית.

בהנתן מאורעות B_k,B_j כאשר $k\neq j$ אז

$$B_k=\Omega_0^{k-1}\times A_k\times\Omega_0^{n-k}$$
$$B_j=\Omega_0^{j-1}\times A_j\times\Omega_0^{n-j}$$
$$\underbrace{k>j}_{without\ \ loss\ \ of\ \ generality}$$

נסתכל בחיתוך המאורעות ונקבל:

$$B_k\cap B_j=\Omega_0^{j-1}\times A_j\times\Omega_0^{k-j-1}\times A_k\times\Omega_0^{n-k}$$

$$P(B_k)=\sum_{\omega\in B_k}P(\{\omega\})=\sum_{\substack{a_{i_1}\in\Omega_0\\ a_{i_k}\in A_k\\ a_{i_n}\in\Omega_0}}p_{i_1}\cdot\dots\cdot p_{i_n}=1\cdot1\cdot\dots\cdot P_0(A_k)\cdot1\cdot1$$

ומכאן עולה המבוקש:

$$P(B_k\cap B_j)=P(B_k)P(B_j)$$

<u>העשרה</u>

נסתכל ב $\Omega=[0,1]$

נניח שרוצים לבנות פונקציית הסתברות שהיא אינווריאנטית תחת הזזות (mod 1)

זה בלתי אפשרי, מדוע?

ניתן להציג את הקטע כאיחוד זר בן-מנייה של קבוצות שנבדלות זו מזו בהזזה $[0,1]=\coprod_{n=1}^{\infty}A_n$

מה ההסתברות של A_1 ? אם נבחר אותה להיות 0 – ההסתברות של הקטע כולו יהיה 0, אם נבחר אותה להיות מספר כלשהו, לא משנה כמה קטן, ההסתברות של הקטע כולו תהיה הסכום של ה A_n (P אינווריאנטית תחת הזזה ולכן ההסתברויות של כל ה A_n שוות) ולכן אינסופית.

מכאן – לא ניתן להגדיר פונקציית הסתברות על הקטע.

<u>דוגמה</u>

$$\Omega_0=\{0,1\}$$

$P(\{1\})=p$
$P(\{0\})=1-p=q$
ניסוי Bernoulli

נתבונן בחזרה אינסופית של חזרות בת"ל, ונשאל את עצמנו כל מיני שאלות מעניינות:
מה ההסתברות שהיתה לפחות הצלחה אחת ב k החזרות הראשונות?
מה ההסתברות שהיו בדיוק k הצלחות ב n החזרות הראשונות?
מה ההסתברות שהיו רק הצלחות?

לגבי השאלה הראשונה – המשלים הוא רצף k כשלונות, בגלל הנחת אי התלות $P(\{(0,0,\ldots,0)\})=q^k$
ומכאן שההסתברות שהיתה לפחות הצלחה אחת היא $1-q^k=1-\left(\frac{1}{1-p}\right)^k$
לגבי השאלה השניה – נסתכל במספר הקומבינציות לקבל k הצלחות, שהיא $\binom{n}{k}$ ואז נכפול הסתברויות ונקבל
$\binom{n}{k}p^k q^{n-k}$
לגבי השאלה השלישית:
A_n = הצלחה בחזרה ה n

$$P(\bigcap_{n=1}^{\infty}A_n)=P(\bigcap_{n=1}^{\infty}\bigcap_{k=1}^{n}A_n)=P(\lim_{n\to\infty}\bigcap_{k=1}^{n}A_k)=\lim_{n\to\infty}P(\bigcap_{k=1}^{n}A_k)=\lim_{n\to\infty}p^n=\begin{cases}1 & if \quad p=1\\ 0 & if \quad p<1\end{cases}$$

נמשיך בדוגמה קלאסית
<u>The gambler's ruin = *חורבנו של המהמר*</u>
לשחקן א' יש i שקלים
לשחקן ב' יש n-i שקלים

בכל צעד מטילים מטבע כך שההסתברות לראש היא p וההסתברות לזנב היא q.
אם יוצא H – שחקן ב' משלם שקל אחד לשחקן א', אחרת – להיפך.
המשחק נגמר כשאחד מהם מתרושש.

מהי ההסתברות ששחקן א' ינצח?
מרחב המדגם – הסדרות האינסופיות, מדוע? כי אנחנו לא בטוחים שהמשחק יגמר, ובכל מקרה נוכל להסתכל רק על הרישא של הסדרה ולהתעלם ממה שקרה אחרי ששחקן ב' או א' התרוששו.
$\Omega=\{H,T\}^{\mathbb{N}}$
ההסתברות לכל יחידון סופי היא מכפלת ההסתברויות לפי מה שמופיע ביחידון, כלומר:
$$P(\{H,T,H,H,\ldots,T\}\times\underbrace{\{H,T\}^{\mathbb{N}}}_{\text{tail of the sequence}})=p\cdot q\cdot p\cdot p\cdot\ldots\cdot q$$
E = השחקן הראשון ניצח
F = ההטלה הראשונה היתה $\{H\}$

צריך למצוא את $P(E)$
נשתמש בחוק ההסתברות השלמה עבור הזוג F ו F^C אשר פורשים את המרחב:
$$P(E)=P(E|F)P(F)+P(E|F^C)P(F^C)=p\,P(E|F)+q\,P(E|F^C)$$
נשתמש באינטואיציה שלנו – אחרי ההטלה הראשונה ניתן להסתכל על הבעיה מחדש, אם יצא H אז נדמיין שבמשחק החדש לשחקן א' יש כעת i+1 ולשחקן ב' יש n-i-1

נסמן α_i = ההסתברות ששחקן א' ניצח כאשר היו לו i שקלים בהתחלה

טענה:

$$\alpha_i = p\alpha_{i+1} + q\alpha_{i-1}$$

$$\alpha_0 = 0 \quad \alpha_n = 1$$

$$\alpha_{i+1} = \frac{1}{p}\cdot\alpha_i - \frac{q}{p}\cdot\alpha_{i-1}$$

$$\alpha_{i+1} - \alpha_i = (\frac{1}{p} - 1)\alpha_i - \frac{p}{q}\cdot\alpha_{i-1} = \frac{q}{p}\cdot(\alpha_i - \alpha_{i-1})$$

$$\alpha_2 - \alpha_1 = \frac{q}{p}\alpha_1$$

$$\alpha_3 - \alpha_2 = \left(\frac{q}{p}\right)^2 \alpha_1$$

$$\alpha_n - \alpha_{n-1} = \left(\frac{q}{p}\right)^n \alpha_1$$

$$\Rightarrow \quad 1 - \alpha_1 = \left(\frac{q}{p} + \frac{q^2}{p^2} + \ldots + \left(\frac{q}{p}\right)^{n-1}\right)\alpha_1$$

$$1 = \alpha_1 \cdot \left(1 + \frac{q\cdot(\frac{q}{p})^{n-1} - 1}{p\cdot\frac{q}{p} - 1}\right)$$

$$= \frac{\frac{q}{p} - 1 + \frac{q}{p}\cdot((\frac{q}{p})^{n-1} - 1)}{\frac{q}{p} - 1}$$

$$\Rightarrow \quad a_1 = \frac{\frac{q}{p} - 1}{\frac{q^n}{p^n} - 1}$$

$$\alpha_i - \alpha_1 = \left\{\frac{q}{p} + \ldots + (\frac{q}{p})^{i-1}\right\}\cdot\alpha_1$$

$$\alpha_i = \alpha_1\cdot(1 + \ldots + (\frac{q}{p})^{i-1}) = \frac{(\frac{q}{p})^i - 1}{(\frac{q}{p})^n - 1}$$

ההסתברות ש א' ניצח = $\frac{(\frac{q}{p})^i - 1}{(\frac{q}{p})^n - 1}$

ההסתברות ש ב' ניצח = $\frac{(\frac{q}{p})^{n-i} - 1}{(\frac{q}{p})^n - 1}$

ההסתברות שאף אחד לא ינצח לעולם היא אפס.

<u>**שיעור שמיני – 27.11.08**</u>

<u>אי תלות</u>
האם מאורע יכול להיות בלתי תלוי בעצמו?
האם יש מאורע שהמידע שהוא אירע (או לא) לא תשפיע על הערכתנו את ההסתברות שהוא אירע?

בעצם רק אם ההסתברות שלו היא 1 או 0, למה?

$$P(A)=P(A\cap A)=P(A)P(A)\ \Rightarrow\ P(A)^2=P(A)\ \Rightarrow\ P(A)=0,1$$

<u>חוקי 0-1</u> (Kolmogorov)
מושיבים קוף ליד מכונת כתיבה, הוא מקליד באקראי, שם לכל המקשים יש הסתברות חיובית.
מרחב המדגם הוא שרשרת אינסופית של תווים של מכונת כתיבה.
נלך לאורך השרשרת ונשאל האם יש מקטע בשרשרת שהוא בדיוק הספר "מובי דיק"?

מהי ההסתברות שיופיעו ∞ עותקים של "מובי דיק"?
רז: אם כבר מצאנו שיטה להוצאה לאור, אז למה לא

נגדיר את המאורעות הבאים:
H = הודפסו אינסוף עותקים של "מובי דיק"
H_k = הודפסו לפחות k עותקים

H_k סדרה יורדת ולכן גבולה הוא H.

$H_{m,k}$ = הודפסו לפחות k עותקים עד התו ה m.

זו סדרה עולה כפונקציה של m וגבולה $\lim_{m\to\infty} H_{m,k}=H_k$

H^m = החל מהתו ה $m+1$ הודפסו אינסוף עותקים של "מובי דיק", ברור ש $H^m=H$ כי הרישא של הסדרה לא משנה תכונות אינסופיות.
H הוא <u>**מאורע זנב**</u> (Tail event)

נקבע m ו k ונסתכל ב: $H_{m,k}\cap H^m$ אלו מאורעות בלתי תלויים כי ההקלדות השונות הן בלתי תלויות ולכן:

$$P(H_{m,k}\cap H^m)=P(H_{m,k})P(H^m)=P(H_{m,k})P(H)$$

משיקולי אריתמטיקה של גבולות, אפשר להשתכנע ש:

$$\Rightarrow\ \lim_{m\to\infty}(H_{m,k}\cap H)=H_k\cap H$$

ולכן:

$$\lim_{m\to\infty} P(H_{m,k}\cap H)=P(\lim_{m\to\infty} H_{m,k}\cap H)=P(H_k\cap H)$$

וזה גם שווה מצד שני:

$$\lim_{m\to\infty} P(H_{m,k}\cap H)=\lim_{m\to\infty} P(H_{m,k})P(H)=P(\lim_{m\to\infty} H_{m,k})P(H)=P(H_k)P(H)$$

$$P(H)=P(H_k\cap H)=P(H_k)P(H)\ \underset{k\to\infty}{\rightarrow} P(H)=P(H)^2\ \Rightarrow\ P(H)=0,1$$

רז: אפשר לשאול כמה זמן זה יקח, אבל זה לא... זה לא... זה יקח הרבה זמן

<u>המשתנה המקרי</u>

<u>דוגמה</u>

מטילים שתי קוביות, $\Omega=\{1,\dots,6\}^2$

$\mathscr{F}=P(\Omega)$

$P(\{\omega\})=\frac{1}{36}$

נוכל להסתכל על פונקציה מעל Ω כמו למשל – הסכום, או – זוגיות הקוביה הכחולה, בהסתברות יש שם fancy לפונקציה – המשתנה המקרי.

<u>הגדרה זמנית</u>

בהנתן מרחב הסתברות – משתנה מקרי (ממשי) הוא פונקציה $\Omega\to\mathbb{R}$

נחזור לדוגמה שלנו עם הקוביות:

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

$X:\Omega\to\underbrace{\{2,\dots,12\}}_{S\subseteq\mathbb{R}}\qquad X((i,j))=i+j$

מרגע שהגדרנו פונקציה, אולי נוכל להגדיר פונקציית הסתברות ל S, אשר תהיה תלויה באופן כלשהו בפונקציית ההסתברות שכבר הגדרנו במרחב.

נשאל $P(A)=?$ כאשר $A\subseteq S$

זה כמובן חסר משמעות כי P מוגדרת על מרחב המדגם ולא על מרחב הסכומים, נסמן את הפונקציה המבוקשת שלנו P_X ונבקש שהיא תשקף את ההסתברות לכך שנקבל את הסכום, כלומר:

$$P_X(A)=P(\{\omega\in\Omega:X(\omega)\in A\})=P(\underbrace{X^{-1}(A)}_{\text{אוסף מקורות ולא פונקציה הופכית}})$$

לפונקציה P_X נקרא ה**<u>התפלגות</u>** של המ"מ (=משתנה מקרי)

S היא קבוצה שגם עליה אפשר לבנות סיגמה אלגברה, נסמנה $\mathscr{F}_S$ אז אפשר לחשוב על X כהעתקה לא של קבוצה לקבוצה אלא של מרחב מדיד למרחב מדיד.

נסתכל על מאורע $A\in\mathscr{F}_S$

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

נשאל מה ההסתברות של $X^{-1}(A)$, אם כי לא מובטח שזה מאורע.

<u>משתנה מקרי</u>

בהנתן שני מרחבים מדידים $(\Omega,\mathscr{F}),(S,\mathscr{F}_S)$ ופונקציה $X:\Omega\to S$ יקרא **משתנה מקרי** אם לכל $A\in\mathscr{F}_S$,

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

הערה לתשומת לב: $X:\Omega\to S, X^{-1}:\mathscr{F}_S\to\mathscr{F}$

<u>טענה</u>: X^{-1} מתחלפת עם כל הפעולות על קבוצות, דהיינו:

$$X^{-1}(\bigcap_{n=1}^{\infty}A_n)=\bigcap_{n=1}^{\infty}X^{-1}(A_n)\qquad X^{-1}(A^C)=(X^{-1}(A))^C$$

וגם A,B זרים אז $X^{-1}(A),X^{-1}(B)$ זרים

<u>הוכחה</u>

$$X^{-1}(A_i)=\{\omega:X(\omega)\in A^C\}=\{\omega:X(\omega)\notin A\}=\{\omega:X(\omega)\in A\}^C=(X^{-1}(A))^C$$

זה כמעט טאוטולוגיה וכך מוכיחים גם את השאר.

<u>שיעור 9, 1.12.08</u>

דוגמה למשתנה מקרי:

יהא $(\Omega, \mathscr{F}, P)$ מרחב הסתברות $A\in\mathscr{F}$

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

$$I_A:\Omega\to\{0,1\}=S$$

I_A היא הפונקציה המציינת של המאורע A.

אזי הסיגמה אלגברה של S היא $2^S=\{\emptyset,\}0\{,\{1\},\{0,1\}\}$

$$X^{-1}(\{1\})=A \quad , \quad X^{-1}(\{0\})=A^C$$

$$P_X(\{1\})=P(\{\omega\in\Omega:X(\omega)\in\{1\}\})=P(X^{-1}(\{1\}))=P(A)$$

<u>משפט:</u> יהי $X:(\Omega,\mathscr{F},P)\to(S,\mathscr{F},P_X)$ מ"מ מעל מרחב הסתברות $(\Omega,\mathscr{F},P)$ אז ההתפלגות של X היא פונקצית הסתברות מעל $\mathscr{F}_S$

<u>הוכחה:</u>

לכל $A\in\mathscr{F}_S$

$P_X(A)=P(X^{-1}(A))\in[0,1]$ (תכונה 1 של פונקציית הסתברות)

$P_X(S)=P(X^{-1}(S))=P(\Omega)=1$ (תכונה 2)

תהי סדרת מאורעות זרים $A_1, A_2, \ldots\in\mathscr{F}_S$

$X^{-1}(A_1), X^{-1}(A_2), \ldots\in\mathscr{F}$ קבוצות זרות

$$X^{-1}(\coprod_{i=1}^{\infty}A_i)=\coprod_{i=1}^{\infty}X^{-1}(A_i)$$

$P_X(\coprod_i A_i)=P(X^{-1}(\coprod A_i))=P(\coprod X^{-1}(A_i))$ (תכונה 3 – אדיטיביות בת מניה)

❑

S קבוצה בת מניה

P_X התפלגות מוגדרת ע"י ערכה עבור כל היחידונים.

אבל $P(\{\omega\in\Omega:X(\omega)=k\})=P_X(\{k\}) \quad k\in S$

לשם הנוחות נסמן $p_x(k)=P_X(\{k\})$

מכיוון שזו ההתפלגות כשהיא מוערכת ביחידון – נכנה אותה **התפלגות נקודתית של משתנה מקרי X**.

יש בזה טעם רק כאשר מרחב ההסתברות הוא בן מניה, במקרים אחרים – נשתמש במשהו אחר.

נחדד ונזכיר:

$$P:\mathscr{F}\to[0,1]$$

$$P_X:\mathscr{F}_S\to[0,1]$$

$$p_x:S\to[0,1]$$

דוגמה:

נתונים 20 כדורים ממוספרים, מוציאים שלושה, מה ההסתברות שלפחות אחד מהשלושה הוא 17 או יותר?

$$\Omega=\{(i,j,k):1\le i<j<k\le 20\}$$

הסתברות אחידה – יש 18 דרכים לבחור את i, ואז 19-i דרכים לבחור את j וכן הלאה.

נגדיר משתנה מקרי שיתן את הכדור הגדול בכל שלשה:

$$X:\Omega\to\{3,\ldots,20\}$$

$$X((i,j,k))=k$$

נתבקשנו לענות על השאלה:

$$P(\{(i,j,k)\in\Omega:k\ge 17\})=P(\{(i,j,k)\in\Omega:X(\omega)\ge 17\})=P_X(\{17,18,19,20\})$$
$$=p_X(17)+p_X(18)+p_X(19)+p_X(20)$$

ולאחר כל המשחקים האלו (שבהם לא ביצענו שום פעולה אלא רק הצגנו את הבעיה בכמה דרכים) נחזור לשאלה מה ההסתברות שהגדול ביותר הוא k?

$$p_X(k)=P(X=k)=\frac{\binom{k-1}{2}}{\binom{20}{3}}$$

ונשוב לעוד דוגמה קלאסית:

<u>בעיית אספן הקופונים (coupon collector)</u>

יש N קופונים שונים

כשנלך לחנות אנחנו מקבלים באקראי אחד מהקופונים הללו, כשאנחנו מגיעים הביתה, בודקים אם כבר יש לנו אותו, אם לא – אנחנו מדביקים אותו באלבום.

נשאלת השאלה: כמה קופונים עלינו לקנות עד להשלמת הרצף???

$\Omega=\{1,\ldots,N\}^{\mathbb{N}}$ = סדרות אינסופיות בתחום המספרים הטבעיים הקטנים או שווים ל N

ניקח את כל המאורעות המתייחסים למספר סופי של קופונים, כלומר כל אלו שמספיק להסתכל על רישא של הסדרות האינסופיות כדי לוודא את קיום האירוע, למשל – המאורע "הקופון הראשון הוא 5", כלומר:

$\mathscr{F}$ = כל תת הקבוצות של Ω המתקבלות על ידי מספר בן מניה של פעולות על תת קבוצות המתייחסות למספר סופי של קופונים.

למשל :

$$A_{i,j}=\{(a_1,a_2,\ldots):a_i=j\}$$

פונקציית ההסתברות היא כזאת שכל המאורעות האפשריים המתארים קומבינציה של k הקופונים הראשונים הם שווי הסתברות.

T מספר הקופונים שנאסף עד קבלת אוסף מלא.

$$T(a_1,a_2,\ldots)=\min_k\{\forall 1\le j\le N \quad j\in\{a_1,\ldots,a_k\}\}$$

$$T:\Omega\to\{N,N+1,\ldots\}\cup\{\infty\}$$

שאלתנו היא : עבור $N\le n$ מה ההסתברות ש $T(\omega)=n$ ($\omega\in\Omega$)

$$P_T(\{n\})=p_T(n)$$

$P_T(\{n+1,n+2,\ldots\})$ = ההסתברות שלא היה אוסף מלא לאחר n צעדים

A_j = לאחר n צעדים הקופון ה j היה חסר.

$$T>n$$

$$P_T(\{n+1,n+2,\ldots\})=P(\bigcup_{j=1}^{N}A_j)=\sum_{j=1}^{N}P(A_j)-\sum_{i<j}P(A_i\cap A_j)+\sum_{i<j<k}P(A_i\cap A_j\cap A_k)-\ldots$$

$$=N\cdot\frac{(N-1)^n}{N^n}-\binom{N}{2}\cdot\frac{(N-2)^n}{N^n}+\ldots \;\Rightarrow\; P_T(\{n+1,n+2,\ldots\})=\sum_{k=1}^{N}\binom{N}{k}(-1)^{k+1}\left(\frac{N-k}{N}\right)^n$$

$$p_T(n)=P_T(\{n\})=P_T(\{n,n+1,\ldots\}\setminus\{n+1,n+2,\ldots\})$$

$$=P_T(\{n,n+1,\ldots\})-P_T(\{n+1,n+2,\ldots\})=\sum_{k=1}^{N}\binom{N}{k}(-1)^{k+1}\left(\frac{N-k}{N}\right)^{n-1}\cdot\frac{k}{N}$$

$$\sum_{n=1}^{\infty}p_T(n)=1$$

$$\Rightarrow\; p_T(\infty)=0$$

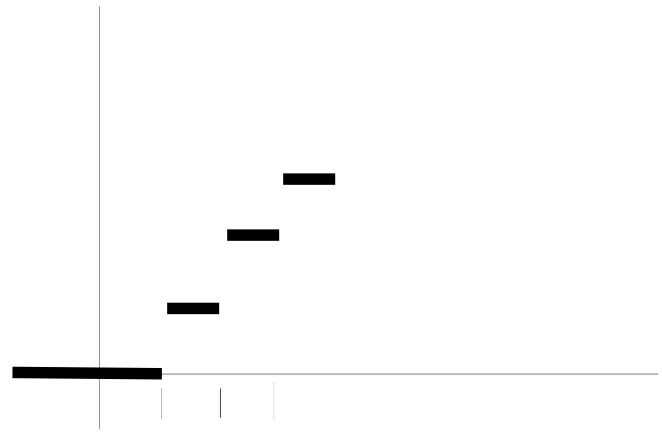
הגדרה: X משתנה מקרי מעל $(\Omega,\mathscr{F},P)$

לכל $a\in\mathbb{R}$ נגדיר את **פונקציית ההתפלגות של X** ע"י:

$$F_X(a)=P(\{\omega\in\Omega:X(\omega)\le a\})=p_X((-\infty,a])$$

$$F_x:\mathbb{R}\to[0,1]$$

דוגמה: הטלת שתי קוביות, X הסכום, אז F_X תראה כך:



טענה: כל פונקציית התפלגות מקיימת :

1. F_X אינה יורדת

.2 $\lim_{a\to-\infty} F_X(a)=0$

.3 $\lim_{a\to\infty} F_X(a)=1$

4. F_X רציפה מימין

<u>הוכחה</u>

1) $a>b$

$$F_X(a)=P_X((-\infty,a])=P_X((-\infty,b]\uplus[b,a])=F_X(b)+\underbrace{P_X((b,a])}_{\geq 0}\geq F_X(b)$$

(2

$$\lim_{a\to\infty} F_X(a)=\lim_{a\to-\infty} P_X((-\infty,a])=P_X(\lim_{a\to-\infty}(-\infty,a])=P_X(\emptyset)=0$$

3 – באותו אופן

(4

$$\lim_{n\to\infty} F_X(a+\frac{1}{n})=\lim_{n\to\infty} P_X(\left(-\infty,a+\frac{1}{n}\right])=P_X(\lim_{n\to\infty}\left(-\infty,a+\frac{1}{n}\right])=P_X((-\infty,a])=F_X(a)$$

טענה: F_X מגדירה את P_X

$$P_X((a,b])=P_X((-\infty,b]\setminus(-\infty,a])=F_X(b)-F_X(a)$$

$$P_X((a,b))=P_X(\lim_{n\to\infty}\left(a,b+\frac{1}{n}\right])=\lim_{n\to\infty}\left[F_X(b-\frac{1}{n})-F_X(a)\right]$$

<u>**שיעור עשירי – 4.12.08**</u>

ניגע היום במשתנים מקריים מסויימים – אלה שנתקלים בהם מספיק פעמים בחיים כדי שיהיה להם שם משלהם.

הגדרה

משתנה מקרי נקרא **משתנה ברנולי (Bernoulli)** אם טווח הערכים שלו הוא $\{0,1\}$
ההסתברות שלו מוגדרת באופן יחיד ע"י ההתפלגות של אחד היחידונים, נאמר $P_X(\{1\})=p$ כי ההסתברויות האחרות תהיינה:

$$P_X(\{0,1\})=1 \quad P_X(\{0\})=1-p \quad P_X\{\emptyset\}=0$$

הדגשנו בפעמים הקודמות את החשיבות שבבניית מרחב הסתברות, אבל הפעם ההגדרה של משתנה ברנולי בעצם בונה מרחב הסתברות ע"י $(\{0,1\}, 2^{\{0,1\}}, P_X)$

הגדרה:

תהליך ברנולי הוא ניסוי המורכב מחזרות בלתי תלויות של n ניסויים בינאריים (=ברנוליים...).
כלומר אם:

$$\Omega=\{0,1\}^n \quad n\in\mathbb{N}$$

כך שבכל תת ניסוי מהניסויים הללו ההסתברות ל"הצלחה" (= $\{1\}$) הוא p, כלומר שווה בכולם.
כל הסתברות של יחידון תהיה $P(\{a_1, a_2, \dots, a_n\})=p^{|\text{הצלחות}|}\cdot(1-p)^{|\text{כשלונות}|}$

נגדיר משתנה מקרי X להיות $|\text{הצלחות}|$ (=מספר ההצלחות)
נתהה מהי ההתפלגות של המשתנה המקרי?
הטווח של X הוא $S=\{0,1,\dots,n\}$
ההתפלגות:

$$P_X(\{k\})=p_X(k)=\binom{n}{k}p^k(1-p)^{n-k} \quad (***)$$

הגדרה

משתנה מקרי יקרא **משתנה בינומי עם פרמטרים (n,p)** אם טווח הערכים שלו הוא S וההתפלגות שלו (***).
נסמן $X\sim B(n,p)$
כדי לומר ש X מתפלג כמו משתנה בינומי עם פרמטרים אלו.

תרגיל

יצרנית של חיתולים מייצרת חיתול פגום בהסתברות 0.01
כל שני חיתולים הם בלתי תלויים...
החברה מוכרת חבילות של עשרה חיתולים ומחזירה ללקוח את כספו רק אם יותר מחיתול אחד פגום.
מה ההסתברות שלקוח שקנה חבילה יחזיר אותה?
רז: התשובה היא 0 כמובן כי יאשימו אותו שהוא זה שהרס את החיתול הפגום אבל בואו נהיה אופטימיים ונענה על השאלה.

$X\sim B(10,0.99)$ מספר החיתולים התקינים

$$\Omega=\{0,\dots,10\}$$

$$p_X(k)=P_X(\{k\})=\binom{10}{k}0.99^k\cdot 0.01^{10-k}$$

תשובה = $1-p_x(10)-p_x(9)$ = בערך 0.07

תרגיל קצת מרתיע

מנוע של מטוס שובק במהלך טיסה בהסתברות $1-p$
מטוס נוחת בשלום רק אם לפחות מחצית ממנועיו תקינים, מה עדיף – מטוס דו מנועי או 4-מנועי?

$X_1\sim B(2,p)$ מספר המנועים התקינים

$X_2\sim B(4,p)$ כנ"ל

A_1 = יצאנו בשלום עם מטוס דו מנועי

A_2 = יצאנו בשלום עם מטוס 4-מנוע

כדי שהעסק הזה יראה מתמטית כמו שצריך, אז אפשר לבנות שני מרחבי הסתברות שונים עבור כל אחד מהמקרים, ואפשר להסתכל על מכפלה שלהם. אבל אפשר פשוט לשאול כל שאלה בנפרד ואז:

$$A_1=\{1,2\}$$

$$A_2=\{2,3,4\}$$

$$P_{X_1}(\{1,2\})=p_{X_1}(1)+p_{X_1}(2)=2\mathrm{p}^1(1-p)^1+p^2(1-p)^0=2p-p^2$$

$$P_{X_2}(\{2,3,4\})=6p^2(1-p)^2+4p^3(1-p)^1+p^4(1-p)^0=3p^4-8p^3+6p^2\geq 2p-p^2$$

כלומר יש להעדיף מטוס 4-מנועי רק אם $p>\frac{2}{3}$

<u>שאלה</u>

מטילים מטבע הוגן 100 פעמים, מה ההסתברות שיצא H בדיוק 50 פעם.

$2n$ הטלות, אז המשתנה המקרי $X\sim B(2n,\frac{1}{2})$

$$P_X(\{n\})=p_X(n)=\binom{2n}{n}\frac{1}{2}^n\cdot\frac{1}{2}^{2n-n}=\frac{(2n)!}{(n!)^2 2^{2n}}$$

ניעזר בקירוב סטרלינג.

$$n!\sim\sqrt{2\pi}\,n^{n+\frac{1}{2}}e^{-n}$$

ולכן נקבל:

$$\frac{\sqrt{2\pi}(2n)^{2n}\sqrt{2n}e^{-2n}}{(\sqrt{2\pi}n^n\sqrt{n}e^{-n})^2\cdot 2^{2n}}=\ldots=\frac{\sqrt{2n}}{\sqrt{2\pi n^2}}=\frac{1}{\sqrt{\pi n}}$$

ועבור המקרה שלנו זה יצא בערך 0.08

<u>טענה:</u>

ההתפלגות הנקודתית של משתנה בינומי עולה עד הערך $k=\lfloor(n+1)p\rfloor$ ויורדת לאחר מכן:

<u>הוכחה</u>

נסתכל ב $\frac{p_x(k)}{p_x(k-1)}$ ונקבל:

$$\frac{\binom{n}{k}p^k(1-p)^{n-k}}{\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1}}=\frac{(k-1)!(n-k+1)!\,p}{k!(n-k)!(1-p)}=\frac{(n-k+1)p}{k\cdot(1-p)}$$

הפונקציה עולה אם $(n-k+1)p\geq k(1-p)$

כלומר $k<p(n+1)$ כנדרש.

❑

<u>המשך שעורים עם רז קופרמן בהסתברות (1)</u>

<u>שיעור 11, 8.12.08</u>

משתנים מקריים $X:(\Omega,\mathscr{F},P)\rightarrow(S,\mathscr{F}_S,P_X)$

S מרחב בדיד

P_X בתפלגות

p_X התפלגות נקודתית

<u>הגדרה:</u> משתנה מקרי X יקרא **בעל התפלגות** Poisson עם פרמטר λ (נסמן $X\sim Poi(\lambda)$)
אם הוא מקבל ערכים בקבוצה $\{0,1,2,...\}$ והתפלגותו הנקודתית היא:

$$p_x(k)=\frac{e^{-\lambda}}{k!}\cdot\lambda^k$$

אפשר לראות שזה אי שלילי לכל k ושהטור על הטבעיים עם אפס נותן 1.
אבל למה להגדיר דבר כזה?
למשל התפרקות רדיואקטיבית.
אם אקח חומר רדיואקטיבי (רז: לא לפחד) ואשים לידו מונה גייגר שסופר כמה פגיעות בשניה היו, זה משתנה מקרי שאני לא יכול לחזות מראש. אבל אם אבנה היסטוגרמה של שכיחות מספר הפגיעות, ההיסטוגרמה תתכנס למשהו שיהיה התפלגות פואסונית, כאשר λ תלוי בכל מיני גורמים פיסיקליים שאינם ממין העניין כרגע.
למה זה קורה?

נניח שאני מונה פגיעות לאורך ציר הזמן:

0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 1

כאשר השנתות בדרך הן $\Delta t=\frac{1}{n}$

בכל מקטע זמן Δt נוסף למונה אפס או אחד.
ההסתברות שיתווסף 1 במקטע Δt היא $\lambda\Delta t$
ההסתברות לתוספת בכל מקטע זמן בלתי תלוייה במקטעי הזמן האחרים.
X = מספר הפגיעות ביחידת זמן.
$X\sim B(n,\frac{\lambda}{n})$ (נקבע את n להיות מספיק גדול כך שזה יתקיים, כלומר תהיה פגיעה אחת לכל היותר בכל יחידת זמן)
ואז:

$$p_X(k)=\binom{n}{k}\left(\frac{\lambda}{k}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}=\frac{n!}{(n-k)!\,k!}\cdot\frac{\lambda^k}{n^k}\cdot\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}=\frac{\lambda^k}{k!}\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{n^k}\cdot\frac{1}{\left(1-\frac{\lambda}{n}\right)^k}$$

$$=\frac{\lambda^k}{k!}\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\frac{1\cdot(1-\frac{1}{n})(1-\frac{2}{n})\cdot\ldots\cdot(1-\frac{k-1}{n})}{(1-\frac{\lambda}{n})^k}\ \Rightarrow\ \underset{n\to\infty}{\rightarrow}\frac{\lambda^k}{k!}\cdot e^{-\lambda}$$

<u>דוגמה:</u>
מספר שגיאות הדפוס בעמוד מרשימותיו של רז (2006) הוא משתנה $Poi\left(\frac{1}{2}\right)$
מה ההסתברות שיש לפחות שגיאה אחת בעמוד נתון?
נשים לב שברגע שמדברים על משתנה פואסוני, אנחנו מדברים על מרחב הסתברות שהוא הטבעיים ואפס.
ולכן השאלה שנשאל:

$$S=\{0,1,\ldots\}$$

$$P_X(S\setminus\{0\})=1-p_X(0)=1-e^{\frac{-1}{2}}$$

מוטיבציה להגדרת **<u>משתנה מקרי גיאומטרי</u>** -

נתונה סדרה אינסופית של ניסויי ברנולי בלתי תלויים עם פרמטר P.

X = מספר החזרות עד להצלחה ראשונה

$$S=\{1,2,...\}=\mathbb{N}$$

מה ההסתברות שצריך בדיוק k הטלות? הסיפור הוא ש k-1 פעמים נכשלנו ובפעם ה k הצלחנו לראשונה, ואת ההסתברות של זה אנחנו מכירים: $p_X(k)=(1-p)^{k-1}p$

ואם אנחנו לא מאמינים שזו התפלגות, נסכום:

$$\sum_{k=1}^{\infty} p_X(k)=p\cdot\sum_{k=1}^{\infty}(1-p)^{k-1}=p\sum_{k=0}^{\infty}(1-p)^k=1$$

ומכיוון שזה נהיה טור גיאומטרי, X כנ"ל יהיה **<u>משתנה מקרי גיאומטרי</u>** (כלומר שהתפלגותו כנ"ל והוא מקבל ערכים טבעיים)

נסמן $X\sim Geo(p)$

הערה: בספרות לעיתים יופיע k ולא k-1 כי השאלה תהיה "מה ההסתברות שהיו k כשלונות ואז הצלחה"

<u>תרגיל:</u>

N כדורים לבנים

M כדורים שחורים

אנחנו מוציאים כדור באקראי, אם הוא שחור – ניצחנו, אם הוא לבן – מחזירים אותו פנימה, בוחשים ומוציאים כדור אחר.

1) מה ההסתברות שנזקקנו ל k הוצאות?
2) מה ההסתברות שנזקקנו לפחות ל n הוצאות?

1) זה די מיידי, כי זו סדרת ניסויי ברנולי, שבה ההסתברות להוצאת שחור (=הצלחה) היא $\frac{M}{M+N}$

ולכן:

$$p_X(k)=\left(1-\frac{M}{M+N}\right)^{k-1}\frac{M}{M+N}$$

2)

$$\sum_{k=n}^{\infty} p_X(k)=\frac{\sum_{k=n}^{\infty} M}{M+N}\cdot\left(1-\frac{M}{M+N}\right)^{k-1}=\frac{M}{M+N}\cdot\frac{\left(1-\frac{M}{M+N}\right)^{n-1}}{1-\left(1-\frac{M}{M+N}\right)}=(1-\frac{n}{M+N})^{n-1}$$

תכונת **<u>היעדר זיכרון</u>** של המשתנה המקרי הגיאומטרי:

$$X\sim Geo(p)$$

אם n>k אז:

$$P(X=n|X>k)=p_X(n-k)$$

הוכחה:

קודם נסתכל מה אנחנו בעצם מנסים להוכיח לפי הגדרת ההסתברות המותנה:

$$\frac{P(X=n\wedge X>k)}{P(X>k)}=\frac{p_X(n)}{p_X(\{k+1,...\})}=\frac{(1-p)^{n-1}\cdot p}{(1-p)^k}=(1-p)^{n-k-1}p$$

סדרת ניסויי ברנולי עם פרמטר p.

X = מספר הכשלונות עד קבלת n הצלחות

$$p_X(k)=\binom{n+k-1}{k}p^n(1-p)^k=\frac{(n+k-1)!}{k!(n-1)!}p^n(1-p)^k$$

מסתבר שאם נרצה להרחיב את זה לא רק לטבעיים, נוכל להשתמש בפונקציית גאמה, הלא היא הרחבה של פונקציית העצרת:

$$\Gamma(x)=\int_0^\infty e^{-t}t^{x-1}\,dt \quad \Gamma(n)=(n-1)! \quad (n,p)$$

<u>הגדרה:</u>

X נקרא **<u>מתפלג בינומי שלילי</u>** עם פרמטרים (r,p) אם:

$$p_X(k)=\frac{\Gamma(r+k)}{k!\,\Gamma(r)}\,p^r(1-p)^k$$

שעשוע – ניתן הוכחה הסתברותית למשהו מתורת המספרים.

כידוע מאינפי 1 – הטור ההרמוני לא מתכנס, ואילו ריבועו מתכנס.

נסתכל בטור $\sum_{p\ is\ prime}\frac{1}{p}=\infty$

יהא S>1 ונתבונן במשתנה מקרי X עם ערכים טבעיים והתפלגות נקודתית:

$$\zeta(S)=\sum_{n=1}^{\infty}n^{-S} \qquad p_X(k)=\frac{k^{-S}}{\zeta(s)}$$

(זוהי פונקציית זטא של רימן)

תהי A_m = קבוצת הכפולות של m

אזי:

$$P_X(A_m)=\sum_{m|k}\frac{k^{-S}}{\zeta(S)}=\sum_{k=1}^{\infty}\frac{(km)^{-S}}{\zeta(S)}=m^{-S}$$

נטען: אם q, p ראשוניים אז A_p ו A_q בלתי תלויות.

צ"ל:

$$P_X(\underbrace{A_q\cap A_p}_{A_{p\times q}})=P_X(A_p)P_X(A_q)$$

ואגף ימין הוא אכן:

$$p^{-S}\cdot q^{-S}=(pq)^{-S}$$

ולכן אי התלות מתקיימת.

נזכיר ש A_p^C = אוסף המספרים שאינם כפולה של p

אם נחתוך את כל אלו עבור p ראשוני נקבל 1, כי רק הוא אינו כפולה של אף מספר ראשוני, כלומר $\bigcap_{p\ is\ prime}A_p^C=\{1\}$

ומכאן:

$$p_X(1)=P_X(\{1\})=P_X(\bigcap_{p\ is\ prime}A_p^C)=\prod_{p\ is\ prime}P_X(A_p^C)=\prod_{p\ is\ prime}\left(1-\frac{1}{p^S}\right)=\frac{1}{\zeta(S)}$$

והדבר האחרון שקיבלנו נקרא – נוסחת אויילר, שמבטאת זיקה (שהיא לא מיידית מהסתכלות על הפונקציה) בין פונקציית זטא של רימן לראשוניים.

פונקציית זטא שואפת לאינסוף באחד ואז ההופכי ישאף לאפס באחד ולכן:

$$\lim_{S\to 1}\prod_{p\ is\ prime}\left(1-\frac{1}{p^S}\right)=0$$

$$\lim_{S\to 1}\sum_{p\ is\ prime}\log\left(1-\underbrace{\frac{1}{p}}_{x}\right)=-\infty$$

ולכל $0<x<0.6$ מתקיים $\log(1-x)\geq -2x$ ואז:

$$\sum_{p\ is\ prime}\log\left(1-\frac{1}{p}\right)\geq -2\sum_{p\ is\ prime}\frac{1}{p}=\text{“}\infty\text{”}$$

<u>שני משתנים מקריים</u>

נגדיר שני משתנים מקריים

$$X:(\Omega,\mathscr{F},P)\to(S_X,\mathscr{F}_X,P_X)$$
$$Y:(\Omega,\mathscr{F},P)\to(S_Y,\mathscr{F}_Y,P_Y)$$

נניח:

$$B\in\mathscr{F}_Y \quad A\in\mathscr{F}_X$$

מה ההסתברות ש $X\in A$, $Y\in B$

נניח שאני יודע רק את P_X, P_Y אבל לא את P.

כי כמובן שזה $P(\{\omega\in\Omega: X(\omega)\in A, Y(\omega)\in B\})$

הנקודה היא שלא צריך לחשוב על X ועל Y כפונקציות, אלא כהעתקה:

$$(X,Y):(\Omega,\mathscr{F},P)\to(S_X\times S_Y,\mathscr{F}_{X,Y},P_{X,Y})$$

כרגיל נדרוש מהסיגמה אלגברה שהגדרנו שהמקור שלה ע"י ההעתקה יהיה ב $\mathscr{F}$.

ומהי $P_{X,Y}$?

בהנתן $\mathscr{F}_{X,Y}\ni A\subset S_X\times S_Y$ נגדיר:

$$P_{X,Y}(A)=P(\{\omega\in\Omega:(X(\omega),Y(\omega))\in A\})$$

ואז היא תקרא ההתפלגות המשותפת של המשתנים המקריים X ו Y.

$B\in\mathscr{F}_X$ אז:

$$P_{X,Y}(B\times S_Y)=P(\{\omega: X(\omega)\in B, Y(\omega)\in S_Y\})=P_X(B)$$

ובאותו אופן:

$$C\in\mathscr{F}_Y \;\Rightarrow\; P_{X,Y}(S_X\times C)=P_Y(C)$$

ואז:

$$p_{X,Y}(x,y)=P_{X,Y}(\{x,y\}) \quad x\in S_X, y\in S_Y$$
$$p_X(x)=\sum_{y\in S_Y} p_{X,Y}(x,y) \quad p_Y(y)=\sum_{x\in S_X} p_{X,Y}(x,y)$$

פונקציית התפלגות משותפת:

$$F_{X,Y}(a,b)=P_{X,Y}((-\infty,a],(-\infty,b])=P(\{\omega: X(\omega)\le a, Y(\omega)\le b\})$$

דוגמה:

בכד 3 כדורים אדומים

4 לבנים

5 כחולים

מוציאים שלושה באקראי.

X מספר האדומים בהוצאה

Y מספר הלבנים בהוצאה

כל אחד מהם כשלעצמו הוא משתנה מקרי שאנחנו יודעים לחשב את ההתפלגות שלו.

סוג השאלות עבור ההתפלגות המשותפת יהיה – מה ההסתברות ש X = 1 וגם Y=0

כלומר הזוגות הסדורים כך:

$$(X,Y)\in\{(i,j): i,j\ge 0, i+j\le 3\}$$
$$S_X=\{0,1,2,3\}=S_Y$$

אפשר להסתכל בקבוצת המכפלה של המרחבים ולתת הסתברות אפס לכל הדברים ש"לא יתכנו", כלומר שמספר הכדורים הכולל עולה על 3.

נשאל אם כן $P_{X,Y}(i,j)=?$
ניתן להציג זאת כטבלה:

i \ j	**0**	**1**	**2**	**3**
0				
1				-
2			-	-
3		-	-	-

למשל:

$$P_{X,Y}(0,0)=\frac{\binom{5}{3}}{\binom{12}{3}}$$

ואז הסכומים על העמודות יתנו את ההתפלגות של Y וההתפלגות על השורות יתנו את ההתפלגות של X, כי למשל כל השורה הראשונה היא אוסף כל המקרים X=0 וכו'.

לסיכום: ברור שאין שום דבר קדוש במספר 2, ויכול להיות כל מספר שהוא של משתנים מקריים, והדרך לחשוב עליהם היא לא כ n העתקות אלא כהעתקה אחת לעולם של וקטורים n מימדיים.

<u>שיעור 12 – 11.12.08</u>
לרז נגמר הקול – ננסה להבין מה הוא אומר למרות שזה יחסית בשקט.

נמשיך עם ריבוי משתנים מקריים – למשל n.
ואז ההעתקה שלנו תהיה $\mathscr{F}\to S_1\times\ldots\times S_n$
וההתפלגות המשותפת ה n מימדית תהיה:

$$P_{x_1,\ldots,x_n}(A)=P(\{\omega\in\Omega:(X_1(\omega),\ldots,X_n(\omega))\in A\})$$

אז מה עם $B\subseteq S_1\times\ldots\times S_{n-1}$?

$$P_{x_1,\ldots,x_{n-1}}(B)=P(\{\omega\in\Omega:(X_1(\omega),\ldots,X_{n-1}(\omega))\in B\})$$
$$=P(\{\omega\in\Omega:(X_1(\omega),\ldots,X_n(\omega))\in B\times S_n\})=P_{x_1,\ldots,x_n}(B)$$

בהנתן התפלגות n מימדית, ניתן ליצור ממנה n התפלגויות חד ממדיות אבל ההיפך אינו נכון.

נתבונן במשתנה מקרי $X:\Omega\to S_x$ (בעצם זה לא ממש מאומגה אלא מהסיגמה-אלגברה המקורית ל $\mathscr{F}_X$)

לכל $A\in\mathscr{F}_X$ יש $X^{-1}(A)\in\mathscr{F}$
ואז:

$$\sigma(X)=\{X^{-1}(A):A\in\mathscr{F}_X\}\subseteq\mathscr{F}$$

<u>טענה:</u>
$\sigma(X)$ היא תת- σ -אלגברה של $\mathscr{F}$, תת קבוצה ממש.
<u>דוגמה:</u>
אם נסתכל בסכום שתי הטלות של קוביה בתור ה $\sigma(X)$ שלנו – זוהי קבוצה שהמקורות שלה הם סכומים, כלומר מתארת מאורעות שהמשותף להם הוא הסכום, ולכן אם $(4,3)\in B\in\sigma(X)$ אז גם $(5,2)\in B$ ולכן היחידון $\{(4,3)\}$ אינו נמצא ב $\sigma(X)$ כי כל קבוצה שמכילה אותו – צריכה להכיל את כל המאורעות שנותנים את אותו סכום.

שני משתנים מקריים X,Y

כל אחד מגדיר σ -אלגברה: $\sigma(X), \sigma(Y)$
X ו Y יקראו בלתי תלויים, אם כל זוג מאורעות $B\in\sigma(Y), A\in\sigma(X)$ בלתי תלויים.

<u>דוגמה:</u>
$\Omega=\{1,\dots,6\}^2$ דהיינו מטילים קוביה פעמיים
$X(i,j)=i+j$
$Y(i,j)=i$

ראינו כבר שהמאורע $X^{-1}(\{7\})$ (יצא סכום שבע) בלתי תלוי במאורע $Y^{-1}(\{2\})$ (הקוביה הראשונה היא 2)

אבל ראינו גם ש $X^{-1}(\{6\})$ ו $Y^{-1}(\{2\})$ הם תלויים וזה מספיק כדי להכריע ש X ו Y הם תלויים.
בעצם זה ברור – כי למשל הידיעה שהסכום הוא 12 משנה את הערכתנו לגבי "האם הקוביה הראשונה יצאה 2".

בשלב זה נראה שיהיה לא פשוט לגלות אם X ו Y הם תלויים או לא, כי בעצם נצטרך לבדוק את כל האפשרויות...
התשובה היא שקל מאד לקבל קריטריון שיכריע בדבר.

נתבונן בהתפלגות המשותפת של X ו Y על ידי: $P_{X,Y}(A)=P(\{\omega:(X(\omega),Y(\omega))\in A\})$
$S_X\times S_Y\supseteq A\in\mathscr{F}_{X,Y}$
$A=B\times C$
$B\in\mathscr{F}_X, C\in\mathscr{F}_Y$

אז ההתפלגות המשותפת תהיה:
$$P_{X,Y}(B\times C)=P(x\in B, y\in C)=P(\{\omega\in\Omega: X(\omega)\in B, Y(\omega)\in C\})$$
$$=P(\{\omega\in\Omega: X(\omega)\in B\}\cap\{\omega\in\Omega: Y(\omega)\in C\})=\underbrace{P(\{\omega: X(\omega)\in B\})\cdot P(\{\omega: Y(\omega)\in C\})}_{\text{אי תלות}}$$
$$=P_X(B)\cdot P_Y(C)$$

בהכללה $X_1,\dots,X_n$ בלתי תלויים אם"ם $P_{X_1,\dots,X_n}(A_1\times\dots\times A_n)=\prod_{k=1}^{n}P_{X_k}(A_k)$

<u>למה: חלק ב' של Borel Cantelli</u>
חלק א' אמר שאם יש סדרת מאורעות $A_1,\dots$ שסכום ההסתברויות שלהם סופי, אז ההסתברות של ה lim sup היא 0.
החלק השני אומר:
אם מתקיים $\sum_{n=1}^{\infty}P(A_n)=\infty$ והמאורעות בלתי תלויים אז $P(\limsup A_n)=1$

<u>הוכחה</u>
$$(\limsup_{n\to\infty} A_n)^C=(\bigcap_{n=1}^{\infty}\bigcup_{n=1}^{\infty}A_k)^C=\bigcup_{n=1}^{\infty}\bigcap_{n=1}^{\infty}A_k^C=\liminf_{n\to\infty}A_n^C$$
כעת נתבונן במשלים
$$P(\bigcap_{k=n}^{\infty}A_k^C)=\prod_{k=n}^{\infty}P(A_k^C)=\prod_{k=n}^{\infty}(1-P(A_k))\le\underbrace{\prod_{k=n}^{\infty}e^{-P(A_k)}}_{since\quad 1-x\le e^{-x}}=\exp\left(-\sum_{k=n}^{\infty}P(A_k)\right)=0$$
כיוון שהמשלים הוא 0 – אזי $P(\limsup A_n)=1$

<u>שיעור 12 – 15.12.08</u>
בהמשך ללמה – דוגמה:
נתונה סדרה אינסופית של ניסויי ברנולי ב"ת עם הסתברות הצלחה p.
$\Omega=\{0,1\}$

מה ההסתברות שהשרשרת 1,0,1 חוזרת אינסוף פעמים.
נגדיר את המאורע $A_{3n}=\{(a_k)\in\Omega : a_{3n}=1 , a_{3n+1}=0 , a_{3n+2}=1\}$
כלומר המאורע הזה אירע אם במקום ה $3n$ יש שרשרת כמבוקש.
ולכן $(A_{3n})_{n=1}^{\infty}$ סדרה של מאורעות בלתי תלויים.
$p(A_{15})=p(1-p)p=p(A_{3n})\forall n$ ואם נסכם נקבל $\sum_{n=1}^{\infty} P(A_{3n})=\infty$ ולפי הלמה של בורל קנטלי נקבל:
$P(\limsup A_n)=1$ וזו ההסתברות לקבלת המאורע "שרשרת 1,0,1" שמתקיים בכל איבר בסדרה.

שימוש נוסף – נסתכל בסדרה של משתנים מקריים (X_n) בלתי תלויים (=כל סדרה סופית שלהם היא בלתי תלויה)
המקבלים ערכים חיוביים לפי החוק הבא:

$$F_{X_n}(x)=P(X_n\leq x)=\begin{cases}0 & x\leq 0\\ 1-e^{-x} & x>0\end{cases}$$

ולכן:

$$P(X_n>x)=\begin{cases}1 & x\leq 0\\ e^{-x} & x>0\end{cases}$$

בפרט: נציב $x=\alpha\log n$ כאשר $n\in\mathbb{N}$ ו $\alpha>0$ נקבל:

$$P(X>alph\log n)=e^{-\alpha\log n}=n^{-\alpha}$$

שאלה: בהנתן סדרה אינסופית X_n כנ"ל מה ההסתברות ש $X_n>\alpha\log n$ עבור אינסוף nים.
הערה – אפשר לומר שכל X הוא משתנה שלא "אוהב" להיות גדול, כי ההסתברות שלו להיות די גדול דועכת ממש מהר.
נסמן:

$$A_n=\{\omega\in\Omega : X_n(\omega)>\alpha\log n\}$$

אלו מאורעות בלתי תלויים.

כרגע אם נסתכל בסכום ההסתברויות אם הוא יהיה סופי אז הלמה הראשונה תבטיח לנו שההסתברות לכך שאינסוף מהם אירעו היא אפס, מצד שני – אם הסכום הוא אינסופי לפי הלמה השניה נקבל שההסתברות היא 1.

$$\sum_{n=1}^{\infty}P(A_n)=\sum_{n=1}^{\infty}n^{-\alpha}=\begin{cases}<\infty & \alpha>1\\ \infty & \alpha\leq 1\end{cases}$$

קיבלנו ש

$$P(X_n>\alpha\log n \quad \underbrace{i.o.}_{\infty \ often})=\begin{cases}1 & \alpha\leq 1\\ 0 & \alpha>1\end{cases}$$

❑

<u>סכום של משתנים מקריים</u>

יהיו X,Y משתנים מקריים
נזכיר שאנחנו חושבים על זה כעל $\Omega\to S_X\times S_Y$ ועל ההסתברות של יחידון כ $p_{X,Y}(x,y)$
נגדיר משתנה מקרי Z כך: $Z(\omega)=X(\omega)+Y(\omega)$
מה ניתן לומר על ההתפלגות של Z?
נשאל מה הערך של $F_Z(z)=P(Z\leq z)=P(\{\omega : Z(\omega)\leq z\})$?
נכתוב אחרת:

$$\{\omega : Z(\omega)\leq z\}=\coprod_{x\in S_X}\{\omega : X(\omega)=x , Y(\omega)\leq z-x\}$$

$$F_Z(z)=\sum_{x\in S_X}P(X=x , Y ,=z-x)=\sum_{x\in S_X}\sum_{y\leq z-x}P(X=x , Y=y)$$

נשים לב שכאן עשינו שימוש ברישום:

$$\{\omega : X(\omega)=x , Y(\omega)\leq z-x\}=\coprod_{y\leq z-x}\{\omega : X(\omega)=x , Y(\omega)=y\}$$

$$\Rightarrow F_Z(z)=\sum_{x\in S_X}\sum_{y\leq z-x}p_{X,Y}(x,y)$$

$$p_Z(z)=\sum_x \sum_{y=z-x} p_{X,Y}(x,y)=\sum_x p_{X,Y}(x,z-x)$$

אם X,Y בלתי תלויים אז

$$p_Z(z)=\sum_x p_X(x)\,p_Y(z-x)$$

$$F_Z(z)=\sum_x p_X(x)\sum_{y\le z-x} p_Y(y)=\sum_x p_X(x)\cdot F_Y(z-x)$$

דוגמה:

$$X\sim Poi(\lambda_1)\qquad Y\sim Poi(\lambda_2)$$

אם הם בלתי תלויים אז:

$$p_{X,Y}(x,y)=p_X(x)\,p_Y(y)=\frac{e^{-\lambda_1}}{x!}\lambda_1^x\cdot\frac{e^{-\lambda_2}}{y!}\lambda_2^y$$

מהי ההתפלגות של Z=X+Y?

$$p_Z(z)=\sum_{x=0}^{\infty} p_{X,Y}(x,z-x)=\sum_{x=0}^{z} e^{-(\lambda_1-\lambda_2)}\cdot\frac{\lambda_1^x\lambda_2^{z-x}}{x!(z-x)!}\cdot\frac{z!}{z!}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{z!}\cdot\sum_{x=0}^{z}\binom{z}{x} lambda_1^x\lambda_2^{z-x}=\frac{e^{-(\lambda_1+\lambda_2)}}{z!}(\lambda_1+\lambda_2)^z$$

כלומר גם זו התפלגות פואסונית עם מקדם שהוא סכום המקדמים.

<u>התפלגות מותנה</u>

הגדרה:

$$P_{X|Y}(x|y)=P(X=x|Y=y)=\frac{P(X=x,Y=y)}{P(Y=y)}=\frac{p_{X,Y}(x,y)}{\sum_w p_{X,Y}(w,y)}$$

דוגמה:

X,Y משתני ברנולי שנציג את התפלגותם המשותפת ע"י טבלה:

1	**0**	**Y\X**
0.1	*0.4*	**0**
0.3	*0.2*	**1**

אזי:

$$P_{X|Y}(0|0)=\frac{0.4}{0.4+0.1}$$

$$P_{X|Y}(0|1)=\frac{0.2}{0.2+0.3}$$

<u>דוגמה</u>

מספר הביצים שחרק מטיל הוא משתנה מקרי $Poi(\lambda)$

כל ביצה מתפתחת בהסתברות p

מה ההסתברות שבהטלה מסויימת יתפתחו <u>k חרקים</u>?

X מספר הביצים שהוטלו

Y מספר החרקים שהתפתחו

והשאלה היא $P(Y=k)$

לפי הנתון $X\sim Poi(\lambda)$ וגם $Y|X=n\sim B(n,p)$

$$P(Y=k)=\sum_{n=0}^{\infty} P(Y=k|X=n)P(X=n)$$ (לפי חוק ההסתברות השלמה להתפלגויות)

בעצם אין טעם לסכום כאשר $n<k$ ולכן זה:

$$\sum_{n=k}^{\infty}\binom{n}{k}p^k(1-p)^{n-k}\cdot\frac{e^{-\lambda}}{n!}\lambda^n$$

נשנה את הסכימה כך שתתחיל מ 0 (וכך נחליף את ה n ב n+k):

$$\sum_{n=0}^{\infty}\binom{n+k}{k}p^k(1-p)^{n}\cdot\frac{e^{-\lambda}}{(n+k)!}\lambda^{n+k}$$

$$=e^{-\lambda}\cdot\frac{(p\lambda)^k}{k!}\sum_{n=0}^{\infty}\frac{1}{n!}[\lambda(1-p)]^n=e^{-\lambda p}\cdot\frac{(\lambda p)^k}{k!}\sim Poi(\lambda p)$$

<u>הערות</u>

1.

$$P_{X,Y}(x,y)=p_{X|Y}(x|y)\,p_Y\,p_Y(y)$$

2.

$$p_{X,Y|Z}(x,y|z)=P(X=x,Y=y|Z=z)$$

$$p_{X|Y,Z}(x|y,z)=P(X=x|Y=y,Z=z)$$

3.

$$p_{X,Y,Z}(x,y,z)=p_{X|Y,Z}(x|y,z)\,p_{Y|Z}(y|z)\,p_Z(z)$$

<u>תרגיל</u>

יהיו שני משתנים מקריים פואסוניים בלתי תלויים:

$$X\sim Poi(\lambda_1)$$

$$Y\sim Poi(\lambda_2)$$

מהי ההתפלגות המותנה של X בהנתן X+Y?

$$p_{X|X+Y}(k|n)=\frac{P(X=k,X+Y=n)}{P(X+Y=n)}=\underbrace{\frac{P(X=k)P(Y=n-k)}{P(X+Y=n)}}_{\text{אי תלות}}$$

$$=\frac{\frac{e^{-\lambda_1}}{k!}\cdot\lambda_1^k\frac{e^{-\lambda_2}}{(n-k)!}\cdot\lambda_2^{(n-k)}}{\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\cdot(\lambda_1+\lambda_2)^n}=\binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(1-\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-k}$$

$$X|X+Y=n\sim B(n,\frac{\lambda_1}{\lambda_1+\lambda_2})$$

$$X\sim Poi(20)\qquad Y\sim Poi(30)\qquad X+Y=70\Rightarrow\quad X\sim B(70,\frac{2}{5})$$

<u>שיעור 13</u>

<u>התפלגויות מותנות</u>

$$P_{X,Y|Z,W,\ldots}(x,y|z,w,\ldots)$$

כלל הכפל:

$$p_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=p(x_n|x_1\ldots x_{n-1})\cdot p(x_{n-1}|x_1\ldots x_{n-2})\cdot\ldots\cdot p(x_1)$$

מחלקה מיוחדת של סדרות של משתנים מקריים:

$$p(x_k|x_0, x_1, \dots x_{k-1}) = p(x_k|x_{k-1})$$

כלומר "העתיד אינו תלוי בעבר בהנתן ההווה) במתמטיקה זה נקרא **<u>מרקוביות</u>** (Markov)

אם $p_{X_k|X_{k-1}}$ פונקציה שאינה תלויה ב k אז הסדרה נקראת **<u>הומוגנית בזמן</u>** במקרה זה:

$$p_{X_k|X_{k-1}}(x_k|x_{k-1}) = M_{X_k, X_{k-1}} \text{ =a matrix}$$

לסדרה של מ"מ כנ"ל קוראים **<u>שרשרת מרקוב</u>** (Markov Chain)

$$p(x_1, \dots, x_n) = M_{X_n, X_{n-1}} M_{X_{n-1}, X_{n-2}} \cdot \dots \cdot M_{X_1, X_0} p(x_0)$$

נסכום על כל הערכים האפשריים של $x_1, \dots, x_n$

$$p(x_n) = \sum_{x_0} (M^n)_{X_1, X_0} p(x_0)$$

<u>תוחלת (Expectation)</u>

$$X : \Omega \to S \subseteq \mathbb{R}$$

$$\sum_{\omega \in \Omega} X(\omega) P(\{\omega\})$$

ממוצע משוקלל של $X(\omega)$, המשקל הוא ההסתברות.

אם הסכום הזה **מתכנס בהחלט** – אז נקרא לו **<u>התוחלת של המשתנה המקרי X.</u>**

והוא מסומן ב $E[X]$

$$\Omega = \coprod_{x \in S} X^{-1}(\{x\})$$

$$\sum_{\omega \in \Omega} \Rightarrow \sum_{x \in S} \sum_{\omega \in X^{-1}(\{x\})}$$

$$E[X] = \sum_{x \in S} \sum_{\omega \in X^{-1}(\{x\})} X(\omega) P(\{\omega\}) = \sum_{x \in S} x \sum_{\omega \in X^{-1}(\{x\})} P(\{x\}) = \sum_{x \in S} x\, p_X(x)$$

<u>דוגמאות</u>

.1

זורקים קוביה – מה התוחלת של התוצאה?

X תוצאת ההטלה, אז $S = \{1 \dots 6\}$ ו $p_X(k) = \frac{1}{6}$ והתוחלת:

$$E[X] = \sum_{k=1}^{6} k\, p_X(k) = \frac{21}{6} = 3.5$$

.2

$$X \sim B(n, p)$$

$$E[X] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!\, p^k (1-p)^{n-k}}{(n-k)!(k-1)!}$$

נשחק עם האינדקסים:

$$= \sum_{k=0}^{n-1} \frac{n!\, p^{k+1}(1-p)^{n-k-1}}{(n-k-1)!\, k!} = n\, p \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-k-1} = \underbrace{n\, p\, (p+1-p)^{n-1}}_{\text{הבינום של ניוטון}} = n\, p$$

.3

$$X \sim Poi(\lambda)$$

$$E[X] = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda}}{k!} \lambda^k = \sum_{k=1}^{\infty} \frac{e^{-\lambda}}{(k-1)!} \lambda^k = \sum_{k=0}^{\infty} \frac{e^{-\lambda}}{k!} \lambda^{k+1} = \lambda$$

.4

$$X \sim Geo(p)$$

$$E[X]=\sum_{k=1}^{\infty} k\cdot \underbrace{(1-p)}_{q}{}^{k-1} p = p\sum_{k=1}^{\infty} k\, q^{k-1} = p\cdot\frac{d}{dq}\left(\sum_{k=1}^{\infty} q^k\right) = p\frac{d}{dq}\left(\frac{q}{1-q}\right) = p\frac{d}{dq}\left(\frac{q-1+1}{1-q}\right)$$

$$= p\frac{d}{dq}\left(\frac{1}{1-q}-1\right) = p\frac{-1}{(1-q)^2}\cdot \underbrace{-1}_{\text{נגזרת פנימית לפי } q} = p\frac{1}{(1-q)^2} = \frac{1}{p}$$

נעיר כי אפשר להוכיח כי $E[X]=\frac{1}{p}$ גם ע"י שינוי אינדקסים כך ש k יתחיל מ 0 ופיתוח אחר, אפשר למצוא את זה בסיכומים של רז וזה אולי פחות מסובך. הפיתוח לעיל עם גזירה של טור חזקות הוא זה שראינו בכיתה.

<u>שיעור 15</u>

נזכיר שהגדרנו את התוחלת: $E[X]=\sum_{\omega\in\Omega} X(\omega)P(\{\omega\})$ והוכחנו שזה שווה ל $\sum_{x\in S} x\, p_X(x)$

נדגיש את ההבדל בין תוחלת ובין ממוצע - אנחנו לוקחים X תוצאת ניסוי:
n פעמים, ניסוי זהה, הניסויים בלתי תלויים.
ממוצע = $\frac{1}{n}\cdot \sum_{k=1}^{\infty} k\cdot$ k מספר הפעמים שיצא

<u>דוגמה</u>
X משתנה מקרי המקבל ערכים 0,1,2
את ההסתברות אפשר להציג כטבלה:

x	0	1	2
$p_X(x)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

וכעת התוחלת תהיה $0\cdot\frac{1}{2}+1\cdot\frac{1}{3}+2\cdot\frac{1}{6}=\frac{2}{3}$

נגדיר משתנה מקרי נוסף $Y(\omega)=g(X(\omega))$ לדוגמה $Y(\omega)=X^2(\omega)$
מהי התוחלת של Y.

x	0	1	4
$p_Y(x)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

וכעת התוחלת תהיה $0\cdot\frac{1}{2}+1\cdot\frac{1}{3}+4\cdot\frac{1}{6}=1$

נראה כאילו חישבנו בשני המקרים משקל כפול הסתברות לכל ערך אפשרי וסכמנו.
עושה רושם שאם יש משתנה מקרי X עם התפלגות $p_X(x)$ ויש Y המוגדר ע"י $Y(\omega)=g(X(\omega))$ אז נקבל
$$E[Y]=\sum_{x\in S_x} g(x)\, p_X(x)$$

אבל זו **<u>בכלל לא</u>** ההגדרה של התוחלת, שאמורה להיות $E[Y]=\sum_{y\in S_y} y\, p_Y(y)$

רק "במקרה" זה מתלכד... ובעצם העובדה שהם מתלכדים היא:
<u>משפט הסטטיסטיקאי הלא-מודע</u>
X מ"מ מעל מרחב הסתברות $(\Omega,\mathscr{F},P)$
עם התפלגות נקודתית $p_X(x)$
$g:\mathbb{R}\to\mathbb{R}$ אזי:

$$E[g(x)]=\sum_{x\in S} g(x)\,p_X(x)$$
(בהנחה שאגף ימין סופי)

<u>הוכחה:</u>

נגדיר $Y(\omega)=g(X(\omega))$

הטווח של g = $S_y=g(S_X)$

אז:

$$p_Y(y)=P(\{\omega:Y(\omega)=y\})=P(\{\omega:g(X(\omega))=y\})=P(\{\omega:X(\omega)\in \underbrace{g^{-1}(y)}_{set\ \ of\ \ y's\ \ sources}\})$$
$$p_X(g^{-1}(y))=\sum_{x\in g^{-1}(y)} p_X(x)$$

והתוחלת היא: $E[Y]=\sum_y y\sum_{x\in g^{-1}(y)} p_X(x)$

אבל נטען כי אם $x\in g^{-1}(y)$ אז $y=g(x)$ ואז התוחלת תהיה $\sum_y\sum_{x\in g(y)} g(x)\,p_X(x)$ שזה בעצם

$$\sum_x g(x)\,p_X(x)$$
❑

<u>הוכחה חליפית</u>

$$E[g(x)]=\sum_{\omega\in\Omega} g(X(\omega))P(\{\omega\})=\sum_x\sum_{\omega\in X^{-1}(x)} g(X(\omega))P(\{\omega\})=\sum_x\sum_{\omega\in X^{-1}(x)} g(x)P(\{\omega\})$$
$$=\sum_x g(x)\,p_X(x)$$

אבל כאן השתמשנו באופן סמוי בכך ש Ω היא קבוצה בת מניה, ולכן זה פחות מוצלח.

<u>הערות</u>

.1

אם $a\in\mathbb{R}$ נוכל לתת משמעות לתוחלת של a אם נגדיר $X(\omega)=a$ ומשתנה מקרי כזה יש לו התפלגות:

$$p_X(x)=\begin{cases}1 & x=a\\ 0 & otherwise\end{cases}$$

.2

X,Y מ"מ $p_{X,Y}(x,y)$ כאשר $g:\mathbb{R}^2\to\mathbb{R}$

$$E[g(X,Y)]=\sum_x\sum_y g(x,y)\,p_{X,Y}(x,y)$$

והרחבה לכל מספר סופי של משתנים מקריים.

<u>מסקנה</u>

תוחלת היא פונקציונל לינארי

כלומר:

$$E[aX+bY]=aE[X]+bE[Y]$$

<u>הוכחה</u>

$$E[aX+bY]=\sum_{x,y}(ax+by)\,p_{X,Y}(x,y)=a\sum_{x,y}x\,p_{X,Y}(x,y)+b\sum_{x,y}y\,p_{X,Y}(x,y)=\ldots$$

הסיבה העמוקה יותר היא שתוחלת היא אינטגרל, ואינטגרל הוא פונקציונל לינארי.

הרבה דברים מאד מסובכים נעשים טריוויאליים אם זוכרים שתוחלת היא פונקציונל לינארי.

<u>הגדרות</u>

יהא X מ"מ

ה**<u>מומנט</u>** ה k של X יוגדר $M_k[X]=E[X^k]$

ה<u>מומנט המרכזי</u> ה k של X יוגדר $C_k[X]=E[(X-E[X])^k]$

<u>שונות</u> של X היא המומנט השני של X שתסומן $Var[X]=E[(X-E[X])^2]$

<u>סטיית התקן</u> היא השורש של השונות שתסומן $\sigma_X=\sqrt{VAR[X]}$

נסמן $\mu_X=E[X]$ ונפתח סוגריים בהגדרת השונות:

$$Var[X]=E[(X-\mu_X)^2]=E[X^2-2\mu_X X+\mu_X^2]=E[X^2]-2\mu_X E[X]+E[\mu_X^2]=E[X^2]-(E[X])^2$$

יהיו a,b קבועים אז

$$Var[aX+b]=E[(aX+b-E[aX+b])^2]=E[a^2(X-E[X])^2]=a^2 Var[X]$$

<u>טענה</u>

אם השונות היא 0 אז $X(\omega)=E[X]$ בהסתברות 1.

<u>הוכחה</u>

$Var[X]=\sum_x (x-\mu_X)^2 p_X(x)$ וזה אפס אם"ם $p_X(p_X)=1$

דוגמה:

$X\sim B(n,p)$

ראינו ש $E[X]=np$

$$E[X(X-1)]=\sum_{k=0}^{n} k(k-1)\binom{n}{k}p^k q^{n-k}=\sum_{k=2}^{n}\frac{n!}{(n-k)!(k-2)!}p^k q^{n-k}$$

$$=\sum_{k=0}^{n-2}\frac{n!}{(n-2-k)!k!}p^{k+2}q^{n-2-k}=n(n-1)p^2$$

$$Var[X]=E[X^2]-E^2[X]$$

$$E[X^2]=\sum_{k=0}^{n}k^2\binom{n}{k}p^k q^{n-k}=E[X(X-1)]+E[X]=p^2 n(n-1)+np$$

$$Var[X]=p^2 n(n-1)+np-(np)^2=-p^2 n+np=np(1-p)$$

דוגמה:

$X\sim Poi(\lambda)$

התפלגות פואסון היא גבול של $B(n,p)$

$n\to\infty$

$p\to 0$

$np\to\lambda$

$Var[X]=\lambda$

$E[X]=\lambda$

<u>הגדרות</u>:

Y,X משתנים מקריים, השונות המשותפת שלהם Covariance תהיה:

$$Cov[X,Y]=E[(X-E[X])(Y-E[Y])]$$

$$=\sum_{x,y}(x-\mu_X)(y-\mu_Y)p_{X,Y}(x,y)$$

קבוע הקורלציה שלהם הוא:

$$\rho_{X,Y}=\frac{Cov[X,Y]}{\sigma_X\cdot\sigma_Y}$$

אם לשני משתנים מקריים שונות משותפת אפס, נאמר שהם חסרי קורלציה.

נזכיר משהו שאולי הוכחנו (אולי בתרגיל...) אז בהנתן שני משתנים מקריים בלתי תלויים, בהנתן כל שתי פונקציות ממשיות

$g(Y)$ ו $h(X)$ אזי גם h ו g הם משתנים מקריים בלתי תלויים, זה גובל בטריוויאליות.

<u>טענה</u>

אם X,Y משתנים מקריים בלתי תלויים אז $E[g(X)h(Y)]=E[g(X)]\cdot E[h(Y)]$

<u>הוכחה קצרצרה</u>

$$E[g(X)h(Y)]=\sum_{x,y} g(x)h(y)\,p_{X,Y}(x,y)=\sum_{x} g(x)\,p_X(x)\sum_{y} h(u)\,p_Y(y)$$

<u>מסקנה</u>

אי תלות ⇐ אין קורלציה

מדוע? כי אי תלות משמעה:

$$Cov[X,Y]=E[(X-\mu_X)(Y-\mu_Y)]=E[X-\mu_X]\cdot E[Y-\mu_Y]=0$$

<u>טענה</u>

אין קורלציה לא בהכרח גורר אי תלות!

2	**1**	**0**	**X\Y**
$\frac{1}{3}$	0	$\frac{1}{3}$	**0**
0	$\frac{1}{3}$	0	**1**

X,Y תלויים לפי הטבלה.

$$\mu_X=\frac{1}{3}$$

$$\mu_Y=1$$

$$Cov[X,Y]=\frac{1}{3}\cdot\left(\frac{-1}{3}\right)(-1)+\frac{1}{3}\left(1-\frac{1}{3}\right)(1-1)+\frac{1}{3}\cdot\left(\frac{-1}{3}\right)(2-1)=\frac{1}{9}-\frac{1}{9}=0$$

<u>שימושים בלינאריות של התוחלת</u>

1.

<u>טענה</u>

$$Var[X+Y]=E[(X+Y-\mu_X-\mu_Y)^2]=E[(X-\mu_X)^2+(Y-\mu_Y)^2+2(X-\mu_X)(Y-\mu_Y)]$$
$$=Var[X]+Var[Y]+2\,Cov[X,Y]$$

כלומר השונות אינה לינארית אלא אם אין קורלציה בין המשתנים המקריים ובפרט אם הם בלתי תלויים.

<u>הרחבה</u>

$$Var[\sum_{i=1}^{n} X_i]=\sum_{i=1}^{n} Var[X_i]+2\sum_{i<j} Cov[X_i,X_j]$$

2.

$$X\sim B(n,p)$$

$$E[X]=np \qquad Var[X]=np(1-p)$$

$(X_i)_{i=1}^n$ n משתני ברנולי בלתי תלויים עם פרמטר p, נסמן זאת **<u>iid</u>** כלומר Independent identically distributed

אז X יסמן כעת את המשתנה המקרי של סכום המשתנים המקריים לעיל כלומר $X=\sum_{i=1}^{n} X_i\sim B(n,p)$

ואז:

$$E[X]=E[\sum_{i=1}^{n} X_i]=\sum_{i=1}^{n} E[X_i]=n\,E[X_1]=np$$

X_1 הוא משתנה ברנולי עם פרמטר p אז

$$E[X_1]=p \qquad E[X_1^2]=p \;\Rightarrow\; Var[X_1]=p-p^2$$

$$=p(1-p) \;\Rightarrow\; Var[X]=Var[\sum_{i=1}^{n} X_i]=\sum_{i=1}^{n} Var[X_i]=n\,p(1-p)$$

.3

המזכירה המבולבלת, מה תוחלת מספר המכתבים שהגיעו ליעדם?

$$X_i=\begin{cases}1 & letter\ \ reached\ \ destination\\ 0 & otherwise\end{cases}$$

$$E[X_i]=\frac{1}{n}$$

מספר המכתבים שהגיעו ליעדם = $X=\sum_{i=1}^{n} X_i \;\Rightarrow\; E[X]=n\cdot\frac{1}{n}=1$

<u>שיעור 16</u>

בהמשך לשאלת המזכירה המבולבלת/ המלחים השיכורים / מסיבת הפיראטים:

מה השונות של X?

$$Var[X]=\sum_{i=1}^{n} Var[X_i]+2\sum_{i<j} Cov[X_i, X_j]$$

$$=Var[X]=n\,Var[X_1]+2\binom{n}{2}cov[X_1, X_2]$$

ניזכר בהגדרה של שונות משותפת:

$$Cov[X_1, X_2]=E[(X_1-E[X_1])(X_2-E[X_2])]=E[X_1X_2-X\,E[X_2]-X_2E[X_1]+E[X_1]E[X_2]]$$

$$=E[X_1X_2]-E[X_1]E[X_2]$$

ולכן נשים לב שמכפלה של המשתנים המקריים כאן גם היא משתנה מקרי והיא ההסתברות שגם המכתב ה i וגם המכתב ה j הגיעו ליעדם ולכן המשך החישוב הקודם:

$$Var[X_1]=\frac{1}{n}\cdot(1-\frac{1}{n}) \qquad Cov[X_1, X_2]=\underbrace{E[X_1X_2]}_{P(X_1=1,\,X_2=2)\to P(X=2|X=1)P(X_1=1)=\frac{1}{n-1}\cdot\frac{1}{n}} \;-\underbrace{E[X_1]E[X_2]}_{\frac{1}{n^2}}$$

$$Var[X]=\frac{1}{n}(1^{i}\frac{1}{n})\cdot n+2\cdot\frac{n(n-1)}{2}(\frac{1}{n}\cdot\frac{1}{n-1}-\frac{1}{n^2})=1-\frac{1}{n}+1-\frac{n-1}{n}=1$$

הערה: נשים לב שמכאן אפשר לראות שסיפור המזכירה המבולבלת (או המלחים וכו'...) הוא המשתנה המקרי הכי דומה למשתנה פואסוני שהוא עדיין סופי.

<u>תרגיל</u>

2n קלפים:

$1\,2\,...\,n$

$1\,2\,...\,n$

לוקחים m קלפים אקראיים, מהי <u>התוחלת</u> של מספר הזוגות שנשארו שלמים?

$$X_k=\begin{cases}1 & \text{the k couple remained intact}\\ 0 & otherwise\end{cases}$$

מספר הזוגות השלמים יהיה = X = $\sum_{k=1}^{n} X_k$

$$E[X]=\underbrace{n\,E[X_1]}_{P(X_1=1)=\frac{\binom{2n-2}{m}}{\binom{2n}{m}}}$$

חוזרים לבעיית אוסף הקופונים

X = מספר הקופונים שיש לרכוש עד השלמת אוסף מלא בגודל n.

נשאל כעת מהי התוחלת.

נחשוב על התהליך.

בהתחלה אני רוכש קופונים עד שאני מקבל קופון כלשהו (זה 1) ואחר כך עד שאני מקבל קופון שונה מהקופון הראשון וכן הלאה, לכן אפשר לחשוב על התהליך כ n צעדים שבהם בכל צעד אנחנו רוצים להגדיל את אוסף הקופונים שלנו ב 1.

X_k = מספר הקופונים שקניתי בין הרגע בו היה לי אוסף בגודל k ועד שהיה לי אוסף בגודל k+1

נזכיר כי מרחב המדגם הוא סדרות אינסופיות של קופונים.

$$X=\sum_{k=0}^{n-1} X_k \ \rightarrow\ E[X]=\sum_{k=0}^{n-1} E[X_k]$$

נטען ש X_k הוא משתנה גיאומטרי, היינו – מספר הפעמים שיש לחזור על ניסוי ברנולי עד להצלחה, ולכן:

$$X_k \sim Geo\left(\tfrac{n-k}{n}\right)$$

$$E[X_k]=\frac{n}{n-k}$$

$$E[X]=\sum_{k=0}^{n-1} \tfrac{n}{n-k}=n\sum_{k=1}^{n} \tfrac{1}{k}$$

זהו המספר ההרמוני, אשר מתבדר כמו $\log n$ ומכאן שבהנתן שצריך לאסוף 10,000 קופונים – בממוצע נצטרך לאסוף $10{,}000\cdot\log(10{,}000)$ קופונים כדי להשלים את האוסף.

תכונות נוספות של תוחלת

1.

$X(\omega)\ge a$ בהסתברות 1 $\Leftarrow$ $E[X]\ge a$ מדוע? כי $E[X]=\sum_x x\, p_X(x)\ge\sum_x a\, p_X(a)=a$

2.

$$\underbrace{E[|X|]}_{if\quad defined}\ge|E[X]|$$

הוכחה

$$|E[X]|=\left|\sum_x x\, p_X(x)\right|\le\sum_x |x|\, p_X(x)=E[|X|]$$

3.

אם

$$\sum_{n=1}^{\infty} E[|X_n|]<\infty$$

אז

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

לא נוכיח זאת כי זה קצת מהתחום של תורת המידה, אבל לעיתים נשתמש כדי להיעזר באדיטיביות במקרים אינסופיים.

3.

נניח ש $X(\omega)$ משתנה מקרי המקבל ערכים טבעיים ונגדיר:

$$X_i(\omega)=\begin{cases}1 & X(\omega)\ge i\\ 0 & otherwise\end{cases}$$

וכעת נגדיר את X:

$$\sum_{i=1}^{\infty} X_i(\omega)=\sum_{i\le X(\omega)} 1=X(\omega)$$

$$E[X]=\sum_{i=1}^{\infty}E[X_i]=\sum_{i=1}^{\infty}P(X\geq i)=\sum_{i=1}^{\infty}(1-P(X<i))=\sum_{i=1}^{\infty}(1-P(X\leq i-1))$$

ובאגף ימין קיבלנו משהו שמזכיר את פונקציית ההתפלגות של המשתנה המקרי X, ובעצם:

$$\underbrace{E[X]}_{\sum_{i=1}^{\infty} i\,p_X(i)}=\sum_{i=1}^{\infty}(1-F_X(i-1))$$

במעבר האחרון באופן פורמלי עשינו סכימה בחלקים, כמו אינטגרציה בחלקים רק דיסקרטי, מדוע? כי:

$$p_X(i)=F_X(i)-F_X(i-1)$$

<u>שיעור 17</u>

<u>תוחלת מותנה</u>

אחת ההגדרות שנתנו לתוחלת היא $E[X]=\sum_x x\,p_X(x)$

בנוסף היתה לנו התפלגות מותנה של X בהנתן Y שהוגדרה: $p_{X|Y}(x|y)=P(X=x|Y=y)$ וראינו שזו אכן התפלגות.

ברור שנוכל לבנות תוחלת של משתנה מקרי שזו התפלגותו.

הגדרה <u>**זמנית**</u>:

X, Y משתנים מקריים

התוחלת המותנה של X בהנתן Y=y היא:

$$E[X|Y=y]=\sum_x x\,p_{X|Y}(x|y)$$

דוגמה:

$X, Y\sim B(n,p)$ בלתי תלויים

מהי התוחלת של X בהנתן ש X+Y=m ?

ברור ש $m\leq 2n$ ולכן $max(0,n-m)\leq X\leq min(m,n)$

$$E[X|X+Y=m]=\sum_k k\,p_{X|X+Y}(k|m)$$

$$p_{X|X+Y}(k|m)=P(X=k|X+Y=m)$$

$$=\frac{P(X=k,X+Y=m)}{P(X+Y=m)}=\frac{P(X=k,Y=m-k)}{P(X+Y=m)}=\frac{\binom{n}{k}p^k(1-p)^{n-k}\binom{n}{m-k}p^{n-k}(1-p)^{n-m+k}}{\binom{2\text{n}}{n}p^m(1-p)^{2\text{n}-m}}$$

ולכן:

$$E[X|X+Y=m]=\sum_{k=max(0,num)}^{min(n,m)}k\cdot\frac{\binom{n}{k}\binom{n}{m-k}}{\binom{2\text{n}}{m}}$$

נחשוב על הבעיה אחרת, נתון לנו כד ובו 2n כדורים, כל אחד לבן בהסתברות p ושחור בהסתברות $1-p$

נשים לב ש:

$$E[X|X+Y=m]=E[Y|X+Y=m]=\frac{1}{2}\cdot[E[X|X+Y=m]+E[Y|X+Y=m]]$$

וזה שווה ל:

$$\frac{1}{2}[E[X+Y|X+Y=m]]=\frac{m}{2}$$

וזה נכון למרות שזה לגמרי לא שקוף מהביטוי הבינומי שהגענו אליו קודם.

<u>הגדרה סופית:</u>
X, Y משתנים מקריים
התוחלת של X בהנתן Y הוא משתנה מקרי

$$E[X|Y](\omega)=\underbrace{E[X|Y=y]}_{y=Y(\omega)}$$

או באופן שקול

$$E[X|Y](\omega)=\underbrace{\phi(Y(\omega))}_{\phi(y)=E[X|Y=y]}$$

ניתן לחשוב על זה כך – אתם תספרו לי את Y ואני אצטרך לומר מה התוחלת של X, אבל אתם עוד לא מספרים לי, אז אני מכין רשימה כך שכאשר תספרט לי מה הוא Y אוכל ישר לומר מה התוחלת של X...

<u>טענה</u>

$$E[\underbrace{E[X|Y]}_{\text{משתנה מקרי}}]=E[X]$$

<u>הוכחה</u>

$$E[E[X|Y]]=\sum_{\omega} E[X|Y](\omega)P(\{\omega\})=\sum_{y} E[X|Y=y]\,p_Y(y)$$

$$=\sum_{y}\sum_{x} x\,\underbrace{p_{X|Y}(x|y)\,p_Y(y)}_{p_{X,Y}(x,y)}=E[X]$$

זו טענה שימושית, כי לעיתים נרצה למצוא תוחלת של משתנה מקרי שקשה לחשב, ונבנה בניית עזר של משתנה מקרי נוסף, אשר נסתייע בו ע"י חישוב התוחלת המותנה.

<u>יישומים</u>
האגדה מספרת על כורה פחם שהלך לאיבוד במכרה שלו ומצא את עצמו מול 3 פתחים – A, B ו C.
אם הוא ילך ל A – הוא ימצא את דרכו החוצה בתוך 3 שעות.
אם הוא ילך ל B – הוא ימצא את עצמו באותו חדר שבו הוא כעת בתוך 5 שעות.
אם הוא ילך ל C – כמו B רק תוך 7 שעות.

העניין הוא שהוא כל כך מבולבל, שהוא לעולם לא זוכר מה בחר קודם, ולכן הוא מגריל בהסתברות $\frac{1}{3}$ כל פעם דלת אחרת.

נשאלת השאלה – מה התוחלת של זמן היציאה?

מרחב המדגם הטבעי הוא שרשרת אינסופית של אותיות כנ"ל.
$X(\omega)$ הוא זמן היציאה, מה התוחלת של X?
$Y(\omega)$ הבחירה הראשונה אז:

$$E[X]=\sum_{y} E[X|Y=y]\,p_Y(y)=\frac{1}{3}\cdot\left(E[X|Y=\text{"}A\text{"}]+E[X|Y=\text{"}B\text{"}]+E[X|Y=\text{"}C\text{"}]\right)$$

המחובר הראשון קל לחישוב

$$E[X|Y=\text{"}A\text{"}]=3$$

נאמר שבחרתי ב B – אז הלכתי 5 שעות ושבתי לנקודת המוצא.
ניתן לחלק את הזמן לזמן שהלכתי ולשאר הזמן, כלומר ל 5 שעות ולתוחלת הכללית (מאדיטיביות):

$$E[X|Y=\text{"}B\text{"}]=5+E[X]\qquad E[X|Y=\text{"}C\text{"}]=7+E[X]$$

ולכן:

$$=\frac{1}{3}\cdot\left(3+5+E[X]+7+E[X]\right)=5+\frac{2}{3}\cdot E[X]=E[X]\ \Rightarrow\ E[X]=15$$

<u>דוגמה נוספת</u>

מבצעים סדרה של ניסויי ברנולי כאשר ההסתברות להצלחה היא p.
נשאל מה התוחלת של מספר ההטלות עד לשתי הצלחות רצופות.
מרחב המדגם הוא גם כאן סדרות אינסופיות בינאריות.

$X(\omega)$ = מספר ההטלות עד לשתי הטלות רצופות

$Y_j(\omega)$ = תוצאת ההטלה ה j-ית

$$E[X]=E[E[X|Y_1]]=E[X|Y_1=0]\cdot \underbrace{p_{Y_1}(0)}_{1-p}+E[X|Y=1]\underbrace{p_{Y_1}(1)}_{p}$$

נשים לב ש:

$$E[X|Y_1=0]=1+E[X]$$

מדוע? כי נניח שקיבלנו 0 וזה ממש לא שינה את התוחלת כעת, כי הניסוי החדש (ללא ההטלה הראשונה=0) הוא למעשה הניסוי המקורי.

נעשה טריק דומה כדי להמשיך:

$$E[X|Y_1=1]=E[E[X|Y_1=1,Y_2]]=E[X|Y_1=1,Y_2=0](1-p)+E[X|Y_1=1,Y_2=1]p$$
$$=(1-p)(2+E[X])+p\cdot 2$$

נסכם:

$$E[X]=(1-p)(1+E[X])+p((1-p)(2+E[X])+2p)=(1-p)(1+E[X])+p(2+E[X]-pE[X])$$
$$=(1-p)+2p+E[X]\left((1-p)+p-p^2\right) \Rightarrow E[X]=1+p+E[X](1-p^2)=\frac{1+p}{p^2}$$

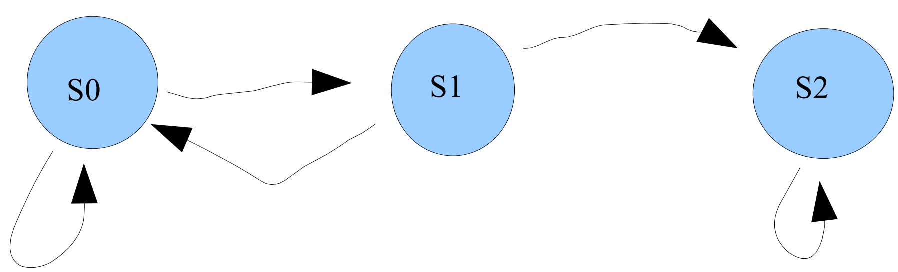
זהו אכן הפתרון, אולם נתבונן באותה בעיה מחדש קצת אחרת.
בכל נקודה לאורך השרשרת יש שלושה מצבים אפשריים המייצגים את מצבנו "לקראת" שתי אחדות רצופות:

S_0 = ראינו 0 כרגע או שאנחנו בהתחלה

S_1 = צריך עוד 1 כדי להשלים 2 אחדות רצופות

S_2 = ראינו שתי אחדות רצופות

ובאמת אפשר לחשוב על העתקה מהסדרות האינסופיות לעולם תלת-מצבי (על מכונות מצבים – בקורס בחישוביות):



$$E[X|S_0]=?$$
$$E[X|S_1]=?$$
$$E[X|S_2]=0$$

$$E[X|S_1]=p+(1-p)E[X|S_0]$$
$$E[X|S_0]=1+pE[X|S_1]+(1-p)E[X|S_0]$$

<u>פונקציה יוצרת מומנטים</u>

הגדרה:

X משתנה מקרי, נגדיר:

$$M_X(t)=E[e^{tX}]=\sum_x e^{tx}p_X(x)$$

$$M_X(0)=1$$
$$M_X'(0)=\sum_x xe^{tx}p_X(x)=E[X]$$
$$\left.\frac{d^n M_X(t)}{dt^n}\right|_{t=0}=E[X^n]$$

כלומר הנגזרות שלה יתנו לנו את המומנטים.

דוגמאות:
$$X\sim B(n,p)$$
אז:
$$M_X(t)=\sum_{k=0}^{n} e^{kt}\binom{n}{k}p^k(1-p)^{n-k}=\underbrace{(1-p+pe^t)^n}_{\text{binomial where } e^{kt}p^k=(pe^t)^k}$$
$$M_X'(t)=n(1-p+pe^t)^{n-1}pe^t$$
$$M_X'(0)=np$$
$$E[X^2]=Var[X]+(E[X])^2$$
$$Var[X]=np(1-p)$$

ניקח משתנה פואסוני:
$$X\sim Poi(\lambda)$$
$$M_X(t)=\sum_{k=0}^{\infty} e^{kt}e^{-\lambda}\frac{\lambda^k}{k!}=e^{-\lambda}e^{\lambda e^t}=e^{\lambda(e^t-1)}$$
$$M_X'(t)=e^{\lambda(e^t-1)}\lambda e^t$$
$$M_X'(0)=\lambda$$
$$M_X''(t)=e^{\lambda(e^t-1)}(\lambda e^t)^2+e^{\lambda(e^t-1)}\lambda e^t$$
$$M_X''(0)=\lambda^2+\lambda$$
(את החישוב לעיל העתקתי מהר מדי וסביר להניח שיש בהם טעויות)

<u>משתנים מקריים רציפים</u>

הגדרה: יהי מרחב הסתברות $(\Omega,\mathscr{F},P)$ ומשתנה מקרי $X:\Omega\rightarrow\mathbb{R}$
X נקרא **משתנה מקרי רציף** אם קיימת פונקציה אינטגרבילית אי שלילית f_X (צפיפות ההתפלגות של X) כך ש:
$$P(\{\omega:X(\omega)\le x\})=F_X(x)=\int_{-\infty}^{x} f_X(t)dt$$

$$P_X((a,b])=P_X((\infty,b]\setminus(-\infty,a])=F_X(b)-F_X(a)=\int_a^a f_X(t)\,dt$$

נגדיר $A=\coprod_{j=1}^{\infty}(a_j,b_j]$ אז $P_X(A)=\int_A f_X(t)dt$ כלומר $P_X(A)=\sum_{j=1}^{\infty}\int_{a_j}^{b_j} f_X(t)\,dt$

ובנוסף $\int_{-\infty}^{\infty} f_X(t)\,dt=1$

<u>מסקנה:</u>

$$P_X(\{a\})=P_X([a,b]\setminus(a,b))=0$$

<u>משתנה מקרי אחיד (Uniform)</u>

$$X \sim U[a,b]$$

$$f_X(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & otherwise\end{cases}$$

זה פשוט קו ישר כך שהאינטגרל שלו יהיה בדיוק 1.

דוגמה:

רכבת מגיעה לתחנה בכל רבע שעה עגולה, אדם מגיע לתחנה בזמן אקראי המתפלג אחיד בין 7 ל 7:30, מה ההסתברות שהמתין לרכבת פחות מ 5 דקות?

אזי זו ההסתברות שהוא הגיע בדיוק ב 7, או בין 7:10 ל 7:15 או בין 7:25 ל 7:30.

זה סכום של אינטגרלים על שלושה קטעים, הראשון באורך 0, ושני האחרים כך שהאינטגרל שלהם $\frac{5}{30}$ כל אחד.

ולכן ההסתברות היא $\frac{10}{30}=\frac{1}{3}$

<u>קטעים נבחרים מהתרגול</u>

כדי להוכיח את השוויון (שיופיע בתרגיל):

$$Var(\sum_{i=1}^{n} X_i)=\sum_{i=1}^{n} Var(X_i)+2\sum_{i<j} Cov(X_i,X_j)$$

כדאי להוכיח את הלמה הבאה:

$$Cov(\sum X_i,\sum Y_j)=\sum_i\sum_j Cov(X_i,Y_j)$$

טענה טריוויאלית:

X, Y מ"מ ב"ת אז:

$$E[X|Y=y]=E[X]$$

המלצות לתרגיל – לראות את ההוכחה ל:

X,Y משתנים מקריים דיסקרטיים

תהי A קבוצה מדידה בסיגמה אלגברה הנוצרת ע"י y.

$$\sum_{\omega\in A} X(\omega)P(\omega)=\sum_{\omega\in A} E[X|Y(\omega)]P(\omega)$$

לדברי המתרגל הנקודה החשובה היא להבין ש $E[X|Y(\omega)]$ זו תוחלת של משתנה מקרי.

<u>שיעור 18</u>

$$X \sim U[a,b]$$

$$f_X(x)=\begin{cases}\frac{1}{b-a} & a\le x\le b\\ 0 & x\notin[a,b]\end{cases}$$

<u>פרדוקס ברטרנד</u>

נעביר מיתר במעגל באופן אקראי, מה ההסתברות שאורך המיתר גדול מצלע משולש שווה צלעות חסום במעגל.

יצא לו גם חצי ושליש... איך? בפעם הראשונה הוא אמר – ניקח רדיוס ונבחר באקראי נקודה על הרדיוס ונעביר ניצב, ההסתברות שזה יהיה מיתר שארכו כנ"ל היא חצי.

אפשרות שניה – מעבירים משיק ומגרילים זווית ביחס למשיק.

הגדרה:

$X:\Omega\to\mathbb{R}$ יקרא **משתנה מקרי נורמלי** (או גאוס) עם פרמטרים (μ,σ^2) אם:

$$f_X(x)=\frac{1}{\sqrt{2\Pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

נסמן זאת: $X\sim N(\mu,\sigma^2)$

אם $\mu=0,\sigma^2=1$ כלומר $X\sim N(0,1)$

X יקרא **משתנה מקרי נורמלי תקני (סטנדרטי)** אם:

$$F_X(x)=\frac{1}{\sqrt{2\Pi}}\cdot\int_{-\infty}^{x}e^{-\frac{y^2}{2}}dy=\Phi(x)$$

Euler's error function

<u>טענה:</u>

אם $X\sim N(\mu,\sigma^2)\ ,\ Y=aX+b$

אז $Y\sim N(a\mu+b,a^2\sigma^2)$

מסקנה: $Y=\frac{X-\mu}{\sigma}$

(כלומר $a=\frac{1}{\sigma},b=\frac{\mu}{\sigma}$)

אז:

$Y\sim N(0,1)$

<u>הוכחה</u>

נניח ש a>0

$$F_Y(y)=P(Y\le y)=P(aX+b\le y)=P(X\le\frac{y-b}{a})=\frac{1}{\sqrt{2\Pi\sigma^2}}\int_{-\infty}^{\frac{y-b}{a}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$u=ax+b\qquad x=\frac{u-b}{a}\qquad dx=\frac{du}{a}$$

$$\Rightarrow\ \frac{1}{\sqrt{2\Pi\sigma^2}}\int_{-\infty}^{y}e^{-\frac{\frac{u-b}{a}-\mu}{2\sigma^2}}\frac{du}{\sqrt{a^2}}=\frac{1}{\sqrt{2\Pi\sigma^2a^2}}\int_{-\infty}^{y}e^{-\frac{(u-b-a\mu)^2}{2\sigma^2a^2}}du$$

לקראת סוף סמסטר נלמד את משפט הגבול המרכזי שאומר את הדבר הבא:

אם נתונים לנו המון משתנים ב"ת ושווי התפלגות X_k אז $\frac{1}{n}\sum_{k=1}^{n}\frac{X_k-E[X_k]}{\sqrt{Var[X_k]}}$ נקבל משתנה מקרי שהתוחלת שלו 0 והשונות 1. כאשר נשאיף את n לאינסוף המשפט מספר שבמובן כלשהו זה מתכנס למשתנה נורמלי $N(0,1)$

נשים לב שבמונה אנחנו מאפסים את התוחלת והחלוקה נותנת כבר שונות 1 (זה בעצם נירמול).

<u>משפט דה-מואבר – לפלס (deMoiver-Laplace)</u>

X_k סדרה של משתני ב"ת ברנולי עם פרמטר p.

נסמן:

$$Y_n=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\frac{X_k-p}{\sqrt{p(1-p)}}$$

אזי

$Y_n \to N(0,1)$ במובן הבא:

$$\lim_{n\to\infty} P(a \le Y_n \le b) = \underbrace{\Phi(b) - \Phi(a)}_{\text{משתנה נורמלי תקני}}$$

זו התכנסות במידה, או התכנסות בהתפלגות (Convergence in law) או שאיפה בהתפלגות, אין תמימות דעים על התרגום של המושג הזה.

<u>הוכחה</u>

$$P(a \le Y_n \le b) = P\left(a \le \frac{1}{\sqrt{n}} \sum_k \frac{X_k - p}{\sqrt{(1-p)p}} \le b\right)$$

$$P\left(np + \sqrt{p(1-p)n} \;\; a \le \sum_{k=1}^{n} X_k \le np + \sqrt{p(1-p)n} \;\; b\right)$$

$$P(a \le Y_n \le b) = \sum_{k=np+\sqrt{p(1-p)n}\;\;a}^{np+\sqrt{p(1-p)n}\;\;b} \binom{n}{k} p^k (1-p)^{n-k}$$

צ"ל:

$$n \to \infty \Rightarrow = \int_a^b \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\Pi}} dy$$

נשים לב שמשתנה הסכימה לעיל הוא

$$k = np + \sqrt{p(1-p)n} \;\; m$$

כאשר m הולך מ a ל b בקפיצות של $\Delta m = \frac{1}{\sqrt{p(1-p)n}}$

ולכן נעשה רק שינוי משתנים, בלי שום מניפולציה נוספת ונקבל:

$$P(a \le Y_n \le b) = \sum_{m=a}^{b} \frac{1}{\Delta m} \binom{n}{k} p^k (1-p)^{n-k} \Delta m$$

נסתכל על הפיסה הראשונה בסכום:

$$\frac{1}{\Delta m} \binom{n}{k} = \frac{n!}{k!(n-k)!} \sqrt{p(1-p)n}$$

(נשתמש בקירוב סטירלינג, כי מעניינת אותנו ההתנהגות של הביטוי בשאיפה לאינסוף)

$$= \underbrace{\frac{\sqrt{2\Pi}\, n^{n+\frac{1}{2}} e^{-n} \sqrt{p(1-p)n}}{\sqrt{2\Pi}\, k^{k+\frac{1}{k}} e^{-k} \sqrt{2\Pi} (n-k)^{n-k+\frac{1}{2}} e^{-n-k}}}_{n\to\infty} = \frac{n^n n \sqrt{p(1-p)}}{\sqrt{2\Pi} (np)^{np} [n(1-p)]^{n(1-p)}} = \frac{n^n n \sqrt{p(1-p)}}{\sqrt{2\Pi}\, n^{np} p^{np} n^{n(1-p)} (1-p)^{n(1-p)}}$$

קצת הלכנו לאיבוד – אבל נחזור משם בשיעור הבא.
(שנה טובה)

<u>שיעור 19</u>

רז משתעשע בצרפת, אביתר (המתרגל) מחליף אותו.
נדבר על משפט דה-מואבר לפלאס, בסוף לא נשלים את ההוכחה (שניתן למצוא ברשימות של רז) מכיוון שמדובר בעניין טכני בעיקרו, אשר בסופו של דבר הוא מקרה פרטי של משפט הגבול המרכזי שנראה בסוף הסמסטר.

$$Z \sim B(n, p)$$

$$Z=x_1+\ldots+x_n$$

$$S_n=\frac{\sum_{k=1}^{n} x_k-np}{\sqrt{np(1-p)}}=\frac{Z-np}{\sqrt{np(1-p)}}$$

הערה:

$$X\sim(\mu,\sigma^2)$$

$$F_X(x)=P(X\le x)=P\left(\frac{X-\mu}{\sigma^2}\le\frac{x-\mu}{\sigma^2}\right)=\Phi\left(\frac{X-\mu}{\sigma^2}\right)$$

דוגמאות: משך הריון רגיל הוא משתנה מקרי $X\sim N(270,100)$

מסתבר שבעלה לא היה בבית בין 290 לבין 240 יום לפני הלידה.

מה ההסתברות שהתינוק ממנו?

$$P(\{X>290\}\cup\{X<240\})=P(\{\frac{X-270}{10}>\frac{20}{10}\}\cup\{\frac{X-270}{10}<-3\})=\Phi(-3)+[1-\Phi(2)]=0.241$$

מטילים מטבע 40 פעמים, מה ההסתברות לבדיוק 20 H.

X – מספר הפעמים שיצא H.

$$X\sim B(40,\tfrac{1}{2})$$

$$P_X(20)=\binom{40}{20}(\tfrac{1}{2})^{20}(\tfrac{1}{2})^{20}=0.1268$$

נשתמש במשפט דה-מואבר לפלאס:

$$Y=\frac{X-40\cdot\frac{1}{2}}{\sqrt{40\cdot\frac{1}{2}(1-\frac{1}{2})}}=\frac{X-20}{\sqrt{10}}$$

$$P(20\le X<21)=P(0\le\frac{X-20}{\sqrt{10}}<\frac{1}{\sqrt{10}})\approx\Phi(\frac{1}{\sqrt{10}})-\Phi(0)=0.127\ldots$$

וראינו שגם במספר "קטן" כמו 40 המשפט נותן קירוב לא רע.

<u>התפלגות אקספוננציאלית</u>

הגדרה: $X\sim\exp(\lambda)$ אם:

צפיפות ההתפלגות = PDF = Probability Density Function:

$$f_X(x)=\begin{cases}\lambda e^{-\lambda x} & x\ge 0\\ 0 & x<0\end{cases}$$

ופונקציית ההתפלגות היא:

$$F_X(x)=\begin{cases}1-e^{-\lambda x} & x\ge 0\\ 0 & x<0\end{cases}$$

דוגמה:

אורך שיחת טלפון הוא X

$X\sim\exp(0.1)$ מה ההסתברות ששיחה תהיה ארוכה מ 10 דקות?

$$P(X>10)=1-F_X(10)=e^{\frac{-10}{10}}=\frac{1}{e}\approx 0.368$$

נניח שאנו יודעים שהשיחה כבר מתרחשת 10 דקות, השאלה הופכת להיות $P(X>20|X>10)=?$

$$=\frac{P(X<20\,,X>10)}{P(X>10)}=\frac{P(X>20)}{P(X>10)}=\frac{e^{-2}}{e^{-1}}=e^{-1}\approx 0.368$$

<u>טענה</u>

X מ"מ המקיים:

$P(X>t|X>s)=P(X>t-s)$ כאשר $t>s>0$ אזי X הוא משתנה אקספוננציאלי.

<u>הוכחה</u>

$$\frac{P(X>t\,,X>s)}{P(X>s)}=P(X>t-s)$$

$$\frac{1-F_X(t)}{1-F_X(s)}=1-F_X(t-s)$$

$$g(t)=1-F_X(t)$$

$$g(t)=g(s)-g(t-s)\ \Rightarrow\ g=\exp$$

ההוכחה של המעבר האחרון תרגיל לבית (אפשר באמצעות טורים וכו')

נזכיר את פונקציית גאמה (הרצפה של עצרת):

$$\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}\,dt$$

<u>טענה</u>

$$n\in\mathbb{N}\ \Rightarrow\ \Gamma(n+1)=n!$$

<u>הוכחה:</u>

נשים לב ש:

$$\Gamma(1)=1$$

הטענה שקול ל:

$$\Gamma(n+1)=n\,\Gamma(n)$$

נראה זאת:

$$\Gamma(n+1)=\int_0^\infty t^n e^{-t}\,dt$$

נבצע אינטגרציה בחלקים:

$$-t^n e^{-t}\Big|_0^\infty+\int_0^\infty n\,t^{n-1}e^{-t}\,dt=n\,\Gamma(n)$$

<u>הגדרה</u>

$X\sim Gamma(r\,,\lambda)$ אם פונקציית צפיפות ההתפלגות היא:

$$f_X(x)=\begin{cases}\frac{\lambda^r}{\Gamma(r)}X^{r-1}e^{-\lambda x} & x\geq 0\\ 0 & x<0\end{cases}$$

אם r=1 אז $X\sim\exp$

נראה שאכן הפונקציה באינסוף נותנת 1 (כלומר זו אכן התפלגות):

$$\lambda x=t$$
$$\lambda dx=dt$$
$$\int_0^\infty f_X(x)dx=\frac{1}{\Gamma(r)}\int_0^\infty \lambda^r X^{r-1}e^{-\lambda x}dx=\frac{1}{\Gamma(r)}\int_0^\infty \lambda^{r-1}x^{r-1}e^{-\lambda x}\lambda dx=\frac{1}{\Gamma(r)}\int_0^\infty t^{r-1}e^{-t}dt=\frac{\Gamma(r)}{\Gamma(r)}=1$$

התפלגות β - המתרגל מתקשה קצת להסביר למה זה טוב, אבל נרשום זאת בכל זאת:

אם: $X\sim\beta(K,l)$

$$f_X(x)=\frac{\Gamma(k+l)}{\Gamma(k)\Gamma(l)}x^{k-1}(1-x)^{l-1}$$

כאשר $x\in[0,1]$

פונקציות של משתנים מקריים:

X משתנה מקרי

$$f:\mathbb{R}\to\mathbb{R}$$
$$Y=f(x)$$

מהי פונקציית צפיפות ההתפלגות (שתקרא לפעמים "הצפיפות") של Y?

דוגמה:

$$X\sim U(0,1)$$
$$Y=x^n$$

נובע ש Y מקבל ערכים בקטע [0,1]

$$F_Y(y)=P(Y\le y)=P(X^n\le y)\underbrace{=}_{\text{מונוטוניות}}P(X\le Y^{\frac{1}{n}})=F_X(Y^{\frac{1}{n}})$$

$$F_X(x)=\int_{-\infty}^{x}\underbrace{f_X(t)}_{1}dt=\begin{cases}0 & x<0\\ x & 0\le x\le 1\\ 1 & x>1\end{cases}$$

ולכן:

$$F_Y y=\begin{cases}0 & y=0\\ y^{\frac{1}{n}} & 0\le y\le 1\\ 1 & y>1\end{cases}$$

ואז הנגזרת תהיה:

$$f_Y(y)=F_Y'(y)=\begin{cases}\frac{1}{n}y^{\frac{1}{n}-1} & 0\le y\le 1\\ 0 & else\end{cases}$$

וזו כמובן הפונקציה שחיפשנו.

דוגמה: X משתנה מקרי רציף עם צפיפות $f_X(x)$, מה הצפיפות של $y=x^2$

$$F_Y(y)=P(Y\le y)=P(X^2\le y)=P(-\sqrt{y}\le X,X\le\sqrt{y})$$
$$=F_X(sqrtt\ y)-F_X(-\sqrt{y})\ \Rightarrow\ f_Y(y)=F_Y'(y)=\frac{f_X(\sqrt{y})+f_X(-\sqrt{y})}{2\sqrt{y}}$$

משפט

אם X משתנה מקרי רציף עם צפיפות נתונה f_X

ונתונה פונקציה ממשית $g:\mathbb{R}\to\mathbb{R}$, מונוטונית ממש וגזירה כך ש $Y(\omega)=g(X(\omega))$
אזי נקבל:

$$f_Y(y)=\begin{cases}\left|(g^{-1})'(y)\right| f_X(g^{-1}(y)) & y\in Range(g)\\ 0 & else\end{cases}$$

הערה:
אם g לא מונוטונית ממש אזי $g^{-1}(y)$ היא קבוצה (ולא נקודה)

$$\sum_{z\in g^{-1}(y)}\left|(g^{-1})'(y)\right| f_X(z)$$

נשלים את השאר עם רז.

המשך סיכומים בהסתברות(1)

<u>שיעור-ספק-תירגול 21 (אביתר מחליף את רז)</u>
דוגמה למשתנה אקספוננציאלי:
בסניף דואר שני פקידים, זמן הטיפול של פקיד בלקוח הוא $X\sim\exp(\lambda)$.
שני הפקידים תפוסים, מר כהן מגיע והוא היחיד בתור.
מה ההסתברות שמר כהן יצא אחרון?
נשתמש בתכונת חוסר הזיכרון של המשתנה המקרי המעריכי:

$$P(X>s+t|X>t)=P(X>t)$$

בזמן t הלקוח הראשון עוזב, התפלגות זמן הטיפול של הלקוח אחרי זמן t היא $Y\sim\exp(\lambda)$

$$P(X>Y)=P(Y>X)=\frac{1}{2}$$

נזכיר טענה מהשיעור הקודם -
תכונת חוסר הזיכרון גוררת $g(s+t)=g(s)g(t)$
הטענה היתה שאם g רציפה אז בהכרח $g(x)=e^{-\lambda x}$

הוכחה:

$$g\left(\tfrac{2}{n}\right)=g\left(\tfrac{1}{n}+\tfrac{1}{n}\right)=\left(g\left(\tfrac{1}{n}\right)\right)^2$$

$$g\left(\tfrac{m}{n}\right)=\left(g\tfrac{1}{n}\right)^m$$

$$g(1)=\underbrace{\tfrac{1}{n}+\tfrac{1}{n}+\dots}_{\text{n times}}=\left(\tfrac{1}{n}\right)^n$$

$$g\left(\tfrac{1}{n}\right)=(g(1))^{\frac{1}{n}}$$

$$g\left(\tfrac{m}{n}\right)=(g(1))^{\frac{m}{n}}$$

מהרציפות מימין:

$$g(x)=(g(1))^x$$

$$g(1)=g\left(\tfrac{1}{2}\right)^2$$

כעת אם נגדיר

$$\lambda=-\log(g(1))$$

וזה כמובן בסדר כי כאמור $g(1)$ חיובי, אז

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

❑

דוגמה:

$$F(x)=\begin{cases} 1-a\,e^{x/b} & x\geq 0 \\ 0 & x<0 \end{cases}$$

$$0<F(x)<1 \;\Rightarrow\; b\geq 0\,,\,0\leq a\leq 1$$

נשאל – מהם התנאים על a כך ש x מ"מ מקרי או דיסקרטי:

$a=0$ אז X מ"מ דיסקרטי

$a=1$ אז X מ"מ רציף

$0<a<1$ אז מעורב

דבר אחר – **טבלת התפלגות סטנדרטית**

$$X\sim N(0,1)$$

$$F_X(t)=\Phi(t)$$

(להוסיף ציור מויקיפדיה של התפלגות נורמלית ושל טבלת התפלגות)

באופן עקרוני אי אפשר לחשב ערכים של Φ במדויק, כיוון שהיא אינטגרל של פונקציה שאין לה פונקציה קדומה אלמנטרית:

$$\int_{-\infty}^{t}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}dx$$

לכן יש לנו טבלה של ערכי האינטגרל הנ"ל בנקודות שונות, וע"י $\Phi(a)-\Phi(b)$ נקבל את האינטגרל בתחום $[a,b]$.

<u>דוגמה אקולוגית</u>

באגם יש נזילת כספית.

ריכוז הכספית בדג:

$$X\sim N(0.25\,,0.08^2)$$

דג אינו מסוכן לאכילה אם ריכוז הכספית נמוך מ 0.3.

א. מה אחוז הדגים שניתן לאכילה.

$$P(X\leq 0.3)=P\left(\frac{X-0.25}{0.08}\leq\frac{0.3-0.25}{0.08}\right)=\Phi\left(\frac{0.05}{0.08}\right)=\Phi(0.625)\approx 73$$

ב. רשות הדיג רוצה לדעת מה אחוז הכספית של ה 20% הכי מורעלים.

מהו ה t כך ש:

$$P(X\leq t)=0.8$$

$$P\left(\frac{X-0.25}{0.08}\leq\frac{t-0.25}{0.08}\right)=0.2$$

$$\frac{t-0.25}{0.08}=\Phi^{-1}(0.8)$$

$$\Rightarrow\quad t=0.08\cdot\Phi^{-1}(0.8)+0.25=0.08\cdot 0.85+0.25=?$$

<u>שיעור 22</u>

<u>משתנים מקריים רציפים</u>

נזכיר את המשמעות של פונקציית צפיפות ההסתברות -

זו לא פונקציית הסתברות, במובן זה שההסתברות לכל ערך היא תמיד 0, אבל ההסתברות לטווח מסויים של ערכים היא האינטגרל בטווח זה.

<u>פונקציות של משתנים מקריים</u>

X מ"מ רציף עם צפיפות f_X

$g:\mathbb{R}\to\mathbb{R}$, ונגדיר $Y(\omega)=g(X(\omega))$
מהי ההתפלגות של Y?

דוגמה:
$X\sim U[0,1]$ ונגדיר $Y=X^n$
נגלה שפעמים רבות הכי קל להתעסק קודם עם פונקציית ההתפלגות, לכן נשאל מהו F_Y ?

$$F_Y(y)=P(Y\le y)=P(X^n<y)$$

$$y\ge 0 \quad P(X\le Y^{\frac{1}{n}})=F_X(Y^{\frac{1}{n}})$$

$$F_Y(y)=\begin{cases}0 & y<0\\ y^{\frac{1}{n}} & y\in[0,1]\\ 1 & y>1\end{cases}$$

$$f_Y(y)=\begin{cases}\frac{1}{n}y^{\frac{1}{n}-1} & y\in[0,1]\\ 0 & otherwise\end{cases}$$

דוגמה:
X מ"מ רציף עם צפיפות $f_X(x)$ כאשר נגדיר $Y(\omega)=X^2(\omega)$
מהי ההתפלגות של Y?

$$F_Y(y)=P(Y\le y)=P(X^2\le y)=P(-\sqrt{y}\le X\le\sqrt{y})=F_X(\sqrt{y})-F_X(-\sqrt{y})$$

$$f_Y(y)=\frac{d}{dy}F_Y(y)=\frac{f'_X(\sqrt{y})}{2\sqrt{y}}-\frac{f_X(-\sqrt{y})}{2\sqrt{y}}$$

<u>משפט:</u> X מ"מ רציף עם צפיפות f_X ו $Y(\omega)=g(X(\omega))$ ונניח ש g מונוטונית.
אזי $f_Y(y)=(g^{-1})'(y)\cdot f_X(g^{-1}(y))$
אם g אינה מונוטונית (אבל מונוטונית במספר סופי של קטעים) אזי:

$$f_Y(y)=\sum_{g^{-1}}(g^{-1})'(y)f_X(g^{-1}(y))$$

<u>הוכחה</u>
נניח בה"כ ש g עולה ממש ולכן הפיכה.

$$F_Y(y)=P(Y\le y)=P(g(x)\le y)=P(X\le g^{-1}(y))=F_X(g^{-1}(y))$$

וגוזרים.

<u>סימולציה של מספרים אקראיים</u>
איך מייצרים מספרים אקראיים על מחשב?
זה פרדוקס בפני עצמו, כי איך יוצרים משהו אקראי ממכונה דטרמיניסטית.
התשובה היא שאנחנו מזייפים קצת ונותנים מספרים פסאודו אקראיים.

דרך אחת – מייצרים סדרה $x_{n+1}=p\,x_n+q\ mod\ r$
כאשר p, q ו r די גדולים מקבלים משהו שנראה די מופרע, אבל כמובן שבהנתן נקודת ההתחלה – הוא קבוע מראש.
זו סדרה <u>פסאודו אקראית</u>.

בהנתן סדרה שכזו – יש מבחנים לבדוק "עד כמה הם אקראיים", לא נתעכב עליהם.
בסופו של דבר, לכל מחשב יש פונקציית rand שאמורה לתת מספרים בין 0 ל 1 בהסתברות שווה, כלומר
$rand\sim U[0,1]$

נניח שהמחשב אכן מייצר מספרים אקראיים, ונציג את הבעיה של לייצר מכך התפלגויות אחרות, למשל התפלגות אקספוננציאלית.

<u>בעיה</u>: בהנתן $X\sim U[0,1]$ כיצד לבנות מ"מ בעל פונקציית התפלגות F?

תהי $F:\mathbb{R}\to[0,1]$ ונקל על עצמנו בדרישה ש F עולה ממש ולכן ההופכי שלה מוגדר היטב.

נתבונן במשתנה המקרי:

$$Y(\omega)=F^{-1}(X(\omega))$$

ונשאל שוב:

$$F_Y(y)?$$

$$F_Y(y)=P(Y\le y)=P(F^{-1}(X)\le y)=\underbrace{P(X\le F(Y))}_{X\sim U[0,1]}=F(Y)$$

דוגמה: נניח שרוצים לחולל מ"מ $\exp(\lambda)$ ע"י ה X הנתון לעיל.

נזכיר כי $f_Y(y)=\begin{cases}\lambda e^{-\lambda y} & y\ge 0\\ 0 & otherwise\end{cases}$

ולכן:

$$F_Y(y)=\begin{cases}1-e^{-\lambda y} & y\ge 0\\ 0 & otherwise\end{cases}$$

כעת צריך להפוך את F:

$$x=1-e^{-\lambda y}\ \Rightarrow\ e^{-\lambda y}=1-x\ \Rightarrow\ y=\frac{-1}{\lambda}\log(1-x)$$

ואזי אם נגדיר $Y=\frac{-1}{\lambda}\log(1-X)$ נקבל אכן $Y\sim\exp(\lambda)$ כפי שרצינו. ❑

בהמשך נעבור על אותם נושאים שנגענו בהם בקשר למשתנים מקריים בדידים, רק קצת יותר מהר, כי אנחנו כבר מכירים את החומר באופן בסיסי.

<u>ריבוי משתנים מקריים</u>

נאמר שלמשתנים המקריים X,Y התפלגות משותפת רציפה אם קיימת פונקציה אי-שלילית ואינטגרבילית, נסמנה $f_{X,Y}(x,y)$ כך ש: $P((X,Y)\in A)=\int_A f_{X,Y}(x,y)\,dx\,dy$

נזכיר שזהו סימון מקוצר ל:

$$P((X,Y)\in A)=P(\{\omega\in\Omega:(X(\omega),Y(\omega))\in A\})$$

<u>הערות</u>

$$F_{X,Y}(x,y)=P_{X,Y}((-\infty,x]\times(-\infty,y])=\int_{-\infty}^{y}\int_{-\infty}^{x}f_{X,Y}(r,s)\,dr\,ds$$

$$f_{X,Y}(x,y)=\frac{\partial^2}{\partial x\partial y}F_{X,Y}(x,y)$$

$$F_X(x)=P(X\le x)=P(X\le x,Y\in\mathbb{R})=P((X,Y)\in(-\infty,x]\times\mathbb{R})=\int_{-\infty}^{x}\underbrace{\int_{-\infty}^{\infty}f_{X,Y}(r,s)\,ds}_{f_X(x)}\,dr$$

תרגיל

$$f_{X,Y}(x,y)=\begin{cases} C & x^2+y^2<R^2 \\ 0 & otherwise \end{cases}$$

- מהו C?
- מהי ההתפלגות של X?
- מהי ההסתברות ש $X^2+Y^2\le a^2\le R^2$

מכיוון שזו הסתברות אזי:

$$\int_{-\infty}^{\infty}\int f_{X,Y}(x,y)\,dx\,dy=C\pi R^2=1$$

$$C=\frac{1}{\pi R^2}$$

$$f_X(x)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy=\frac{1}{\pi R^2}\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy=\frac{2}{\pi R^2}\sqrt{R^2-x^2}$$

ולכן התשובה לשאלה השלישית תהיה $\frac{a^2}{R^2}$

<u>אי תלות של משתנים רציפים</u>

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

אם"ם הם בת"ל תמיד נכון ובמקרה הספציפי הזה

$$\int_B\int_A f_{X,Y}(x,y)\,dx\,dy=\int_A f_X(x)\,dx\int_B f_Y(y)\,dy \quad\Rightarrow\quad f_{X,Y}(x,y)=f_X(x)f_Y(y)$$

דוגמה:: $X,Y,Z\sim U[0,1]$ בלתי תלויים.

מהי ההסתברות ש $X>Y\cdot Z$?

כולם מתפלגים אחיד ולכן, למשל $f_X(x)=\begin{cases} 1 & x\in[0,1] \\ 0 & otherwise \end{cases}$

ולכן $f_{X,Y,Z}(x,y,z)=\begin{cases} 1 & (x,y,z)\in[0,1] \\ 0 & otherwise \end{cases}$

נתבונן בשאלה ואז:

$$P(X\ge Y\cdot Z)=\int_{\substack{(x,y,z)\in[0,1]^3\\ x>yz}}\int\int 1\cdot dx\,dy\,dz=\int_0^1\int_0^1\int_{yz}^1 1\,dx\,dy\,dz$$

$$=\int_0^1\int_0^1(1-yz)\,dy\,dz=\int_0^1\left(1-z\frac{1}{2}\right)dz=1-\frac{\frac{1}{2}\cdot 1}{2}=\frac{3}{4}$$

<u>סכומים של משתנים מקריים</u>

X,Y מ"מ עם צפיפות $f_{X,Y}$, מהי ההתפלגות של X+Y?

נגדיר משתנה מקרי שהוא הסכום $Z=X+Y$

כדרכנו נשאל מהי פונקציית ההתפלגות:

$$F_Z(z)=P(X+Y\le z)=\int_{x+y\le z}\int f_{X,Y}(x,y)\,dx\,dy=\int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_{X,Y}(x,y)\,dx\,dy$$

אם בנוסף X,Y בלתי תלויים אז:

$$F_Z(z)=\int_{-\infty}^{\infty} f_Y(y)\int_{-\infty}^{z-y} f_X(x)\,dx\,dy=\int_{-\infty}^{\infty} F_X(z-y)f_Y(y)\,dy$$

הצפיפות היא הנגזרת של זה לפי z ולכן:

$$f_Z(z)=\int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy$$

לפעולה הזו (יצירת פונקציה שמוגדרת ע"י סכום של שתי פונקציות) קוראים קונבולוציה והיא מסומנת בכוכבית. אם כן:

$$f_{X+Y}=f_X*f_Y$$

<u>דוגמה</u>

$$X,Y\sim U[0,1]$$

מה נוכל לומר על X+Y כמשתנה מקרי?

$$f_{X+Y}(z)=\int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy$$

מכיוון שזהו משתנה מתפלג אחיד, אין טעם להסתכל מעבר לקטע [0,1] ולכן :

$$=\int_0^1 f_X(z-y)1\,dy$$

לכן:

$$f_{X+Y}(z)=\begin{cases}0 & z\le 0\\ z & z\in[0,1]\\ 2-z & z\in(1,2)\\ 0 & z\ge 2\end{cases}$$

<u>החלפת משתנים</u>

X_1,X_2 משתנים מקריים וצפיפות משותפת f_{X_1,X_2} כלומר זו העתקה מאומגה למישור.
היא מספרת לנו את הדבר הבא – נאמר שהיינו בונים ריבוע קטן במישור בתוך תחום ההגדרה, ההסתברות ליפול בתוכו היא בקירוב גודל הריבוע (אם הוא ממש קטן). בצורה יותר ישירה – זהו האינטגרל של הפונקציה בתחומי הריבוע המבוקש.

כעת נגדיר העתקה מהמישור למישור $(X_1,X_2)\to(Y_1,Y_2)$

כמובן זה מגדיר שתי פונקציות $\begin{matrix}Y_1=g_1(X_1,X_2)\\ Y_2=g_2(X_1,X_2)\end{matrix}$ וכדי להקל את החיים, נניח הפיכות כלומר $\begin{matrix}X_1=h_1(Y_1,Y_2)\\ X_2=h_2(Y_1,Y_2)\end{matrix}$

כלומר g מונוטוניות ממש.

כעת נשאל מהי ההתפלגות של Y_1,Y_2 , נתקוף כדרכנו את פונקציית ההתפלגות:

$$F_{Y_1,Y_2}(y_1,y_2)=P((Y_1,Y_2)\in(-\infty,y_1]\times(-\infty,y_2])$$

$$=\int_{\substack{g_1(x_1,x_2)\le y_1\\ g_2(x_1,x_2)\le y_2}} f_{X_1,X_2}(x_1,x_2)dx\,dy$$

$$=\int_{-\infty}^{y_1}\int_{-\infty}^{y_2} \underbrace{f_{X_1,X_2}(h_1(u_1,u_2),h_2(u_1,u_2))\left|\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}\right|(u_1,u_2)}_{f_{Y_1,Y_2}(u_1,u_2)}\quad du_1du_2$$

<u>שיעור 24</u>

ננסח מחדש החלפת משתנים בדו מימד. נגדיר משתנה מקרי ע"י שני משתנים מקריים אחרים: $X=(X_1,X_2)$

$$g:\mathbb{R}^2\to\mathbb{R}^2$$

$$Y(\omega)=g(X(\omega))$$

$$X(\omega)=g^{-1}(Y(\omega))$$

$$F_Y(y)=P(Y\in(-\infty,y_1]\times(-\infty,y_2])=P(g(X)\in(-\infty,y_1]\times(-\infty,y_2])$$

$$=P(X\in g^{-1}((-\infty,y_1]\times(-\infty,y_2]))=\int_{g^{-1}((-\infty,y_1]\times(-\infty,y_2])} f_X(x)dx_1dx_2$$

$$\underbrace{\int_{(-\infty,y_1]\times(-\infty,y_2]} f_X(g^{-1}(y))\left|\frac{\partial g^{-1}}{\partial y}\right| dy_1 dy_2}_{f_Y(y)}$$

דוגמה:

יהיו משתנים מקריים:

$$X\sim Gamma(K,1) \qquad Y\sim Gamma(L,2)$$

$$V=\frac{X}{X+Y} \quad , \quad W=X+Y$$

מהי ההתפלגות המשותפת של V ו W?

$$f_X(x)=\frac{\lambda^k}{\Gamma(K)}x^{k-1}e^{-\lambda x} \qquad x\in[0,\infty)$$

$$X=VW$$

$$Y=W-X=W(1-V)$$

$$g(x,y)\to\left(\frac{x}{x+y},x+y\right)$$

$$g^{-1}(v,w)\to(vw,w(1-v))$$

$$V:\Omega\to[0,1]$$

$$W:\Omega\to[0,\infty)$$

$$\frac{\partial g^{-1}}{\partial(v,w)}=\left\|\begin{matrix}w & -w\\ v & 1-v\end{matrix}\right\|=|w(1-v)+vw|=w$$

$$\Rightarrow \quad f_{V,W}(v,w)=f_{X,Y}(g^{-1}(v,w))\cdot w=w\cdot\frac{1^k}{\Gamma(k)}(vw)^{k-1}\frac{e^{-vw}}{\Gamma(l)}\cdot\frac{1^l}{\Gamma(1)}\cdot[w(1-k)]^{l-1}\cdot e^{-w(1-v)}$$

כאן כבר הלכתי לאיבוד בפיתוח, יש מצב ששווה לבדוק ברשימות של רז עצמו.

עוד דוגמה:

נתונים שני משתנים מקריים X,Y

מהי ההתפלגות של X/Y?

$$V=X \qquad W=\frac{X}{Y}$$

$$g:(x,y)\to\left(x,\frac{x}{y}\right)$$

$$g^{-1}:(v,w)\to\left(v,\frac{v}{w}\right)$$

כעת נשלוף את היעקוביאן כדי להחליף משתנים ונקבל:

$$\frac{\partial g^{-1}}{\partial(w,v)}=\left|\begin{matrix}1 & \frac{1}{w}\\ 0 & \frac{-v}{w^2}\end{matrix}\right|=\frac{-v}{w^2}$$

$$\Rightarrow \quad f_{V,W}(v,w)=\left|\frac{v}{w^2}\right|f_{X,Y}\left(v,\frac{v}{w}\right)$$

אבל רצינו רק את ההתפלגות של W ולכן נטנגרל על פני המשתנה השני:

$$f_W(w)=\int_{-\infty}^{\infty}\left|\frac{v}{w^2}\right|f_{X,Y}\left(v,\frac{v}{w}\right)dv$$

אנחנו תמיד כופלים בערך המוחלט של היעקוביאן, באופן דומה לכך שכפלנו בנגזרת כדי להחליף משתנים במימד אחד. מדוע

בערך מוחלט? במימד אחד הרי שמרנו על סימן!
התשובה – במימד אחד כשאנחנו מטנגרלים אנחנו יכולים לעבוד בשני הכיוונים ולקבל את אותו אינטגרל בסימנים שונים, כשמטנגרלים קבוצה במישור וכן הלאה – תמיד נעבוד עם ערך מוחלט.

דוגמה משעשעת:
$X\sim U[0,1]$
Y מ"מ כלשהו (נניח רציף, למרות שזה לא משנה)
נניח ש X ו Y הם בלתי תלויים.

נבנה מהם משתנה מקרי $W=X+Y \bmod 1$
<u>טענה</u>
$W\sim U[0,1]$

<u>הוכחה</u>
נשאל מה $P(W\le c)$ כאשר $0\le c\le 1$?
ונרצה להוכיח שהתשובה היא בדיוק c.

נוכל תמיד לפרק את Y לחלקו השלם ולשבר ע"י $Y=N+Z$
כאשר N שלם ו $Z\in[0,1)$
אזי:
$$P(W\le c)=P(\{Z\le c, X\le c-Z\}\cup\{Z\le c, X\ge 1-Z\}\cup\{Z>c, 1-Z<X<1-(Z-c)\})$$
המאורעות המאוחדים בסוגריים זרים ולכן ההסתברות שלהם היא סכום ההסתברויות.
נסתכל על שני המאורעות הראשונים ובנפרד על השלישי:
$$P(\{Z\le c, 0\le X\le c-Z\ \vee\ 1-Z\le X\le 1\})+P(\{Z\ge c, 1-Z\le X\le 1-(Z-c)\})$$
נטען שהאיבר הראשון הוא c פעמים ההסתברות ש $z\le C$ ואז:
$$=cP(\{Z\le c\})+c\,P(\{Z\ge c\})=c[P(\{Z\le c\})+P(\{Z\ge c\})]$$
אבל בסוגריים אלו מאורעות משלימים לכן הם 1 והראינו את הדרוש.

<u>שיעור 25</u>
בהמשך להחלפת משתנים, ב 1958 הראו בוקס ומולר כי בהנתן $X_1\,X_2$ משתנים מקריים ב"ת מתפלגים אחיד על הקטע $[0,1]$ אזי נוכל להגדיר שני משתנים מקריים ב"ת המתפלגים נורמלי סטנדרטי ע"י:
$$Y_1=\sqrt{-2\log X_1}\sin X_2$$
$$Y_2=\sqrt{-2\log X_1}\cos X_2$$
<u>צפיפות מותנה</u>

תזכורת – במשתנים בדידים $p_{X|Y}(x|y)=\dfrac{p_{X,Y}(x,y)}{p_Y(y)}$ התפלגות נקודתית מותנה.

נגדיר באופן דומה: $f_{X|Y}(x|y)=\dfrac{f_{X,Y}(x,y)}{f_Y(y)}$
<u>טענה:</u> ל $f_{x|y}$ תכונות של צפיפות (אם מוגדר)
<u>הוכחה:</u>
מקרה 1 – אי שלילי:
$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)\,dx=\frac{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx}{f_Y(y)}=1$$
נראה סביר ש $\int_A f_{X|Y}(x|y)\,dx=P(X\in A|Y=y)$

אבל במשתנים רציפים לכל y מתקיים $P(Y=y)=0$ ולכן אנחנו בבעיה, נסתכל על זה כך:

$$P(X\in A|Y=y):=\lim_{\varepsilon\to 0} P(X\in A|_{y-\varepsilon\le Y\le y+\varepsilon})$$

ובאופן כללי יותר: $y\in B_n\to y$ (B_n סדרת קבוצות שיורדת ל y)

$$P(X\in A|Y=y):=\lim_{\varepsilon\to 0} P(X\in A|Y\in B_n)$$

ואז:

$$P(X\in A|Y=y)=\lim_{\varepsilon\to 0}\frac{P(X\in A, y-\varepsilon\le Y\le y+\varepsilon)}{P(y-\varepsilon\le Y\le y+\varepsilon)}$$

$$=\lim_{\varepsilon\to 0}\frac{\int_A\int_{y-\varepsilon}^{y+\varepsilon} f_{X,Y}(x,u)\,du\,dx}{\int_{y-\varepsilon}^{y+\varepsilon} f_Y(u)\,du}=\int_A\lim_{\varepsilon\to 0}\frac{\int_{y-\varepsilon}^{y+\varepsilon} f_{X,Y}(x,u)\,du}{\int_{y-\varepsilon}^{y+v\varepsilon} f_Y(u)\,du}dx=\int_A\frac{f_{X,Y}(x,y)}{f_Y(y)}dx=\int_A f_{X|Y}(x|y)\,dx$$

<u>*הערה*</u>

בעולם הבדיד אמרנו שהתוחלת של $E[X]$ מוגדרת אם $\sum_x p_X(x)<\infty$ וגם הטור הזה מתכנס בהחלט.

באותו אופן נדרוש מהאינטגרל להתכנס בהחלט, ואז להכניס את הגבול לתוך האינטגרל זה לגיטימי.

<u>תוחלת</u>

<u>תזכורת</u>

$$E[X]=\sum_\omega X(\omega)P(\{\omega\})=\sum_\omega x\,p_X(x)$$

וכאשר X משתנה רציף אזי:

$$E[X]=\int_\Omega X(\omega)P(d\omega)$$

אבל זה קצת חסר משמעות עבורנו (אלו מאיתנו שלא למדנו תורת המידה) ואז ההרחבה המתבקשת (לפחות בעיני רז) של התוחלת עבור המשתנה הבדיד היא משהו כזה

$$E[X]=\int_{\mathbb{R}} x\,\underbrace{p_X(dx)}_{f_X(x)dx}$$

כיוון שאנחנו מנסים למשקל את ההסתברויות.

לכן נקבל זאת כהגדרה מעתה, כלומר

$$E[X]=\int_{\mathbb{R}} x\,f_X(x)\,dx$$

<u>דוגמה:</u>

1. $X\sim U[a,b]$ תחושת הבטן אומרת שהתוחלת תהיה הממוצע של a ו b.

נבדוק זאת:

$$E[X]=\int_a^b x\frac{1}{b-a}dx=\frac{1}{b-a}\left(\frac{1}{2}b^2-\frac{1}{2}a^2\right)=\frac{b+a}{2}$$

2. $X\sim\exp(\lambda)$ אז $f_X(x)=\lambda e^{-\lambda x}$

$x\ge 0$

נחשב:

$$E[X]=\int_0^\infty x\lambda^{-\lambda x}dx=\lambda\left[\frac{-1}{\lambda}x e^{-\lambda x}\Big|_0^\infty+\lambda\int_0^\infty\frac{1}{\lambda^{-\lambda x}}dx\right]=\frac{1}{\lambda}$$

(לא לגמרי עקבתי ויתכן שהעתקתי לא נכון – הרעיון היא אינטגרציה בחלקים, לא משהו יותר מזה)

3. $X\sim N(\mu,\sigma^2)$

$$E[X]=\int_{-\infty}^{\infty} x\, e^{\frac{-(x-\mu)^2}{2\sigma^2}}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}\,dx$$

נחליף ל $x-\mu=y$ ונקבל שזה שווה ל:

$$E[X]=\frac{1}{\sqrt{2\pi\sigma^2}}\cdot\int_{-\infty}^{\infty}(y+\mu)\,e^{\frac{-y^2}{2\sigma^2}}\,dy=\mu$$

האינטגרל של משתנה מקרי נורמלי הוא 1 ולכן המעבר האחרון.

נוכיח כעת את משפט הסטטיסטיקאי הלא מודע, אך קודם נוכיח למת עזר.

<u>למה:</u>

$$E[X]=\int_0^{\infty}[1-F_X(x)-F_X(-x)]dx$$

<u>הוכחה</u>:

נעשה מסאז' תאילנדי לביטוי, ונתעלל בו עד שהוא יראה כמו התוחלת.

נשים לב שאנחנו צריכים לחשב את ההסתברות ש X קטן מ $-x$ ולחבר אותה להסתברות ש X גדול מ x.

$$\int_0^{\infty}P(X>x)\,dx=\int_0^{\infty}\int_x^{\infty}f_X(u)\,du\,dx=\int_0^{\infty}\int_0^{u}f_X(u)\,dx\,du$$

החלפנו סדר אינטגרציה וכעת האינטגרל הפנימי לא תלוי ב x ולכן הוא $u\,f_X(u)$ וכל הביטוי שווה ל:

$$\int_0^{\infty}u\,f_X(u)\,du$$

ולחלקו השני של הביטוי:

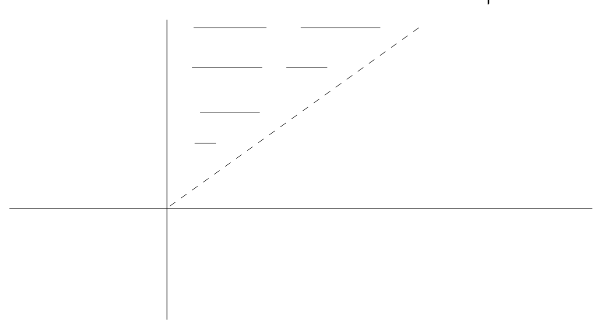
$$\int_0^{\infty}P(X\le -x)\,dx=\int_0^{\infty}\int_{-\infty}^{-x}f_X(u)\,du\,dx=\int_{-\infty}^{0}\int_0^{-u}f_X(u)\,dx\,du=-\int_{-\infty}^{0}u\,f_X(u)\,du$$

וכעת נחבר:

$$=\int_0^{\infty}P(X>x)dx-\int_0^{\infty}P(X\le -x)\,dx=\int_{-\infty}^{\infty}u\,f_X(u)\,du=E[X]$$

כנדרש.

הערה גרפית לגבי ההנמקה להחלפת סדר האינטגרציה:



כאילו אנחנו מטנגרלים לפי משתנה אחד מ 0 עד אינסוף ואז לפי השני מ 0 עד השני, שזה כמו לטנגרל מהשני עד אינסוף, ומתחתיו – מ הראשון עד אינסוף. (כדאי לצייר כדי להשתכנע)

<u>משפט הסטטיסטיקאי הלא-מודע לאינטגרלים</u>

X מ"מ רציף ו $g:\mathbb{R}\to\mathbb{R}$ אז:

$$E[g(X)]=\int_{-\infty}^{\infty} g(x) f_X(x) dx$$

<u>הוכחה:</u>

$$Y=g(X)$$

$$E[Y]=\int_{-\infty}^{\infty} y f_Y(y) dy$$

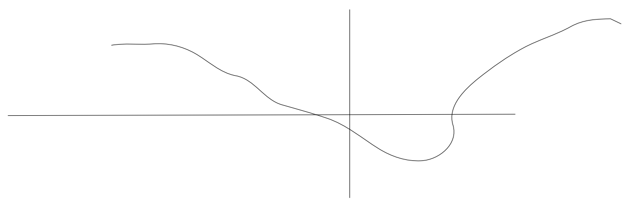
מהלמה

$$E[g(X)]=\int_0^{\infty}[P(g(X)>y)-P(g(X)\le -y)]dy=\int_0^{\infty}\left[\int_{g(x)>y} f_X(x)dx-\int_{g(x)<-y} f_X(x)dx\right]dy$$

נסמן את האינטגרל הראשון ע"י I ואת השני ע"י II ואז:

$$I=\int_0^{\infty}\int_{g(x)>y} f_X(x) dx\, dy=?$$

נטנגרל לפי x ממינוס אינסוף עד אינסוף ואז נטנגרל רק עבור קטעים שבהם g אי-שלילית, אפשר לצייר זאת כניסיון לכסות גרף של פונקציה כלשהי (נצייר משהו שרירותי) בשני שלבים, כמקודם:



נניח שאנחנו רוצים לכסות את השטח התחום מתחת לפונקציה, רק באיזורים בהם היא חיובית...

בכל אופן זה מאפשר לנו להחליף את סדר האינטגרציה ולקבל:

$$I=\int_{-\infty}^{\infty}\int_0^{max(0,g(x))} f_X(x) dy\, dx=\int_{-\infty}^{\infty} f_X(x)\cdot max(0,g(x)) dx$$

באופן דומה נחשב את II כך:

$$II=\int_0^{\infty}\int_{g(x)<-y} f_X(x) dx\, dy=\int_{-\infty}^{\infty} f_X(x) max(0,-g(x)) dx$$

וכעת נסכום:

$$I-II=\int_{-\infty}^{\infty} f_X(x)[max(0,g(x))-max(0,-g(x))]dx=\int_{-\infty}^{\infty} f_X(x) g(x) dx$$

<u>שונויות</u>

$$Var[X]=E[(X-E[X])^2]=\int_{-\infty}^{\infty}(x-E[X])^2 f_X(x) dx$$

למשל $X\sim N(\mu,\sigma)$ ואז:

$$Var[X]=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}(x-\mu)^2 e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

נגדיר $x-\mu=y$ כמקודם וגם $\frac{y}{\sigma}=z$:

$$Var[X]=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} y^2 e^{\frac{-y^2}{2\sigma^2}}dy=\frac{\sigma^3}{\sqrt{2\pi\sigma^2}}\underbrace{\int_{-\infty}^{\infty} z^2 e^{\frac{-z^2}{2}}dz}_{\sqrt{2\pi}}=\sigma^2$$

ניזכר בפונקציה יוצרת המומנטים שהגדרנו עבור משתנים בדידים: $M_X(t)=E[e^{tx}]$

תזכורת – פונקציה זאת, מרגע שחישבנו אותה, מאפשרת לנו ע"י גזירה לקבל תוחלת, ואז נוכל לייצר שונות וכו'.

$$=\int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx$$

זהו טרנספורם לפלס של הצפיפות, אבל זה מעבר לתחום של הקורס ולכן נעצור כאן.

<u>דוגמה</u>

$$X\sim N(\mu,\sigma^2)$$

$$M_x(t)=\int_{-\infty}^{\infty} e^{tx}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$=\frac{1}{\sqrt{2\pi\sigma^2}}\cdot\int_{\mathbb{R}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}+tx\right)dx$$

$$=\frac{1}{\sqrt{2\pi\sigma^2}}\cdot\int_{\mathbb{R}}\exp\left(-\frac{x^2-2\mu x+\mu^2-2t\sigma^2 x}{2\sigma^2}\right)dx$$

$$=\frac{1}{\sqrt{2\pi\sigma^2}}\cdot\int_{\mathbb{R}}\exp\left(-\frac{[x-(\mu+\sigma^2 t)]^2+\mu^2-(\mu+\sigma^2 t)^2}{2\sigma^2}\right)dx$$

$$=e^{\frac{-\mu^2}{2\sigma^2}+\frac{(\mu+\sigma^2 t)^2}{2\sigma^2}}\underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}\cdot\int_{\mathbb{R}} e^{\frac{-[x-(\mu+\sigma^2 t)]^2}{2\sigma^2}}dx}_{1}=\exp\left(\frac{\mu^2}{2\sigma^2}+\mu t+\frac{\sigma^2 t^2}{2}\right)$$

וזו הפונקציה יוצרת המומנטים, ונוכל לגזור אותה ולקבל את הפונקציות שהגענו אליהן קודם, כלומר שהשונות היא אכן σ^2 והתוחלת μ .

(תכל'ס רז עשה את זה על הלוח, סתם אני מתעצל)

לפחות כרגע נראה שהיה קל יותר לחשב ישירות את השונות והתוחלת, בלי לעבור בפונקציה יוצרת המומנטים שהיה די אסון לטנגרל אותה, אבל רז ינסה לשכנע אותנו שזה לא כך בהמשך.

<u>דוגמה</u>

$$X\sim Gamma(\lambda,r)$$

$$f_X(x)=\frac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x}$$

נראה את הפונקציה יוצרת המומנטים במקרה זה:

$$M_X(t)=\frac{\int_0^{\infty}\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x}e^{tx}dx=\frac{\lambda^r}{\Gamma(r)}\int_0^{\infty}x^{r-1}e^{-(\lambda-t)x}dx$$

וכעת נסמן $y=(\lambda-t)x$ ואז נקבל:

$$M_X(t)=\frac{\lambda^r}{\Gamma(r)(\lambda-t)^r}\underbrace{\int_0^{\infty}y^{r-1}e^{-y}dy}_{\Gamma(r)}=\left(\frac{\lambda}{\lambda-r}\right)^r\ \Rightarrow\ X\sim\exp(\lambda):M_X(t)=\frac{\lambda}{\lambda-t}$$

ונוכל לקבל ע"י כך את השונות והתוחלת.

<u>שיעור 26</u>

בשיעור שעבר הזכרנו את הפונקציה יוצרת המומנטים, ונניח שקיימת תוחלת אז $M_X(t)=E[e^{tx}]$

דיברנו על כך שבהנתן מ"מ X עם צפיפות f_X ומ"מ Y עם צפיפות f_Y למשתנה המקרי שהוא הסכום יש צפיפות שהיא הקונבולוציה של הצפיפויות, כלומר $f_{X+Y}=f_X*f_Y$ כאשר

$$(f*g)(x)=\int_{\mathbb{R}} f(y)g(x-y)\,dy=\int_{\mathbb{R}} f(x-y)g(y)\,dy$$

<u>טענה:</u> אם X ו Y בלתי תלויים אז $M_{X+Y}=M_X M_Y$

<u>הוכחה:</u>

$$M_{X+Y}(t)=\int_{\mathbb{R}} e^{tz} f_{X+Y}(z)\,dz$$

$$=\int_{\mathbb{R}} e^{tz}\int_{\mathbb{R}} f_X(y) f_Y(z-y)\,dy\,dz$$

נרשום $e^{tz}=e^{ty}e^{t(z-y)}$

נבצע החלפת משתנים ע"י:

u=y

v=z-y

$$dudv=\left\|\begin{matrix}1 & 0\\ -1 & 1\end{matrix}\right\| dydz$$

ונקבל:

$$\int_{\mathbb{R}}\int_{\mathbb{R}} e^{ty}e^{tv} f_X(u) f_Y(v)\,dudv=\int_{\mathbb{R}} e^{tu} f_X(u)\,du\int_{\mathbb{R}} e^{tv} f_Y(v)\,dv$$

$$=M_X(t)M_Y(t)$$

<u>שימושים</u>

$$X\sim N(\mu_1,\sigma_1^2)$$

$$Y\sim N(\mu_2,\sigma_2^2)$$

טענו ש:

$$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2,\sigma_2^2)$$

$$M_{X+Y}(t)=M_X(t)M_Y(t)=e^{\mu_1 t+\frac{1}{2}\sigma_1^2t^2}e^{\mu_2 t+\frac{1}{2}\sigma_2^2t^2}$$

$$=\exp\left((\mu_1+\mu_2)t+\frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2\right)$$

אם קבוצת משתנים שווי התפלגות בת"ל מתפלגים כמו משתנה אקספוננציאלי עם מ λ נגדיר:

$$Y=\sum_{i=1}^{n} X_i$$

$$f_Y=((f_{X_1}*f_{X_2})*f_{X_3})*\ldots*f_{X_n}$$

$$M_Y=M_{X_1}^n$$

$$\exp(\lambda)\sim Gamma(1,\lambda)$$

$$M_{X_1}=\frac{\lambda}{\lambda-t}$$

$$M_Y=\left(\frac{\lambda}{\lambda-t}\right)^n$$

$$Y\sim Gamma(n,\lambda)$$

דוגמה למה התפלגות גאמה חשובה:
נניח שיש לנו נורה אחת, אזי בקירוב התשובה לשאלה "כמה זמן יעבור עד שאצטרך להחליף אותה" היא משתנה מקרי אקספוננציאלי, אבל אם יש לי מלאי של נורות – אזי התשובה לשאלה על כולן תהיה משתנה מקרי גאמה לפי מספר הנורות עם אותו מקדם λ .

<u>הגדרה</u> משתנה Y יקרא בעל התפלגות X_ν (כאשר ν מספר טבעי) אם התפלגותו שווה לזו של:
$$X_1^2+X_2^2+\ldots+X_\nu^2$$
כאשר אלו משתנים מקריים בלתי תלויים שווי התפלגות שכל אחד מהם הוא משתנה נורמלי סטנדרטי כלומר $N(0,1)$

$$f_{X_1^2}(x)=\frac{f_{X_1}(\sqrt{x})+f_{X_1}(-\sqrt{x})}{2\sqrt{x}}$$
$$=\frac{\frac{1}{\sqrt{2\pi}}\cdot 2}{2\sqrt{x}}\left\{e^{\frac{-x}{2}}\right\}=\frac{1}{\sqrt{2\pi x}}e^{\frac{-x}{2}}$$

נחשב את הפונקציה יוצרת המומנטים של X_1^2 :

$$M_{X_1^2}(t)=\int_0^\infty e^{tx}\frac{1}{\sqrt{2\pi x}}e^{\frac{-x}{2}}dx$$
נבצע החלפת משתנים ע"י
$$x=y^2 \quad , \quad y=\sqrt{x} \quad , \quad dy=\frac{dx}{2\sqrt{x}}$$
$$=\int_0^\infty e^{ty^2}\frac{1}{\sqrt{\pi}}e^{\frac{y^2}{2}}dy=\frac{2}{\sqrt{2\pi}}\int_0^\infty e^{-(1-2t)\frac{y^2}{2}}dy$$
הצבה נוספת ע"י
$$z=y\sqrt{1-2\mathrm{t}}$$
$$=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{\frac{-z^2}{2}}\frac{dz}{\sqrt{1-2\mathrm{t}}}=(1-2\mathrm{t})^{\frac{-1}{2}}=\frac{1}{(1-2\mathrm{t})^{\frac{1}{2}}}=\frac{\frac{1}{2}^{\frac{1}{2}}}{\left(\frac{1}{2}-t\right)^{\frac{1}{2}}}$$

<u>הגדרה:</u>
אם $X\sim X_r^2$ ו $Y\sim X_s^2$
התפלגות המשתנה שהתפלגותו $\frac{X/r}{Y/s}$ מכונה התפלגות פישר.

<u>הגדרה:</u>
$$X\sim N(1,0)$$
$$Y\sim X_\nu^2$$
$$\frac{X}{\sqrt{Y/\nu}}\sim Student \quad t\,\nu \quad distrivbution$$

(רז: זה על שם מישהו ששמו סטודנט, לא על שמכם)

<u>אי שיוויונים</u>

<u>אי-שוויון מרקוב:</u>

<u>משפט</u>:

אם X מ"מ המקבל ערכים אי-שליליים, אז לכל $a>0$:

$$P(X\geq a)\leq\frac{E[X]}{a}$$

נשים לב שאם $a<E[X]$ אז אין פה אינפורמציה כי מקבלים הסתברות גדולה מ 1 באגף ימין ובוודאי אי השוויון מתקיים.

<u>הוכחה</u>:

$$E[X]=\int_0^\infty x f_X(x)dx\geq\int_a^\infty x f_X(x)dx\geq\int_a^\infty a f_X(x)dx=a\cdot P(X\geq a)$$

רז: היה לנו פעם שר אוצר בשם דוד לוי שתבע להשוות את השכר המינימלי לשכר הממוצע במשק.
באותו נושא, אם אתם נתקלים בכתבה שמצהירה שהעשירון העליון מרויח פי 20 מהשכר הממוצע, משמעות הדבר כי:

$$P(X\geq 20E[X])\geq\frac{1}{10}$$

ומצד שני מאי-שוויון מרקוב:

$$P(X\geq 20E[X])\leq\frac{E[X]}{20E[X]}=\frac{1}{20}$$

וזה לא יתכן, אז מעבר לכשפים ברור אינטואיטיבית למה זה נכון, כי אם כל כך הרבה אנשים מרוויחים כל כך הרבה אז הם מושכים את הממוצע מעלה.

<u>שיעור 27</u>

דוגמה עוד יותר אינטואיטיבית לאי שוויון מרקוב – לא יכול להיות שיותר מחצי מתושבי המדינה ירוויחו יותר ממפי שניים מהממוצע, מדוע? אם נקבע את הממוצע לפיהם כבר נקבל סתירה.

דוגמה:

מספר בקבוקי היין שאלכוהוליסט שותה בשבוע הוא מ"מ בעל תוחלת 25.
מה ניתן לומר על ההסתברות שבשבוע נתון ישתה יותר מ 50 בקבוקים?

$$P(X\geq 50)\leq\frac{E[X]}{5}=\frac{1}{2}$$

אם אין לי מידע על ההתפלגות – זה מה שאני יכול לומר.

<u>אי-שוויון צ'בישב</u>

אם X מ"מ עם תוחלת μ (הכוונה קיימת תוחלת והיא סופית) ושונות σ^2 (באותו אופן) אז:

$$P(|X-\mu|\geq a)\leq\frac{\sigma^2}{a^2}$$

הערה: אם $a=k\sigma$ מקבלים:

$$P(|X-\mu|\geq k\sigma)\leq\frac{1}{k^2}$$

<u>הוכחה</u>

$$P(|X-\mu|\geq a)=\underbrace{P((|X-\mu|)^2\geq a^2)\leq\frac{E[(|X-\mu|)^2]}{a^2}}_{\text{מרקוב עבור משתנה מקרי: }(X-\mu)^2}=\frac{\sigma^2}{a^2}$$

הערה: אפשר כמובן להעלות ביותר מריבוע במעבר הראשון בהוכחה ולקבל הערכות שמשתמשות במומנטים אחרים.

דוגמה

$X \sim U[0,10]$

מהי ההסתברות ש $|X-5| \geq 4$

$$P(|X-5| \geq 4) \leq \frac{Var[X]}{16}$$

$$Var[X] = \int f_X(x)(X-5)^2 dx = \frac{1}{10}\int_0^{10}(2-5)^2 dx = \frac{25}{3}$$

$$P(|X-5| \geq 4) \leq \frac{\frac{25}{3}}{16} = 0.52$$

אי שוויון ינסן

עוסק בתוחלות של פונקציות קמורות.

נזכיר שפונקציה קמורה אם כל מיתר בה נמצא מעל הפונקציה.

באופן יותר פורמלי:

f **קמורה** אם לכל x,y ו $0<\lambda<1$ מתקיים $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$

תכונות:

פונקציות קמורות הן רציפות

נגזרות חד-צדדיות קיימות

$f'(x^-) \leq f'(x^+)$ ובפרט הנגזרת השניה אי-שלילית בכל נקודה

משפט (אי-שוויון ינסן):

לכל g קמורה ומ"מ X:

$$E[g(X)] \geq f(g[X])$$

הוכחה

נניח (לרגע) ש g גזירה פעמיים ונסמן את $E[X]$ ע"י μ

נובע (מפיתוח טיילור עם שארית):

$$g(x) = g(\mu)+g'(\mu)(x-\mu)+\underbrace{\frac{1}{2}g''(c)(x-\mu)^2}_{\geq 0}$$
$$\geq g(\mu)+g'(\mu)(x-\mu)$$

נניח ש X רציף ונכפול בפונקציית ההתפלגות (שהיא אי שלילית) f_X ונוציא אינטגרל:

$$\int g(x) f_X(x)dx \geq \int g(\mu) f_X(x)dx + \int g'(\mu)(x-\mu) f_X(x)dx$$

ממשפט הסטטיסטיקאי הלא-מודע אגף שמאל הוא התוחלת של g ולכן נקבל מכך:

$$E[g(X)] \geq g(\mu) = g(E[X])$$

כנדרש.

נניח שהפונקציה אינה גזירה פעמיים בנקודה μ, עדיין נכון שקיימת נגזרת ראשונה ימנית ושמאלית, כאשר הימנית גדולה או שווה מהשמאלית ולכל m ביניהן (במובן החלש) כלומר $g'(\mu^-) \leq m \leq g'(\mu^+)$ יתקיים:

$$g(x) \geq g(\mu)+m(x-\mu)$$

וזה נכון כי מימין יש משיק לפונקציה ומשמאל הפונקציה, ומהגדרת הקמירות זה נובע. מכאן ההמשך הוא כמו במקרה הקודם.

❑

דוגמאות לפונקציות קמורות:

$$-\log, e^x, x^2, |x|$$

ניקח את אי שוויון ינסן עבור הפונקציה המעריכית:

עבור $g(x)=e^{xt}$

$$\underbrace{E[e^{tX}]}_{M_X(t)}\ge e^{tE[X]}$$

וכיוון שחסמנו את הפונקציה יוצרת המומנטים נסתכל במומנט השני (כלומר התוחלת) ונקבל:

$$E[X]\le\frac{1}{t}\log M_X(t)$$

<u>דוגמה</u>

X מ"מ דיסקרטי המקבל ערכים $0<x_1,x_2,\ldots,x_n$ בהסתברות אחידה.

נפעיל את אי-שוויון ינסן עם $g(x)=-\log x$:

$$E[-logX]\ge-\log E[X]$$

מלינאריות התוחלת נוכל לשלוף את המינוס החוצה ולהפוך אגפים:

[E[log X] <= log E[X

ממשפט הסטטיסטיקאי הלא מודע נקבל כי זה נותן:

$$\frac{1}{n}\sum_{i=1}^{n}\log x_i\le\log\frac{1}{n}\sum_{i=1}^{n}x_i$$

$$\log\left(\prod_{i=1}^{n}x_i\right)^{\frac{1}{n}}\le\log\frac{1}{n}\sum_{i=1}^{n}x_i$$

$$\left(\prod_{i=1}^{n}x_i\right)^{\frac{1}{n}}\le\frac{1}{n}\sum_{i=1}^{n}x_i$$

וקיבלנו בעצם הוכחה הסתברותית לאחד האגפים של **<u>אי שוויון הממוצעים</u>**, המסתמכת על אי שוויון ינסן.

אם במקום להניח שההסתברויות הן אחידות היינו נותנים לכל x_i הסתברות p_i אז היינו מקבלים:

$$\sum_i p_i\log x_i\le\log\sum_i p_i x_i$$

$$\prod_i x_i^{p_i}\le\sum_i p_i x_i$$

וזוהי הכללה של אי שוויון הממוצעים, שנכונה לכל קבוצה של x_i אי שליליים כאשר $\sum_i p_i=1$

<u>גבולות של סדרות של משתנים מקריים</u>

אנחנו עובדים במסגרת מרחב ההסתברות Ω ובו סדרות אינסופיות של משתנים מקריים.

<u>הגדרה</u>

נאמר שסדרת המשתנים המקריים (X_n) מתכנסת ל X <u>כמעט תמיד</u> (Almost Surely) אם:

$$P(\{\omega\in\Omega\ :\ \lim_{n\to\infty}X_n(\omega)=X(\omega)\})=1$$

נסמן $X_n\xrightarrow{a.s}X$

באופן שקול $X_n\xrightarrow{a.s}X$ אם:

$$P(\{\omega:\lim_{n\to\infty}X_n(\omega)\ne X(\omega)\})=0$$

והשלילה של הדבר הזה היא:

$$P\left(\left\{\omega:\exists k\in\mathbb{N}:\underbrace{|X_n(\omega)-X(\omega)|}_{\text{אינסוף פעמים}}>\frac{1}{n}\right\}\right)=0$$

$$P\left(\bigcup_{x\in\mathbb{N}}\left\{\omega:\left|X_n(\omega)-X(\omega)\right|>\frac{1}{k}\quad \text{i.o.}\right\}\right)=0$$

אם"ם לכל $\varepsilon>0$

$$P\left(\left\{\omega:\left|X_n(\omega)-X(\omega)\right|>\varepsilon\quad i.o.\right\}\right)=0$$

i.o.=infinitely often

נגדיר:

$$B_n^{\varepsilon}=\left\{\omega:\left|X_n(\omega)-X(\omega)\right|>\varepsilon\right\}$$

ואז $X_n\xrightarrow{a.s}X$ אם"ם לכל $\varepsilon>0$ מתקיים $P(\limsup_{n\to\infty}B_n^{\varepsilon})=0$

<u>הגדרה</u>

נאמר ש X_n מתכנסת ל X <u>**בממוצע ריבועים**</u> (in the mean square) אם:

$$\lim_{n\to\infty}E[(\left|X_n-X\right|)^2]=0$$

סימון: $X_n\xrightarrow{m.s}X$

<u>הגדרה</u>

נאמר ש X_n מתכנסת ל X <u>**בהסתברות**</u> אם לכל $\varepsilon>0$ מתקיים:

$$\lim_{n\to\infty}P(\underbrace{\left|X_n-X\right|>\varepsilon}_{B_n^{\varepsilon}})=0$$

נסמן זאת: $X_n\xrightarrow{P}X$

<u>הגדרה</u>

נאמר ש X_n מתכנסת ל X <u>**בהתפלגות**</u> אם לכל $a\in\mathbb{R}$:

$$\lim_{n\to\infty}F_{X_n}(a)=F_X(a)$$

(בכל נקודת רציפות של F_X)

סימון: $X_n\xrightarrow{D}X$

הערה: במקרה זה המשתנים המקריים לא חייבים לחיות באותו מרחב הסתברות

<u>טענה:</u>

יש יחס היררכי בין ההתכנסויות באופן הבא:

התכנסות כמעט תמיד $\Leftarrow$ בהסתברות $\Leftarrow$ בהתפלגות

התכנסות בממוצע ריבועים $\Leftarrow$ בהסתברות $\Leftarrow$ בהתפלגות

ואלו הגרירות ה<u>יחידות</u>.

נוכיח זאת:

<u>משפט: התכנסות כמעט תמיד גוררת התכנסות בהסתברות</u>

<u>הוכחה</u>: $P(\limsup_{n\to\infty}B_n^{\varepsilon})=0$ על פי הנתון, ומאי-שוויון פאטו ל limsupים נקבל כי:

$$\limsup_{n\to\infty}P(B_n^{\varepsilon})\le P(\limsup_{n\to\infty}B_n^{\varepsilon})=0$$

ולכן גם אגף שמאל (שהוא אי-שלילי) הוא אפס לכל ε וזו בדיוק ההגדרה של התכנסות בהסתברות.

<u>משפט: התכנסות בממוצע ריבועים גוררת התכנסות בהסתברות</u>

<u>הוכחה</u>: יהא $\varepsilon>0$ נתון, ואז:

$$P(|X_n - X| > \varepsilon) = \underbrace{P((X_n - X)^2 > \varepsilon^2) \le \frac{E[X_n - X]^2}{\varepsilon^2}}_{Markov}$$

ואגף ימין שואף ל 0 מהתכנסות בממוצע ריבועים ולכן גם אגף שמאל כך וקיבלנו את המבוקש.

<u>טענה: התכנסות בממוצע ריבועים אינה גוררת התכנסות כמעט תמיד</u>

מספיק לתת דוגמה נגדית:

$$\underbrace{X_n(\omega)}_{\text{ברנולי בלתי תלויים}} = \begin{cases} 0 & probability\ 1 - \frac{1}{n} \\ 1 & probability\ \frac{1}{n} \end{cases}$$

נראה ש $X_n \xrightarrow{m.s.} 0$:

$$E[X_n - 0]^2 = \frac{1}{n} \underset{n\to\infty}{\to} 0$$

נראה ש לא מתקיימת התכנסות כמעט תמיד:

צ"ל קיים $\varepsilon > 0$ כך ש $P(\limsup_{n\to\infty} B_n^{\varepsilon}) > 0$

$$B_n^{\varepsilon} = \{\omega : X_n(\omega) > \varepsilon\}$$

הלמה השניה של בורל קנטלי אומרת שם A_n מאורעות בת"ל וגם $\sum P(A_n) = \infty$ ואז $P(\limsup_{n\to\infty} A_n) = 1$

נבחר אם כן $\varepsilon = \frac{1}{2}$

$$P(\underbrace{|X_n - 0|> \tfrac{1}{2}}_{B_n^{\frac{1}{2}}}) = \frac{1}{n}$$

הראינו שאם יש התכנסות כמעט תמיד – אזי היא לגבול אחר מ-0, ומהגרירות של התכנסות בהסתברות היינו מקבלים שההתכנסות בהסתברות היא לשני הגבולות הללו (לפחות). קל לראות שאם מתקיימת התכנסות בהסתברות אזי הגבול הוא יחיד, ולכן זו סתירה. מכאן נובע התכנסות בממוצע ריבועים אינה גוררת התכנסות כמעט תמיד.

<u>טענה: התכנסות כ"ת אינה גוררת התכנסות בממוצע ריבועים</u>

<u>הוכחה</u>: ע"י דוגמה נגדית גם כאן:

$$\underbrace{X_n(\omega)}_{\text{בלתי תלויים}} = \begin{cases} 0 & probability\ 1 - \frac{1}{n^2} \\ n^3 & probability\ \frac{1}{n} \end{cases}$$

$X_n \xrightarrow{a.s} 0$:

עבור n מספיק גדול: $P(B_n^{\varepsilon}) = \frac{1}{n^2}$ ולכן:

$$\sum_{n=1}^{\infty} P(B_n^{\varepsilon}) < \infty$$

$$\Rightarrow P(\limsup_{n\to\infty} B_n^{\varepsilon}) = 0$$

ולכן התכנסות כמעט תמיד.

מצד שני:

$$E[X_n - 0]^2 = \frac{1}{n^2} \cdot n^6 = n^4$$

ולכן X_n אינו מתכנס בממוצע ריבועים ל 0.

<u>משפט: התכנסות בהסתברות גוררת התכנסות בהתפלגות</u>

<u>הוכחה:</u>

יהא $a\in\mathbb{R}$ ו $\varepsilon>0$ ונתבונן ב $F_{X_n}(a)=P(X_n\le a)$

מחוק ההסתברות השלמה:

$$P(X\le a\,,\,X\le a+\varepsilon)+P(X_n\le a\,,\,X>a+\varepsilon)=P(X_n\le a|X\le a+\varepsilon)P(X\le a+\varepsilon)+P(X_n\le a\,,\,X>a+\varepsilon)$$
$$\le\underbrace{P(X\le a+\varepsilon)}_{F_X(a+\varepsilon)}+P(|X_n-X|>\varepsilon)$$

באופן דומה נסיק:

$$F_X(a-\varepsilon)\le F_{X_n}(a)+P(|X_n-X|>\varepsilon)$$

מכיוון שההסתברות באגף ימין היא חיוביות נעביר אגף ונקבל:

$$F_X(a-\varepsilon)-P(|X_n-X|>\varepsilon)\le F_{X_n}(a)\le F_X(a+\varepsilon)+P(|X_n-X|>\varepsilon)$$

נשאיף את n לאינסוף ונקבל:

$$F_X(a-\varepsilon)\le\liminf_{n\to\infty}F_{X_n}(a)\le\limsup_{n\to infinty}F_X X_n(a)\le F_X(a+\varepsilon)$$

וכעת נשאיף את ε לאפס ונקבל את הטענה מסנדוויץ'.

<u>שיעור 28</u>

<u>חוק המספרים הגדולים</u>

יהיו משתנים מקריים X_i שהם iid ואז **<u>ממוצע רץ</u>** יהיה:

$$Y_n=\frac{X_1+...+X_n}{n}$$

נניח של X_i התוחלת היא מספר סופי כלשהו μ (כמובן היא שווה עבור כולם).

<u>החוק *החלש* של המספרים הגדולים</u>

$$Y_n\xrightarrow{P}\mu$$

$$\lim_{n\to\infty}P(|Y_n-\mu|>\varepsilon)=0$$

אם גם מניחים ש $Var[X]<\infty$ ההוכחה מיידית ומסתמכת על אי-שוויון צ'בישב, אחרת: עדיין נכון אבל צריך לעבוד.

<u>החוק *החזק* של המספרים הגדולים</u>

$$Y_n\underset{a.s.}{\rightarrow}\mu$$

$$P\left(\lim_{n\to\infty}Y_n(\omega)=Y(\omega)\right)=1$$

וכבר אמרנו שזה שקול ל:

$$P\left(|Y_n(\omega)-Y(\omega)|>\varepsilon\text{ ``}i.o\text{''}\right)=0$$

אם מניחים שהמומנט הרביעי סופי, כלומר $E[X^4]<\infty$ ההוכחה "יחסית פשוטה".

<u>משפט הגבול המרכזי CIT -Central Limit Theorem</u>

אם משתנים X_i שהם iid כמקודם עם תוחלת 0 ושונות 1 לכולם.

$$\frac{X_1+...+X_n}{\sqrt{n}}\xrightarrow{D}N(0,1)$$

נגדיר:

$$Y_n = X_1 + ... + X_n$$
$$E[Y_n] = 0$$
$$Var[Y_n] = n$$
$$Var\left[\frac{Y_n}{\sqrt{n}}\right] = 1$$

ראינו שכאשר מוציאים גורם שכופל את המשתנה המקרי מתוך השונות – צריך להעלותו בריבוע, ולכן יש בפנים שורש ולא ממש n. למעשה זה הפקטור הדרוש כדי לנרמל את השונות.

כל משתנה קל להפוך לסטנדרטי, כי אם נחסר ממנו את התוחלת שלו נקבל מ"מ חדש עם תוחלת אפס, ואם נחלק בשורש השונות נקבל מ"מ נורמלי סטנדרטי: $\frac{X - E[X]}{\sqrt{Var[X]}}$

המשך הסבר למשפט, מה משמעותו?

$$\lim_{n\to\infty} P\left(a < \frac{X_1 + ... + X_n}{\sqrt{n}} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-x^2}{2}} dx$$

<u>הערה</u>: משפט דה-מואבר לפלאס שלמדנו הוא מקרה פרטי.

<u>הוכחה</u>

רמאות קטנה - נסתמך על העובדה (שלא נוכיח) שאם לכל t מתקיים:

$$M_{X_n}(t) \to M_X(t)$$

אז

$$X_n \xrightarrow{D} X$$

(כלומר התכנסות טרנס' לפלס גורם התכנסות פונקציות)

הפונקציה יוצרת המומנטים של המשתנה הנורמלי הסטנדרטי היא $e^{\frac{t^2}{2}}$ ואם אנחנו מקבלים את העובדה הנ"ל מספיק להוכיח:

$$M_{\frac{Y_n}{\sqrt{n}}}(t) \underset{n\to\infty}{\to} e^{\frac{t^2}{2}}$$

אם Y_n היא הסכום, למדנו שפונקציה יוצרת מומנטים של סכום משתנים מקריים בת"ל היא מכפלת הפונקציות ולכן:

$$M_{Y_n}(t) = [M_{X_1}(t)]^n$$

(כי המשתנים בסכום הם בעלי אותן פונקציות יוצרות מומנטים)

$$a > 0$$
$$F_{X/a}(x) = P(X/a \le x) = P(X \le ax) = F_X(ax)$$
$$f_{X/a}(x) = \frac{d}{dx} F_{X/a}(x) = a\, f_X(ax)$$
$$M_{X/a}(t) = \int_{\mathbb{R}} e^{xt} f_{X/a}(x)\, dx = a \int e^{\frac{axt}{a}} f_X(ax)\, dx \frac{a}{a} = M_X\left(\frac{t}{a}\right)$$

ולכן:

$$M_{\frac{Y_n}{\sqrt{n}}}(t) = M_{Y_n}\left(\frac{t}{\sqrt{n}}\right) = \left[M_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

נזכור כי הנגזרות של הפונקציה יוצרת המומנטים באפס הם המומנטים, ולכן נפתח טור טיילור סביב אפס:

$$= \left[M_{X_1}(0) + \frac{t}{\sqrt{n}} M_{X_1}'(0) + \frac{1}{2}\frac{t^2}{n} M_{X_1}''(0) + \frac{1}{6}\frac{t_3}{n^{3/2}} M_{X_1}'''(\xi)\right]^n$$

$$M_{Y_n/\sqrt{n}}=\left[1+\frac{t^2}{2n}+\frac{1}{6}\frac{t^3}{n^{3/2}}M_{X_1}(\xi)\right]^n \underset{n\to\infty}{\rightarrow} e^{\frac{t^2}{2}}$$

כאשר ה ξ המזועז הוא שארית טיילור, וזה מתאפס בגבול כי במכנה ה n שואף לאינסוף.

❑

<u>דוגמה מהחיים</u>

יש נסיונאי שעושה ניסוי, ובגלל כל מיני רעשים במערכת התוצאה היא מקרית, רוצה לשערך את התוחלת.

$$Y_n=\frac{X_1+\ldots+X_n}{n}$$

כמה פעמים עליו לחזור על הניסוי באופן ב"ת על מנת שבהסתברות לפחות 95% הסטיה מהתוחלת תהיה פחות מ $\frac{\sigma}{4}$

$$P\left(\mu-\frac{\sigma}{4}<\frac{1}{n}\sum^n X_i<\mu+\frac{\sigma}{4}\right)>0.95$$

$$\frac{-\sigma}{4}<\frac{1}{n}\sum(X_i-\mu)<\frac{\sigma}{4}$$

$$P\left(\frac{-\sqrt{n}}{4}<\frac{1}{\sqrt{n}}\sum_i\frac{X_i-\mu}{\sigma}<\frac{\sqrt{n}}{4}\right)>0.95$$

$$\int_{\frac{-\sqrt{n}}{4}}^{\frac{\sqrt{n}}{4}} e^{\frac{-y^2}{2}}dy>0.95 \quad\Rightarrow\quad n=62$$

ניתן היה לענות גם משיקולים של א"ש צ'בישב:

$$P\left(\left|\frac{X_1+\ldots+X_n}{n}-\mu\right|>\frac{\sigma}{4}\right)<0.05$$

$$P\left(\left|\frac{X_1+\ldots+X_n}{n}-\mu\right|>\frac{\sigma}{4}\right)\le\frac{16\cdot Var[\frac{X_1+\ldots+X_n}{n}-\mu]}{\sigma^2}=\frac{16\sigma^2}{n\sigma^2}=\frac{16}{n}$$

$$n>\frac{16}{0.05}\Rightarrow 320$$

ואכן ניתן לעשות זאת עם א"ש צ'בישב – אבל ההערכה הרבה יותר גסה.

<u>שיעור 29</u>

<u>שרשראות מרקוב (Markov)</u>

S קבוצה סופית או בת מניה, נכנה אותה <u>**מרחב המצבים**</u>

נסמן את איבריה באינדקסים i, j, i_5, j_{19}

<u>**שרשרת מרקוב מעל S**</u> היא סדרה של מ"מ $X_1, X_2, \ldots$ כך ש $X_i\in S$ לכל i, ומתקיים:

$$P(X_n=i_n \mid X_1=i_1, X_2=i_2, \ldots, X_n=i_n)=P(X_n=i_n \mid X_{n-1}=i_{n-1})$$

במילים אחרות – השאלה על ההסתברות המותנה של המשתנה ה n-י נקבעת ע"י המשתנה שלפניו, ולא תלויה במשתנים הקודמים.

הערות

1. שרשרת מרקוב היא מקרה פרטי של תהליך אקראי (Stochastic process) [קבוצה סדורה של מ"מ]

2. מסיבות ישומיות נקרא לאינדקס n – <u>זמן</u> (במקרה שלנו הזמן הוא בדיד)
3. קיימים תהליכים אקראיים בזמן רציף, וקיימים תהליכים אקראיים שמרחב המצבים הוא רציף. (דוגמה ששניהם רציפים – תנועת בראון)
4. השרשרת נקראת הומוגנית בזמן אם הגודל $P(X_n=i_n \mid X_{n-1}=i_{n-1})$ המכונה **<u>הסתברות המעבר</u>** אינו תלוי ב n, במילים אחרות – לכל i,j קיים $p_{i,j}$ יחיד כך ש $P(X_n=i \mid X_{n-1}=j)=p_{i,j}$. המטריצה שאלו הם איבריה מכונה **<u>מטריצת המעבר</u>**.

נסמן ב P את מטריצת המעבר.

P היא מטריצה סטוכסטית:
1. $p_{i,j}\geq 0$
2. $\forall i \ \sum_{j\in S} p_{i,j}=1$

נסמן ב μ את ההתפלגות ההתחלתית $\mu_i=P(X_0=i_0)$
<u>טענה</u>: ההתפלגות ההתחלתית + מטריצת המעבר = מגדירים באופן יחיד את התפלגות התהליך.
$$P(X_n=i_n, X_{n-1}=i_{n-1}, \dots, X_0=i_0)$$

תמיד נכון ש:
$$P(A\cap B\cap C\cap\dots\cap Z)=P(A|B\cap\dots\cap Z)\times P(B|C\cap\dots\cap Z)\times\dots\times P(Y|Z)\times P(Z)$$

ולכן:
$$P(X_n=i_n, X_{n-1}=i_{n-1}, \dots, X_0=i_0)=$$
$$P(X_n=i_n|X_{n-1}=i_{n-1},\dots,X_0=i_0)\times P(X_{n-1}=i_{n-1}|X_{n-2}=i_{n-2},\dots,X_0=i_0)\times P(X_1=i_1|X_0=i_0)\times P(X_0=i_0)$$
$$=\mu_{i_0} p_{i_0,i_1}\times\dots\times p_{i_{n-2},i_{n-1}} p_{i_{n-1},i_n}$$

נסכם על כל הערכים האפשריים למשתנים $i_1,\dots,i_{n-1}$ ונקבל שהם נעלמים מהסוגריים (צריך לחשוב על זה קצת אבל כל הערכים האפשריים למשתנים אלה בעצם נותנים את הכל, כלומר ניתן להתעלם מהאילוץ שהם יוצרים על ההסתברות כי הם לא יוצרים שום אילוץ).
נקבל (אם נפעיל אלגברה לינארית ונתבונן במה שקורה לאיברי המטריצה במכפלה באגף ימין):
$$P(X_n=i_n, X_0=i_0)=\mu_{i_0}(P^n)_{i_0,i_n}$$
ומכאן לפי כללי הסתברות מותנה וכיוון ש $\mu_i=P(X_0=i_0)$:
$$P(X_n=i_n \mid X_0=i_0)=(P^n)_{i_0,i_n}$$
כלומר זו מטריצת המעבר ל n צעדים. נסמן את את איברי P^n ע"י $p^{(n)}_{i,j}$
מפני ש $P^n P^m=P^{n+m}$ אז נובע חוק ההסתברות השלמה:
$$p^{(n+m)}i,j=\sum_{k\in S} p^{(n)}_{i,k} p^{(m)}_{k,j}$$
$$p^{(0)}_{i,j}=\begin{cases}1 & j=1\\ 0 & j\neq 1\end{cases}$$

נסמן ב $\mu^{(n)}$ את ההתפלגות אחרי n צעדים :
$$\mu^{(n)}=\mu^{(0)}P^n$$

דוגמה 1:
S קבוצה סופית
נניח שיש בה 5 אובייקטים, אז התא במטריצה $p_{1,5}$ מסמל את ההסתברות למעבר בצעד אחד מ 1 ל 5.
בעצם זה גרף מכוון כאשר כל קודקוד הוא משתנה מקרי ועל כל קשת יש את ההסתברות למטרה בהנתן המקור.

דוגמה 2:
מהלך מקרי על $\mathbb{Z}^d$
עבור d=1 – מהלך מקרי על הישר, נגדיר את המטריצה כך:

$$p_{i,j}=\begin{cases}1/2 & j=i\pm 1\\ 0 & otherwise\end{cases}$$

נגדיר את הראשית להיות נקודת ההתחלה למרות שזה לא משנה כי כל העסק אינווריאנטי תחת הזזות.

עבור d=2 – מישור בנקודות עם קואורדינטות שלמות נגדיר הסתברות של רבע לכל אחד מהכיוונים.
מסיבות טכניות נגדיר שהצעדים הם לשכנים האלכסוניים ולא המאוזנים ואופקיים. זה כמובן לא משנה.

דוגמה נוספת:
עירבוב של קלפים – כל עירבוב הוא צעד.

עוד דוגמה:
בהנתן גרף **לא** מכוון וקבוצה של צבעים, אפשר להגדיר צביעה חוקית אם אין שני קודקודים המחוברים בצלע וצבועים באותו צבע. אם נתאר את המצב הבא – בוחרים בהתפלגות אחידה קודקוד, ובוחרים צבע בהתפלגות אחידה – אם זה מגדיר צביעה חוקית – משנים את המצב בהתאם.

מה השאלות המעניינות על שרשראות מרקוב?
ראינו ש $\mu^{(n)}=\mu^{(0)}P^n$
האם כאשר n שואף לאינסוף יש התכנסות? זה נקרא <u>התנהגות השרשרת בזמנים ארוכים</u>.
כל עוד אנחנו בעולם סופי – זה באמת אלגברה לינארית כי זה כפל סופי של מטריצות וזה מוגדר ע"י משפט פרובניוס-פרון, (Frobenius -Perron) .

<u>סימון:</u>

$$f_{i,j}^{(n)}=P(X_n=j, X_{n-1}\neq j, \dots X_1\neq j | X_0=i)$$

ההסתברות להגיע בזמן n למצב j בפעם הראשונה בהנתן שהתחלנו במצב i
אם נסכם על כל ה n-ים האפשריים: $\sum_{n=1}^{\infty} f_{i,j}^{(n)}$ זו ההסתברות להגיע למצב j בסופו של דבר, אם התחלנו ב i. נסמן זאת ע"י $f_{i,j}$.

<u>הגדרה</u>
מצב j יקרה נשנה (recurrent) אם $f_{j,j}=1$, כלומר – אם התחלתי ב j – בוודאות אחזור ל j בסופו של דבר.
אם $f_{j,j}<1$ המצב נקרא חולף (transient).

לדוגמה – המהלך המקרי – האם בוודאות אחזור לראשית? או האם אחזור למקום שכבר הייתי בו באופן כללי?

נניח ש $X_0=i$ מהי ההסתברות שהביקור הראשון ב j יהיה בזמן n_1 , הביקור השני בזמן n_2 והביקור ה k בזמן n_k . סימנו שההסתברות לביקור הראשון ב j בזמן n_1 היא $f_{i,j}^{(n_1)}$ ולכן התשובה תהיה:

$$f_{i,j}^{(n_1)} f_{j,j}^{(n_2-n_1)} f_{j,j}^{(n_3-n_2)} \dots f_{j,j}^{(n_k-n_{k-1})}$$

מהי ההסתברות ללפחות k ביקורים ב-j בהנתן שהתחלנו ב i?
התשובה היא לסכם על כל הערכים האפשריים של הסדרה n, כי כך בעצם נקבל את ההסתברויות עבור k ביקורים, כאשר בדרך עושים פרמוטציות על כל "זמני" הביקורים האפשריים – וזה יתן $f_{i,j}\cdot f_{j,j}^{k-1}$ וזה גם מאד אינטואיטיבי כי זו ההסתברות לבקר ב j בהנתן שהתחלנו ב i ואז שביקרנו בו שוב ואז שוב, כך k-1 פעמים – סך הכל – k ביקורים ב j.

מהי ההסתברות לבקר ב j אינסוף פעמים?

$$P(X_n=k \quad "i.o"|X_0=i)=\begin{cases} f_{i,j} & f_{j,j}=1 \\ 0 & f_{j,j}<1 \end{cases}$$

בפרט אם ניקח i=j נקבל:

$$P(X_n=i \quad "i.o"|X_0=i)=\begin{cases} 1 & f_{i,i}=1 \\ 0 & f_{i,i}<1 \end{cases}$$

כלומר – הישנות אחת המובטחת בוודאות – מבטיחה אינסוף הישנויות...

אבל איך מזהים מקרים אלו?

<u>משפט</u>

<u>נשנות</u>

$$f_{j,j}=1 \Leftrightarrow P(X_n=j \quad "i.o" \quad |X_0=j)=1 \quad \Leftrightarrow \sum_{n=1}^{\infty} p_{j,j}^{(n)}=\infty$$

<u>חולפות</u>

$$f_{j,j}<1 \quad \Leftrightarrow \quad P(X_n=j \quad "i.o" \quad |X_0=j)=0 \quad \Leftrightarrow \sum_{n=1}^{\infty} p_{j,j}^{(n)}<\infty$$

הראינו את הגרירה הראשונה ימינה בשני המקרים – לכן מספיק להראות את כל הגרירות דו"צ באחד המקרים – וזה יתן לנו את השני (כך רז טוען, אפשר לבדוק את הלוגיקה שלו אבל נסתפק באמונה עיוורת בשלב זה).

<u>הוכחה</u> (עבור חולפות)

ראינו ש $f_{j,j}<1 \quad \Leftarrow \quad P(X_n=j \quad \text{i.o}|X_0=j)=0$

אבל $P_{j,j}^{(n)}=P(X_n=j|X_0=j)$

ומהלמה הראשונה של בורל-קנטלי נובע שאם הטור מתכנס אז $P(X_n=j \quad "i.o" \quad |X_0=j)=0$

נתבונן ב $p_{i,j}^{(n)}$ ונקבל שזוהי סכימה של כל האפשרויות להגיע ל j בסופו של דבר.

נשים לב ש:

$$p_{i,i}^{(n)}=\sum_{s=1}^{n} f_{i,i}^{(s)} p_{i,i}^{(n-s)}$$

כלומר ההסתברות להתחיל ב i ולסיים ב i בזמן n כוללת את כל האפשרויות לעזוב את i בנקודה כלשהי ולחזור אליה אחר כך – עם ההסתברות להמשך. זה מכונה first passage time decomposition, וזה טריק חשוב.

ואז:

$$\sum_{n=1}^{\infty} p_{i,i}^{(n)}=\sum_{(n=1)}^{\infty} \sum_{s=1}^{n} f_{i,i}^{(s)} p_{i,i}^{(n-s)}$$

$$=\sum_{(s=1)}^{\infty} \sum_{n=s}^{\infty} f_{i,i}^{(s)} p_{i,i}^{(n-s)}=\sum_{(s=1)}^{\infty} f_{i,i}^{(s)} \sum_{n=0}^{\infty} p_{i,i}^{(n)}$$

$$\Rightarrow \sum_{n=1}^{m} p_{i,i}^{(n)}=f_{i,i} \sum_{n=0}^{m} p_{i,i}^{(n)}$$

$$\Rightarrow \sum_{n=1}^{m} p_{i,i}^{(n)}=f_{i,i}+f_{i,i} \sum_{n=1}^{m} p_{i,i}^{(n)}$$

$$(1-f_{i,i}) \sum_{n=1}^{m} p_{i,i}^{(n)}=f_{i,i}$$

נהיתה קצת סמטוכה על הלוח ואני תקווה שמה שכתוב כאן נכון.

שיעור 30

דוגמה

מהלך מקרי על $\mathbb{Z}^d$

נתחיל ב $d=1$

נשאל האם נקודת ההתחלה היא מצב חולף או נשנה.

בעבר הרחוק ראינו שהטלת 2n מטבעות – הסיכוי לקבל בדיוק n מכל סוג היא אסימפטוטית בערך $\frac{1}{\sqrt{\pi n}}$ ולכן

$$p_{0,0}^{(n)} \sim \frac{1}{\sqrt{\pi n}}$$

עבור n מספיק גדול קיימים קבועים c, C כך ש:

$$\frac{c}{\sqrt{\pi n}} < p_{0,0}^{(n)} < \frac{C}{\sqrt{\pi n}}$$

$$\sum_{n=1}^{\infty} p_{0,0}^{(n)} = \infty$$

ואז נקודת הראשית היא מצב נשנה ולכן נחזור אליה אינסוף פעמים.

נסתכל עבור d=2 כלומר מהלך מקרי במישור, נניח שאנחנו מטיילים לאלכסונים, ולכן ניתן לחשוב על זה כהטלה של שני מטבעות, אחד יקבע האם הולכים למעלה ולמטה ואחד האם ימינה או שמאלה, ולכן באותו אופן $p_{0,0}^{(n)} \sim \frac{1}{\pi n}$

וזה עדיין טור מתבדר – ולכן הראשית היא עדיין נקודה מצב נשנה.

$d=3$ באותו אופן נקבל $p_{0,0}^{(n)} \sim \frac{1}{(\pi n)^{\frac{3}{2}}}$ אבל זה כבר טור מתכנס, ועל כן הראשית היא מצב חולף בשלושה מימדים.

מוזר!

הגדרה

שרשרת מרקוב נקראת אי-פריקה (irreducible) אם לכל i,j קיים n כך ש $p_{i,j}^{(n)} > 0$, במילים אחרות – כל מצב אפשרי בהסתברות חיובית.

משפט – בשרשרת מרקוב אי-פריקה מתקיים אחת משתיים:

1. כל המצבים נשנים וגם $P\left(\bigcap_{j=1}^{\infty}\{X_n=j \ \ \text{i.o.}\}: X_0=i\right)=1$ וגם $\sum_{n=1}^{\infty} p_{i,j}^{(n)}=\infty$
2. כל המצבים חולפים וגם $P\left(\bigcup_{j=1}^{\infty}\{X_n=j \ \ \text{i.o.}\}: X_0=i\right)=0$ וגם $\sum_{n=1}^{\infty} p_{i,j}^{(n)}<\infty$

נתמקד בעובדה שנשנות ופריקות היא תכונה של כל השרשרת בהנתן אי-פריקות:

הוכחה

לכל i,j קיימים r,m כך ש $p_{j,i}^{(r)}>0 \quad p_{i,j}^{(n)}>0$ ולכן:

$$p_{i,i}^{(m+n+r)} \geq p_{i,j}^{(m)} p_{j,j}^{(n)} p_{j,i}^{(r)}$$

נניח שהמצב j הוא נשנה, ולכן הטור שלו מתבדר ונקבל כי:

$$\sum_{n=1}^{\infty} p_{i,i}^{(m+n+r)} \geq p_{i,j}^{(m)} p_{j,i}^{(r)} \sum_{n=1}^{\infty} p_{j,j}^{(n)}$$

וגם i מצב נשנה.

בשרשרת אי-פריקה – נשנות/חולפות משותפות לכל המצבים.

נזכיר כי $\mu_i^{(n)} = \sum_j \mu_j^{(0)} P_{j,i}^{(n)}$ האם קיים גבול כאשר n שואף לאינסוף?

הרעיון הוא שבהנתן שרשרת מרקוב – נניח שאנחנו בנקודת התחלה כלשהי, ונראה כעבור n ניסויים מה ההסתברות שלנו

להיות בכל מקום? הטענה היא שכאשר n שואף לאינסוף – לא משנה איפה התחלנו, חלוקת ההסתברויות תתקבל.
לדוגמה נניח שיש שני מצבים 1 ו 2 כאשר ההסתברות לעבור מ 1 לעצמו היא $1-p$ ול-2 היא p.
באותו אופן ההסתברות לעבור מ 2 לעצמו היא $1-q$ ומ 2 ל 1 היא q.
לכן:

$$P=\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

כאשר

$$0<p,q \qquad p+q\neq 2$$

ואז:

$$(1-p-\lambda)(1-q-\lambda)=0$$

$$\lambda=-\lambda(2-p-q)+(1-p-q)=0$$

$$\alpha=1-p-q \qquad \lambda^2-\lambda(1+\alpha)+\alpha=0$$

$$-1<\alpha<1 \qquad \lambda=1,\alpha$$

$$\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=0$$

$$\begin{pmatrix} 1-p-(1-p-q) & p \\ q & 1-q-(1-p-q) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=0$$

$$\begin{pmatrix} q & p \\ q & p \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=0$$

$$\frac{1}{\sqrt{p^2+q^2}}\begin{pmatrix} p \\ -q \end{pmatrix}$$

$$P^n=S\,\Lambda^n\,S^{-1}$$

$$\Lambda^n=\begin{pmatrix} 1 & 0 \\ 0 & \alpha^n \end{pmatrix}$$

$$S=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{p}{\sqrt{p^2+q^2}} \\ \frac{1}{\sqrt{2}} & \frac{-q}{\sqrt{p^2+q^2}} \end{pmatrix}$$

$$\lim_{n\to\infty} P^n_{i,j}=\sum_{k,l} S_{i,k}\,\Lambda^{\infty}_{k,l} S^{-1}_{l,j}=S_{i,1}\,S^{-1}_{i,j}=\begin{pmatrix} q & p \\ q & p \end{pmatrix}$$

$$\lim_{n\to\infty} \mu^{(n)}=(\mu_1^{(0)}\mu_2(0))\begin{pmatrix} q & p \\ q & p \end{pmatrix}\frac{1}{p+q}=\frac{q\,,p}{p+q}$$

כלומר בשרשרת מרקוב אי-פריקה ולא מחזורית (לא הגדרנו את זה) לא משנה איך מתחילים – ההסתברות להיות במצב 1 היא $\frac{q}{p+q}$ וההסתברות להיות במצב 2 היא $\frac{p}{p+q}$, כלומר יש איזושהי התפלגות אשר בחלוף הזמן, בהתעלם מהמצב ההתחלתי, שואפים אליה.