סיכום הסתברות

הגדרות ראשונות

הגדרה: (מרחב מדגם)

מרחב מדגם היא קבוצה Ω (סופית או אינסופית) של כל תוצאות הניסוי האפשריות.

הגדרה: (פונקציית הסתברות נקודתית)

פונקציית הסתברות נקודתית היא $p: \Omega \to [0,1]$, המקיימת:

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

הגדרה: (מאורע)

מאורע הוא תת-קבוצה של Ω. כאשר $\mathcal{F}$ הינה קבוצת כל המאורעות, $\mathcal{F} = 2^{\Omega}$.

הגדרה: (פונקציית ההסתברות)

בהינתן p, נגדיר את $\mathbb{P}: \mathcal{F} \to [0,1]$ להיות פונקציית ההסתברות, $\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)$ כאשר A היא תת-קבוצה של Ω. $\mathbb{P}$ מקיימת:

(i): $\mathbb{P}(\Omega) = 1$

(ii): σ-אדיטיביות: אם $A_1, A_2, \dots$ היא סדרת מאורעות זרים בזוגות ($\forall i \neq j \;\; A_i \cap A_j = \emptyset$) אז:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

הגדרה: (מרחב הסתברות)

ל $(\Omega, F, \mathbb{P})$ נקרא <u>מרחב ההסתברות</u>.

פונקציית ההסתברות הנקודתית מוגדרת על ידי: $p(\omega) = \mathbb{P}(\{\omega\})$.

אם Ω סופית אז המרחב $(\Omega, F, \mathbb{P})$ כאשר $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$ נקרא <u>מרחב הסתברות אחיד</u> ול- $\mathbb{P}$ נקרא פונקציית הסתברות אחידה.

טענה: (תכונות של פונקציית הסתברות – לא בטוח שצריך אבל הוכחות קלילות

יהי Ω מרחב מדגם ותהי $\mathbb{P}$ פונקציית הסתברות. מתקיימות התכונות הבאות:

(i): $\mathbb{P}(\emptyset)=0$

<u>הוכחה:</u>

נגדיר $A_i=\emptyset$ לכל i. מתכונת האדיטיביות, מתקיים:

$$\mathbb{P}(\emptyset)=\mathbb{P}\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}\mathbb{P}(A_i)=\mathbb{P}(\emptyset)+\mathbb{P}(\emptyset)+\cdots=0$$

(ii): אדיטיביות סופית: אם $A_1,\dots,A_n$ מאורעות זרים, מתקיים:

$$\mathbb{P}\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mathbb{P}(A_i)$$

<u>הוכחה:</u> נגדיר סדרת מאורעות B_i ע"י $B_i=A_i$ עבור $1\le i\le n$, ו-$B_i=\emptyset$ עבור $i\ge n+1$.

מאדיטיביות, נקבל שמתקיים:

$$\mathbb{P}\left(\bigcup_{i=1}^{n}A_i\right)=\mathbb{P}\left(\bigcup_{i=1}^{\infty}B_i\right)=\sum_{i=1}^{\infty}\mathbb{P}(B_i)=\sum_{i=1}^{n}\mathbb{P}(A_i)$$

(iii): מונוטוניות: אם $A\subset B$ אז $\mathbb{P}(A)\le\mathbb{P}(B)$

<u>הוכחה:</u> מ-(ii) ומכיוון ש-$B=A\cup(B\setminus A)$ נקבל שמתקיים:

$$\mathbb{P}(B)=\mathbb{P}(A)+\underbrace{\mathbb{P}(B\setminus A)}_{\ge 0}\quad\Rightarrow\quad\mathbb{P}(B)\ge\mathbb{P}(A)$$

(iv): לכל מאורע A מתקיים $\mathbb{P}(A)\le 1$

<u>הוכחה:</u> מכיוון ש-$A\subset\Omega$, מתכונה (iii) נקבל: $\mathbb{P}(\mathrm{A})\le\mathbb{P}(\Omega)=1$

(v): לכל מאורע A מתקיים $\mathbb{P}(A)+\mathbb{P}(A^c)=1$, כאשר $A^c=\Omega\setminus A$.

<u>הוכחה:</u> מ-(ii) עבור A,A^c (מאורעות זרים) מתקיים:

$$\mathbb{P}(A)+\mathbb{P}(A^c)=\mathbb{P}(A\cup A^c)=\mathbb{P}\big(A\cup(\Omega\setminus A)\big)=\mathbb{P}(\Omega)=1$$

משפט: (אי-שוויון בול)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות. הסתברות של מאורעות במרחב היא סיגמה תת-חיבורית.

כלומר, לכל אוסף בן מניה של מאורעות $\{A_n\}_{n\in\mathbb{N}}$ מתקיים:

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty}A_k\right)\le\sum_{k=1}^{\infty}\mathbb{P}(A_k)$$

התפלגויות בינומית ומולטינומית:

הגדרה – התפלגות בינומית:

קבוצת הסתברויות $\{B(k,n,p)\}_{k=0}^{n}$ כאשר:

$$B(k,n,p) \overset{def}{=} \binom{n}{k} p^k q^{n-k} \ ; \quad 0 \le p \le 1 \quad ; q = 1-p$$

נקראת התפלגות בינומית.

הגדרה – התפלגות מולטינומית:

נגדיר $\Omega = \{\omega = (a_1, \dots, a_n) | a_i = b_1, \dots, b_m\}$ כאשר $b_1, \dots, b_m$ הם מספרים נתונים.

$$p(\omega) \overset{def}{=} p_1^{\zeta_1(w)} \dots p_m^{\zeta_m(\omega)} \ ; \ p_1 + \cdots + p_m = 1 \ ; \ p_1, \dots, p_m \ge 0$$

כאשר ζ_i הינו מספר האיברים ברשימה $(a_1, \dots, a_n)$ אשר שווים ל-b_i.

$$\sum_{\omega \in \Omega} p(\omega) = \sum_{\substack{n_1 \ge 0, \dots, n_m \ge 0 \\ n_1 + \cdots + n_m = n}} \binom{n}{n_1, \dots, n_m} p_1^{n_1} p_2^{n_2} \dots p_m^{n_m} = (*)$$

כאשר $\binom{n}{n_1, \dots, n_m}$ הינו מספר הפתרונות של המשוואה $n = n_1 + \cdots + n_m$.

ומתקיים:

$$(*) = (p_1 + \cdots + p_n)^n = 1$$

(ולכן פונקצית ההסתברות מוגדרת היטב.)

נגדיר את המאורע $A_{n_1,\dots,n_m}$ על ידי:

$$A_{n_1,\dots,n_m} = \left\{ \omega = (a_1, \dots, a_n) \middle| \begin{matrix} \zeta_1(\omega) = n_1 \\ \zeta_2(\omega) = n_2 \\ \vdots \\ \zeta_m(\omega) = n_m \end{matrix} \right\}$$

וכעת:

$$\mathbb{P}(A_{n_1,\dots,n_m}) = \binom{n}{n_1, \dots, n_m} p_1^{n_1} \dots p_m^{n_m}$$

קבוצת ההסתברויות $\left\{ \binom{n}{n_1, \dots, n_m} p_1^{n_1} \dots p_m^{n_m} \right\}$ כאשר $n_1, \dots, n_m \ge 0$, $n_1 + \cdots + n_m = n$ נקראת ההתפלגות המולטינומית.

כל מיני סוגים של m מאורעות מתוך n.

משפט: (עקרון ההכלה וההדרה – כשרוצים שלפחות אחד מהמאורעות יתקיים)

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות, $A_1, \dots, A_n \in \mathcal{F}$. עבור:

$$S_k \overset{def}{=} \sum_{1 \le l_1 < \cdots < l_k \le n} \mathbb{P}\left(A_1 \cap \dots \cap A_{l_k}\right) \quad ; i \le k \le n$$

אזי מתקיים: $\mathbb{P}(A_1 \cup \dots \cup A_n) = S_1 - S_2 + \cdots + (-1)^{n-1} S_n = (*)$

הוכחה:

1\\ מספיק להוכיח כי כל מאורע $\omega \in \Omega$ נותן את אותה התרומה לשני אגפי המשוואה $(*)$.

2\\ נניח כי מתקיים $\omega \notin A_1, \dots, \omega \notin A_n$ אזי, $\omega \notin A_1 \cup \dots \cup A_n$.

לכן ω אינו תורם ל $\mathbb{P}(A_1 \cup \dots \cup A_n)$. כמו כן ברור כי ω אינו תורם ל S_k. (עבור כל $1 \le k \le n$). אזי התרומה של ω לשני אגפי המשוואה $(*)$ היא אפס.

3\\ נניח כי $\omega \in A_j$ אבל $\omega \notin A_1, \dots, \omega \notin A_{j-1}, \omega \notin A_{j+1}, \dots, \omega \notin A_n$.

תרומה של ω לביטוי $\mathbb{P}(A_1 \cup \dots \cup A_n)$ היא $\mathbb{P}(\{\omega\})$.

$S_1 = \sum_{k=1}^{n} \mathbb{P}(A_k)$, אזי ω תורמם $\mathbb{P}(\{\omega\})$ ל- S_1.

$S_2 = \sum_{\le l_1 < l_2 \le n}^{n} \mathbb{P}\left(A_{l_1} \cap A_{l_2}\right)$, $\omega \in A_{l_1} \cap A_{l_2}$ עבור כל $l_1 \neq l_2$ מתקיים כי ω אינה תורמת ל-S_2.

ובאותו אופן ω אינה תורמת ל- $S_3, \dots S_n$.

4\\ נניח כי ω שייכת לקבוצות $A_{l_1}, \dots, A_{l_k}$ ו- $\omega \notin A_j$ עבור $j \neq l_1, \dots, l_k$.

התרומה של ω לביטוי $\mathbb{P}(A_1 \cup \dots \cup A_n)$ היא $\mathbb{P}(\{\omega\})$.

התרומה של ω ל S_1 היא: $\binom{k}{1}\mathbb{P}(\{\omega\}) = k \cdot \mathbb{P}(\{\omega\})$.

התרומה של ω ל S_2 היא: $\binom{k}{2}\mathbb{P}(\{\omega\})$.

התרומה של ω ל S_3 היא: $\binom{k}{3}\mathbb{P}(\{\omega\})$.

$\vdots$

התרומה של ω ל S_k היא: $\binom{k}{k}\mathbb{P}(\{\omega\})$.

התרומה של ω ל- $S_{k+1}, \dots, S_n$ היא אפס.

כעת נשים לב כי מהבינום של ניוטון אנו יודעים כי:

$$\underbrace{\binom{k}{0}}_{=1} - \binom{k}{1} + \cdots + (-1)^k \binom{k}{k} = \left(1 + (-1)\right)^k = 0 \quad \underset{\text{נעביר אגפים}}{\Longrightarrow}$$

$$1 = \binom{k}{1} - \binom{k}{2} + \cdots + (-1)^{k-1}\binom{k}{k}$$

כעת אנו יודעים כי התרומה של ω לאגף הימני של $(*)$ הינה $\binom{k}{1}-\binom{k}{2}+\cdots+(-1)^{k-1}\binom{k}{k}$,

ומהבינום של ניוטון נסיק

$$\left[\underbrace{\binom{k}{1}-\binom{k}{2}+\cdots+(-1)^{k-1}\binom{k}{k}}_{=1}\right]\mathbb{P}(\{\omega\})=\mathbb{P}(\{\omega\})$$

לפיכך קיבלנו כי התרומה של ω שווה בשני אגפי המשוואה כנדרש.

$\blacksquare$

משפט: (בידיוק m מאורעות מתוך n)

$(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות, $A_1,\dots,A_n\in\mathcal{F}$.

ההסתברות שיתרחשו בידיוק m מאורעות מתוך $A_1,\dots,A_n$ $\equiv P_m$

טענה: $(**)$ $P_m=S_m-\binom{m+1}{m}S_{m+1}+\cdots+(-1)^{n-m}\binom{n}{m}S_n$

כאשר $S_k=\sum_{1\le l_1<\cdots<l_k\le n}\mathbb{P}\left(A_1\cap\dots\cap A_{l_k}\right)$.

משפט: (לפחות m מאורעות מתוך n)

$(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות $A_1,\dots,A_n\in\mathcal{F}$ מאורעות.

$\overline{P_m}$ – ההסתברות שלפחות m מאורעות יתרחשו:

$$\overline{P_m}=P_m+\cdots+P_n$$

כאשר $P_m=S_m-\binom{m+1}{m}S_{m+1}+\cdots+(-1)^{n-m}\binom{n}{m}S_n$

תכונות ההסתברות המותנית

הגדרה: (ההסתברות המותנית של A בהינתן B)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהי $B\in\mathcal{F}$ מאורע בעל הסתברות חיובית.

לכל מאורע $A\in\mathcal{F}$, נגדיר את ההסתברות המותנית של A בהינתן B על ידי

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

טענה: (תכונות בסיסיות של הסתברות מותנית)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהיו $A,B,D\in\mathcal{F}$ מאורעות בעלי הסתברות חיובית. אזי

(א) אם $A\subset B$ אזי $\mathbb{P}(D|A)=\mathbb{P}(D|A,B)$

(ב) כלל השרשרת $(chain-rule)$: $\mathbb{P}(A\cap B)=\mathbb{P}(B)\cdot\mathbb{P}(A|B)$

(ג) התניה חוזרת: $\mathbb{P}_D(A|B)=\mathbb{P}(A|B,D)=\frac{\mathbb{P}(A\cap B\cap D)}{P(B\cap D)}$

(ד) $\mathbb{P}(A|A)=\frac{\mathbb{P}(A\cap A)}{\mathbb{P}(A)}=1$

(ה) $\mathbb{P}(\emptyset|A)=\frac{\mathbb{P}(\emptyset\cap A)}{\mathbb{P}(A)}=\frac{\mathbb{P}(\emptyset)}{\mathbb{P}(A)}=0$

(ו) $\mathbb{P}(\Omega|A)=\frac{\mathbb{P}(\Omega\cap A)}{\mathbb{P}(A)}=\frac{\mathbb{P}(A)}{\mathbb{P}(A)}=1$

(ז) אם $A\subseteq B$,

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

(ח) יהיו B_1,B_2 מאורעות כך ש $B_1\cap B_2=\emptyset$.

$$\begin{aligned}\mathbb{P}(B_1\cup B_2|A)&=\frac{\mathbb{P}\big((B_1\cup B_2)\cap A\big)}{\mathbb{P}(A)}\\&=\frac{\mathbb{P}\big((B_1\cap A)\cup(B_2\cap A)\big)}{\mathbb{P}(A)}\underset{\text{מאורעות זרים}}{=}\frac{\mathbb{P}(B_1\cap A)}{\mathbb{P}(A)}+\frac{\mathbb{P}(B_2\cap A)}{\mathbb{P}(A)}\\&=\mathbb{P}(B_1|A)+\mathbb{P}(B_2|A)\end{aligned}$$

(ט) $0\leq\mathbb{P}(B|A)\leq1$

(י) יהיו $B_1,B_2\in\mathcal{F}$.

$$\begin{aligned}\mathbb{P}(B_1\cup B_2|A)&=\frac{\mathbb{P}\big((B_1\cup B_2)\cap A\big)}{\mathbb{P}(A)}=\frac{\mathbb{P}\big((B_1\cap A)\cup(B_2\cap A)\big)-\mathbb{P}(B_1\cap B_2\cap A)}{\mathbb{P}(A)}\\&=\mathbb{P}(B_1|A)+\mathbb{P}(B_2|A)-\mathbb{P}(B_1\cup B_2|A)\end{aligned}$$

טענה: (הסתברות מותנית היא הסתברות)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהי $B\in\mathcal{F}$ מאורע בעל הסתברות חיובית.

נגדיר $\mathbb{P}_B(A)=\mathbb{P}(A|B)$. אזי $\mathbb{P}_B$ היא פונקציית הסתברות על Ω. הוכחה מהתכונות שלעיל.

נוסחת ההסתברות השלמה

טענה: (נוסחת ההסתברות השלמה)

יהי מרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. נניח כי $\Omega = A_1 \cup \ldots \cup A_n$ כאשר לכל $i \neq j$ מתקיים $A_i \cap A_j = \emptyset$ - כלומר סדרת המאורעות הינה סדרה של מאורעות זרים.

אזי לכל מאורע B מתקיים:

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

כאשר אנו מפרשים איברים באגף ימין כשווים ל-0 אם $\mathbb{P}(A_i) = 0$.

<u>הוכחה</u>: כל מאורע B נוכל לפרק לפי $B = \bigcup_{i=1}^{n}(B \cap A_i)$

ולכן, בגלל σ-אדיטיביות מתקיים:

$$\mathbb{P}(B) = \mathbb{P}\left(\bigcup_{i=1}^{n}(B \cap A_i)\right) \underset{\text{אדטיביות}}{=} \sum_{i=1}^{n} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i) \cdot \frac{\mathbb{P}(A_i)}{\mathbb{P}(A_i)}$$

$$= \sum_{i=1}^{n} \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$$

■

כלל בייס

טענה: (נוסחת בייס)

יהי $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. ויהיו $A, B \in \mathcal{F}$ שני מאורעות כך ש $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$, אזי

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

הוכחה: מידית מהגדרת ההסתברות המותנית:

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

■

משפט: (משפט בייס)

אם $A_1, A_2, \dots$ סדרה של מאורעות זרים שאיחודם הוא $\bigcup_{i=1}^{\infty} A_i = \Omega$ ו- $A_k \cap A_l = \emptyset$, אזי לכל מאורע $B \in \mathcal{F}$ עם $\mathbb{P}(B) > 0$ מתקיים עבור A_i כך ש $\mathbb{P}(A_i) > 0$:

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

הוכחה:

מנוסחת ההסתברות השלמה נקבל:

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \cdot \mathbb{P}(B|A_i)$$

מהמשפט הקודם נקבל $\mathbb{P}(A_i|B) \cdot \mathbb{P}(B) = \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)$

ומכך נסיק:

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

■

טענה: (נוסחת בייס המותנית)

יהא $(\Omega, \mathbb{P})$ מ"ה, יהיו A, B, C מאורעות כך ש $\mathbb{P}(B \cap C) > 0$ וכן $\mathbb{P}(A \cap C) > 0$ אז:

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

הוכחה: נפתח סוגריים

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

חיתוך של n מאורעות:

טענה: (נוסחת המכפלה)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות, $A_1,\dots,A_n\in\mathcal{F}$, $\mathbb{P}(A_1\cap\dots\cap A_{n-1})>0$

אזי מתקיים:

$$\mathbb{P}(A_n\cap\dots\cap A_1)=\mathbb{P}(A_n|A_1\cap\dots\cap A_{n-1})\cdot\mathbb{P}(A_{n-1}|A_1\cap\dots\cap A_{n-2})\cdot\dots\cdot\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$

<u>הוכחה:</u>

$$\mathbb{P}(A_1\cap\dots\cap A_n)=\mathbb{P}\big((A_1\cap\dots\cap A_{n-1})\cap A_n\big)=\mathbb{P}(A_n|A_1\cap\dots\cap A_{n-1})\cdot\mathbb{P}(A_{n-1}\cap\dots\cap A_1)$$

$$=\mathbb{P}(A_n|A_1\cap\dots\cap A_{n-1})\cdot\mathbb{P}(A_{n-1}|A_1\cap\dots\cap A_{n-2})\cdot\mathbb{P}(A_{n-2}\cap\dots\cap A_1)=$$

$$=\cdots=\mathbb{P}(A_n|A_1\cap\dots\cap A_{n-1})\cdot\mathbb{P}(A_{n-1}|A_1\cap\dots\cap A_{n-2})\cdot\dots\cdot\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)=\mathbb{P}(A_1)$$

כנדרש.

תכונות של אי תלות של מאורעות

הגדרה: (אי-תלות)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות. שני מאורעות $A,B\in\mathcal{F}$ יקראו בלתי-תלויים (ב"ת) אם

$$\mathbb{P}(A\cap B)=\mathbb{P}(A)\cdot\mathbb{P}(B)$$

אחרת המאורעות ייקראו תלויים.

אבחנה: (תכונות בסיסיות של אי-תלות)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות ויהיו $A,B\in\mathcal{F}$ מאורעות ב"ת בעלי הסתברות חיובית, אזי

(א) A בלתי-תלוי ב-Ω וב-$\emptyset$

- $\mathbb{P}(A\cap\emptyset)=\mathbb{P}(\emptyset)=0=\mathbb{P}(\emptyset)\mathbb{P}(A)$
- $\mathbb{P}(A\cap\Omega)=\mathbb{P}(A)=\mathbb{P}(A)\cdot 1=\mathbb{P}(A)\cdot\mathbb{P}(\Omega)$

(ב) $\mathbb{P}(A|B)=\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}=\frac{\mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(B)}=\mathbb{P}(A)$ ובאותו אופן גם $\mathbb{P}(B|A)=\mathbb{P}(B)$

(ג) $\mathbb{P}(A^c\cap B)=\mathbb{P}(A^c)\cdot\mathbb{P}(B)$ כלומר גם A^c ו-B ב"ת:

$$\mathbb{P}(A^c\cap B)=\mathbb{P}(B)-\mathbb{P}(A\cap B)=\big(1-\mathbb{P}(A)\big)\cdot\mathbb{P}(B)=\mathbb{P}(A^c)\cdot\mathbb{P}(B)$$

הגדרה: (אי-תלות של אוסף מאורעות)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות. אוסף מאורעות $\mathcal{A}\subset\mathcal{F}$ יקראו בלתי תלויים אם לכל תת-קבוצה סופית שלהם $\{A_n\}_{n\in[N]}$ מתקיים

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}}A_n\right)=\prod_{n\in\mathbb{N}}\mathbb{P}(A_n)$$

נבחין שזה אומר בפרט שלכל סדרת סימנים $\varepsilon_i\in\{1,C\}$ עבורם $A_i^1=A_i$ מתקיים:

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}}A_n^{\varepsilon_i}\right)=\prod_{n\in\mathbb{N}}\mathbb{P}\left(A_n^{\varepsilon_i}\right)$$

משתנים מקריים בדידים

הגדרה: (משתנה מקרי)

יהא $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. משתנה מקרי הוא פונקציה $X: \Omega \to \mathbb{R}$ ממרחב המדגם לממשיים.

הגדרה: $(supp)$

יהא $(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות. יהא $X: \Omega \to \mathbb{R}$ משתנה מקרי, נסמן את קבוצת כל התוצאות האפשריות שניתן לקבל ב:

$$supp\ (X) = \{x \in \mathbb{R} \mid \mathbb{P}(\{x\}) > 0\}$$

הגדרה (פונקציית התפלגות ופונקציית התפלגות מצטברת):

$(\Omega, \mathcal{F}, \mathbb{P})$ - מרחב הסתברות, כאשר Ω קבוצה סופית. $\zeta: \Omega \to \mathbb{R}$.

1. עבור $B \subseteq \mathbb{R}$ נגדיר:
$$\Phi_\zeta(B) = \mathbb{P}\{\omega: \zeta(\omega) \in B\}$$
פונקציית ההתפלגות של ζ.
2.
$$F_\zeta(x) \overset{def}{=} \mathbb{P}\{\omega: \zeta(\omega) \leq x\}$$
כאשר $x \in \mathbb{R}$.
פונקציית ההתפלגות המצטברת.

תכונות:

1. $\lim\limits_{x \to +\infty} F_\zeta(X) = 1$
2. $\lim\limits_{x \to -\infty} F_\zeta(X) = 0$
3. פונקציה $F_\zeta(X)$ היא פונקציה עולה ב- $\mathbb{R}$.

קונבולוציה:

טענה (נוסחת הקונבולוציה)

יהיו X, Y מ"מ ב"ת כך ש $Im(X) \subset \mathbb{Z}$ וגם $Im(Y) \subset \mathbb{Z}$, אזי לכל $n \in \mathbb{Z}$ מתקיים:

$$\mathbb{P}(X + Y = n) = \sum_{k=-\infty}^{\infty} \mathbb{P}(X = k) \cdot \mathbb{P}(Y = n - k)$$

התפלגויות חשובות: ברנולי, אחידה, בינומית, פואסונית, גאומטרית

הגדרה: (התפלגות ברנולי)

נאמר שמ״מ X מתפלג לפי התפלגות ברנולי עם סיכוי הצלחה p (בקיצור – ברנולי p) ונכתוב $X \sim Ber(p)$

אם $p_X(1) = p$ ו-$p_X(0) = 1 - p$

עבור מרחב הסתברות $\Omega = \{(a_1, \dots, a_n) | a_i \in \{0,1\}\}$

פונקצית התפלגות נקודתית הינה:

$$\mathbb{P}(\{\omega\}) = p^{\sum_{i=1}^{n} a_i} q^{n - \sum_{i=1}^{n} a_i}$$

משתנה ברנולי הינו משתנה מקרי $\delta: \Omega \to \mathbb{R}$ עבורו $\delta_i(\omega) = a_i$.

הגדרה: (התפלגות אחידה)

נאמר שמ״מ X מתפלג לפי התפלגות אחידה על קבוצה סופית $A \subset \mathbb{R}$ ונכתוב $X \sim Unif(A)$

אם $p_X(i) = \frac{1}{|A|} \quad : \forall i \in A$

הגדרה: (התפלגות בינומית)

נאמר שמ״מ X מתפלג לפי התפלגות בינומית על N ניסיונות עם סיכוי הצלחה p ונכתוב $X \sim Bin(N, p)$

אם לכל $n \in \{0, \dots, N\}$ מתקיים $p_X(n) = \binom{N}{n} p^n (1-p)^{N-n}$

טענה: (אפיון התפלגות בינומית)

יהיו $\{X_i\}_{i \in [N]}$ משתני ברנולי p ב״ת, אזי

$$\sum_{i \in [N]} X_i \sim Bin(N, p)$$

מסקנה: (חיבור התפלגויות בינומיות)

אם $X \sim Bin(N, p)$, $Y \sim Bin(M, p)$ ב״ת אז $X + Y \sim Bin(N + M, p)$

הגדרה: (התפלגות פואסון)

נאמר שמ״מ X מתפלג לפי התפלגות פואסון (או פואסונית) עם שכיחות λ ונכתוב $X \sim Poi(\lambda)$ אם לכל $n \in N_0$ מתקיים $p_X(n) = \frac{\lambda^n e^{-\lambda}}{n!}$

הגדרה: (התפלגות גיאומטרית)

נאמר שמ״מ X מתפלג לפי התפלגות גיאומטרית עם סיכוי הצלחה p ונכתוב $X \sim Geo(p)$ אם לכל $n \in \mathbb{N}$ מתקיים $p_X(n) = (1-p)^{n-1} p$

התפלגויות שחזרו על עצמן בתרגילם:

בעיה: (התפלגות בינומית שלילית)

יהיו $X_1, X_2, \dots, X_k \sim Geo(p)$ בלתי תלויים. נגדיר $X = X_1 + \dots + X_k$.

נשים לב כי X_i הינו מספר ההטלות בין ההצלחה ה $i-1$ לבין ההצלחה ה i.

X מ״מ המתפלג בינומית שלילית, נבנה את ההתפלגות:

חישוב ההתפלגות:

נשים לב כי $(1-p)^{n-k} \cdot p^k$ היא ההסתברות של כל סדרת ההטלות באורך n הכוללת $n-k$ כשלונות, ו- k הצלחות. $\binom{n-1}{k-1}$ הוא בדיוק מספר האפשרויות לסדר את k ההצלחות מבין n הנסיונות, באופן שהניסוי האחרון הסתיים בהצלחה. לכן נקבל:

$$p_X(n) = \mathbb{P}(X = n) = \begin{cases} \binom{n-1}{k-1}(1-p)^{n-k}p^k & n \geq k \\ 0 & n < k \end{cases}$$

חישוב התוחלת:

$$\mathbb{E}[X] = \mathbb{E}[X_1 + \dots + X_k] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_k] = \frac{k}{p}$$

כאשר בחישוב נעזרנו בלינאריות התוחלת, ובתוחלת של מ״מ מתפלג גאומטרית.

חישוב שונות:

$$Var(x) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{k} Var(X_i) + 2\sum_{l<j} cov(x_l, x_j) = \sum_{i=1}^{k} Var(X_i) + 0 = \sum_{i=1}^{k} \frac{1-p}{p^2}$$
$$= k \cdot \frac{1-p}{p^2}$$

זאת כיון שהמשתנים X_i הינם מ״מ בלתי תלויים מתפלגים גיאומטרית.

בעיה: (התפלגות היפרגאומטרית)

בכד יש N כדורים מתוכם M כחולים ו $N-M$ כדורים אדומים. מוציאים מהכד ללא החזרה n כדורים.

נסמן ב X את מספר הכדורים הכחולים במדגם.

חישוב ההתפלגות:

ראשית נשים לב $supp\ (X)=\{1,\dots,n\}$.

ניתן פתרון קומבינטורי לבעיה הנ״ל, יש $\binom{M}{k}$ אפשרויות לאוסף הכדורים הכחולים במדגם עם בדיוק k כחולים, ו $\binom{N-M}{n-k}$ לאוסף הכדורים במדגם עם בדיוק k כחולים. לכן מספר האפשרויות למדגם עם בדיוק k כדורים כחולים הינו $\binom{M}{k}\binom{N-M}{n-k}$.

כמו כן, מספר האפשרויות למדגם הינו $\binom{N}{n}$. מאחר שכל האפשרויות למדגם הן עם הסתברות שווה:

$$p_X(k)=\mathbb{P}(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

חישוב התוחלת

נגדיר $X_1,\dots,X_n$ משתנים מקרים אינדיקטורים עבור האם הכדור ה- i שהוצאנו הוא כדור כחול.

$X_1,\dots,X_k$ הינם משתנים מקריים מתפלגים ברנולי עם סיכוי הצלחה $\left(\frac{M}{N}\right)$ ולכן $\mathbb{E}[X_i]=\frac{M}{N}$.

$$\mathbb{E}[X]=\mathbb{E}[X_1+\cdots+X_k]=k\cdot\frac{M}{N}$$

חישוב שונות:

ההסתברות לקבל כדור כחול בשני הניסיונות הראשונים הינה $\frac{M(M-1)}{N(N-1)}$, ולכן $X_1,X_2\sim ber\left(\frac{M(M-1)}{N(N-1)}\right)$

$$cov(X_1,X_2)=\mathbb{E}[X_1X_2]-\mathbb{E}[X_1]\cdot\mathbb{E}[X_2]=\frac{M(M-1)}{N(N-1)}-\frac{M^2}{N^2}$$

ולכן:

$$\begin{aligned}Var(X)=Var\left(\sum_{i=1}^{n}X_i\right)&=\sum_{i=1}^{n}Var(X_i)+2\sum_{l<j}cov(X_l,X_j)\\&=n\cdot\frac{M}{N}\left(1-\frac{M}{N}\right)+2\binom{n}{2}\left(\frac{M(M-1)}{N(N-1)}-\frac{M^2}{N^2}\right)\end{aligned}$$

אי תלות של משתנים מקריים (שניים או יותר)

הגדרה: (אי-תלות של שני מ"מ בדידים)

נאמר שמ"מ X ו-Y בדידים (על אותו מרחב הסתברות) הינם בלתי תלויים, אם לכל שתי קבוצות $A, B \subset \mathbb{R}$ מתקיים שהמאורעות $\{X \in A\}$ ו-$\{Y \in B\}$ בלתי תלויים, כלומר:

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

הגדרה: (אי-תלות של מ"מ בדידים)

יהי X אוסף של מ"מ בדידים (על אותו מרחב הסתברות). נאמר שהמ"מ באוסף בלתי-תלויים אם לכל $\{X_i\}_{i\in[N]}$ (תת-קבוצה סופית של משתנים מקריים ב-X) ולכל $\{A_i\}_{i\in[N]}$ (אוסף קבוצות ב-$\mathbb{R}$) מתקיים:

$$\mathbb{P}(\forall i \in [N]\ X_i \in A_i) = \prod_{i\in[N]} \mathbb{P}(X_i \in A_i)$$

הגדרה: (אי-תלות של שתי קבוצות מ"מ בדידים)

יהיו X, Y שני אוספים של מ"מ בדידים (על אותו מרחב הסתברות)

נאמר שהאוספים בלתי תלויים ונסמן $X \perp Y$ אם לכל $\{X_i\}_{i\in[N]}$ (תת-קבוצה סופית של משתנים מקריים ב-X) ולכל $\{Y_i\}_{i\in[M]}$ (תת-קבוצה סופית של משתנים מקריים ב-Y)

ולכל $\{A_i\}_{i\in[N]}, \{B_i\}_{i\in[N]}$ אוספי קבוצות ב-$\mathbb{R}$ מתקיים:

$$\mathbb{P}(\forall i \in [N]\ \ X_i \in A_i,\ \ \forall i \in [M]\ Y_i \in B_i) = \mathbb{P}(\forall i \in [N]\ \ X_i \in A_i) \cdot \mathbb{P}(\forall i \in [M]\ Y_i \in B_i)$$

שימור אי תלות תחת הפעלת פונקציות

טענה: (שימור אי-תלות תחת הפעלת פונקציות)

יהיו $X_1, \dots, X_n$ ו-$Y_1, \dots, Y_m$ שתי קבוצות בלתי תלויות של מ"מ במרחב הסתברות.

תהיינה $g: \mathbb{R}^m \to \mathbb{R},\ \ f: \mathbb{R}^n \to \mathbb{R}$ פונקציות ממשיות.

אזי המ"מ $f(X_1, \dots, X_n)$ ו-$g(Y_1, \dots, Y_m)$ ב"ת.

<u>אבחנה</u>: (אי-תלות של משתנים מקריים בדידים היא תכונה של פונקציות התפלגות נקודתיות)

יהיו $X_1, \dots, X_n$ מ"מ במרחב הסתברות $(\Omega, \mathcal{F}, \mathbb{P})$. המ"מ ב"ת אם"ם לכל $x_1, \dots, x_n \in \mathbb{R}$ מתקיים

$$\mathbb{P}_{X_1,\dots,X_n}(x_1, \dots, x_n) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)$$

הפונקציה $\mathbb{P}_{X_1,\dots,X_n}(x_1, \dots, x_n)$ נקראת התפלגות הסתברותית של וקטור $(X_1, \dots, X_n)$.

חוסר זיכרון של התפלגות גאומטרית

טענה: (אפיון התפלגות גיאומטרית - חוסר זיכרון)

יהי X משתנה מקרי הנתמך על $\mathbb{N}$. שלושת הבאים שקולים:

(א) $\mathcal{D}$ הנה התפלגות גיאומטרית

(ב) X ו-$(X-1|X>1)$ שווי התפלגות

(ג) X ו-$(X-s|X>s)$ שווי התפלגות לכל $s\in\mathbb{N}$

הוכחה:

(א) גורר את (ב):

לכל $n\in\mathbb{N}$ מתקיים:

$$\mathbb{P}(X-1=n|X>1)=\frac{\mathbb{P}(X-1=n,X>1)}{\mathbb{P}(X>1)}=\frac{\mathbb{P}(X=n+1)}{\mathbb{P}(X>1)}=\frac{(1-p)^n\cdot p}{1-\mathbb{P}(X\le 1)}$$
$$=\frac{(1-p)^n\cdot p}{1-((1-p)^{1-1}\cdot p)}=\frac{(1-p)^n\cdot p}{1-p}=(1-p)^{n-1}\cdot p=\mathbb{P}(X=n)$$

(ב) גורר את (ג):

לפי (ב), $(X-1|X>1)$ ו-X שווי התפלגות. נגדיר $Y\coloneqq\big(X-(s-1)\big)$, נניח באינדוקציה כי המשתנים X ו-$(Y|Y>0)$ שווי התפלגות ונחשב:

$$(Y|Y>0)=X=(X-1|X>1)=(Y-1|Y>0,Y>1)=(Y-1|Y>1)$$
$$=(X-s|X>s)$$

כאשר השוויון השמאלי ביותר נובע מהנחת האינדוקציה והמגדרת Y, השני מתכונה (ב), השלישי שוב משימוש בהנחת האינדוקציה, הרביעי מתכונות בסיסיות של הסתברות מותנית והאחרון מהגדרת Y.

(ג) גורר את (א):

יהי X מ"מ בעל ערכים ב-$\mathbb{N}$ המקיים את תכונה (ג). נרשום לפי כלל השרשרת לכל $s\in\mathbb{N}_0$,

$$\mathbb{P}(X>s)=\mathbb{P}(X>s|X>s-1)\cdot\mathbb{P}(X>s-1|X>s-2)\cdot\ldots\cdot\mathbb{P}(X>1)\overset{(ג)}{=}\mathbb{P}(X>1)^s$$

נסמן $\mathbb{P}(X>1)=1-p$ וקיבלנו את פונקציית ההתפלגות השיורית של מ"מ גיאומטרי עם פרמטר p.

כעת נוכל להשתמש בכך ש-$\{X=s\}\cup\{X>s\}=\{X>s-1\}$ ולחשב

$$\mathbb{P}(X=s)=\mathbb{P}(X>s-1)-\mathbb{P}(X>s)=(1-p)^{s-1}-(1-p)^s\ =\ p(1-p)^s$$

בעקיפין למדנו כי פונקציית ההתפלגות השיורית של התפלגות על השלמים קובעת אותה.

■

התוחלת

הגדרה: (תוחלת)

יהי X משתנה מקרי המוגדר על מרחב הסתברות בדידה $(\Omega, \mathcal{F}, \mathbb{P})$

נזכר כי $Supp(X) = \{x \in \mathbb{R} \ : \ \mathbb{P}(X = x) > 0\}$ ונשים לב שזו היא קבוצה בדידה.

התוחלת של X מוגדרת על ידי

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) = \sum_{s \in Supp(X)} \sum_{\omega \in X^{-1}(s)} X(\omega)\mathbb{P}(\{\omega\}) = \sum_{s \in \mathbb{R}} s \cdot \mathbb{P}(X = s)$$

במידה וטור זה מתכנס בהחלט במובן הרחב. אחרת נאמר שלמשתנה X אין תוחלת.

טענה: (תוחלת של פונקציה של מ"מ – סטטיסטיקאי חסר הכרה)

יהי X מ"מ בדיד ותהי $f: \mathbb{R} \to \mathbb{R}$ פונקציה. אזי המ"מ $Y = f(X)$ מקיים

$$\mathbb{E}(Y) = \sum_{x \in \mathbb{R}} f(x) p_X(x)$$

בתנאי שטור זה מתכנס בהחלט.

תוחלות ההתפלגויות החשובות

תוחלת משתנה מקרי ברנולי $X\sim Ber(p)$ היא $p = \mathbb{E}(X) = p\cdot 1 + (1-p)\cdot 0$

לפיכך התוחלת של האינדיקטור של המאורע A היא $\mathbb{P}(A)$.

תוחלת משתנה מקרי אחיד על $[N]$ היא

$$\mathbb{E}(X) = \sum_{n\in[N]} n\cdot\mathbb{P}(X=n) = \frac{N(N+1)}{2N} = \frac{N+1}{2}$$

תוחלת משתנה מקרי בינומי $X\sim Bin(N,p)$ היא

$$\mathbb{E}(X) = \sum_{n=0}^{N} n\binom{N}{n}p^n(1-p)^{N-n} = \sum_{n=0}^{N} n\cdot\frac{N!}{(N-n)!\,n!}p^n(1-p)^{N-n}$$

$$= \sum_{n=1}^{N}\frac{N!}{(N-n)!\,(n-1)!}p^n(1-p)^{N-n} \underbrace{=}_{m=n-1} \sum_{m=0}^{N-1}\frac{N!}{(N-m-1)!\,m!}p^{m+1}(1-p)^{N-m-1}$$

$$= Np\sum_{m=0}^{N-1}\frac{(N-1)!}{(N-1-m)!\,m!}p^m(1-p)^{N-1-m} \underbrace{=}_{\text{בינום}} Np\big(p+(1-p)\big)^{N-1} = Np$$

תוחלת משתנה מקרי פואסוני $X\sim Po(\lambda)$ היא

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} ne^{-\lambda}\frac{\lambda^n}{n!} = \lambda e^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^{n-1}}{(n-1)!} \underbrace{=}_{m=n-1} \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$$

נשים לב שזו גם תוחלתו של משתנה $Y\sim Bin\left(N,\frac{\lambda}{N}\right)$ כפי שניתן היה לצפות.

תוחלת משתנה מקרי גיאומטרי $X\sim Geo(p)$ היא

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} np(1-p)^{n-1}$$

כדי לחשב טור זה נזכר בנוסחה לסכום טור הנדסי עבור $|x|<1$:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

נגזור את שני הצדדים ונקבל:

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}$$

כאשר הגזירה מוצדקת כיוון שטור הנגזרות מתכנס במידה שווה בסביבת x. נציב $x=1-p$:

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} np(1-p)^{n-1} = p\cdot\frac{1}{p^2} = \frac{1}{p}$$

תכונות התוחלת

טענה: (תכונות התוחלת)

יהיו X, Y מ״מ בדידים בעלי תוחלת סופית המוגדרים באותו מרחב הסתברות, אזי

(א) אם $\mathbb{P}(X \geq 0) = 1$ אז $\mathbb{E}(X) \geq 0$. אם בנוסף $\mathbb{P}(X > 0) > 0$ אז $\mathbb{E}(X) > 0$

(ב) ליניאריות התוחלת: $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ לכל $a, b \in \mathbb{R}$.

(ג) מונוטוניות התוחלת: אם $\mathbb{P}(X \geq Y) = 1$ אז $\mathbb{E}(X) \geq \mathbb{E}(Y)$

הוכחה:

(א) $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) \geq 0$

מכיוון שכל המחוברים אי-שליליים – שהרי אם $X(\omega) < 0$ אז לפי ההנחה שלנו $\mathbb{P}(\{\omega\}) = 0$ ואם לפחות אחד מהם חיובי אז התוחלת חיובית.

(ב) נחשב לפי הגדרה:

$$\mathbb{E}(aX + bY) = \sum_{\omega \in \Omega} \big(aX(\omega) + bY(\omega)\big) \cdot \mathbb{P}(\{\omega\})$$

$$= a \cdot \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}) + b \cdot \sum_{\omega \in \Omega} Y(\omega)\mathbb{P}(\{\omega\}) = a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y)$$

(ג) נרשום $X = (X - Y) + Y$ ונחשב לפי הסעיפים הקודמים:

$$\mathbb{E}(X) \underbrace{=}_{\text{סעיף ב}} \mathbb{E}(X - Y) + \mathbb{E}(Y) \underbrace{\geq}_{\text{סעיף א}} \mathbb{E}(Y)$$

■

טענה: (נוסחא לתוחלת של מ״מ טבעי באמצעות פונקציית ההתפלגות השיורית)

יהי X מ״מ המקיים $Supp(X) \subset \mathbb{N}$, אזי

$$\mathbb{E}(X) = \sum_{n \in \mathbb{N}} \mathbb{P}(X \geq n)$$

הוכחה: נחשב תוחלת לפי ההגדרה השקולה,

וננצל את העובדה שהמחוברים בטור אי-שליליים כדי להחליף סדר סכימה.

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} n\mathbb{P}(X = n) = \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mathbb{P}(X = n) = \sum_{\substack{k,n \in \mathbb{N} \\ k \leq n}} \mathbb{P}(X = n)$$

$$= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \mathbb{P}(X = n) = \sum_{n \in \mathbb{N}} \mathbb{P}(X \geq n)$$

■

תוחלת של מכפלת משתנים מקריים בלתי תלויים היא מכפלת התוחלות

טענה: (תוחלת של מכפלת מ"מ בלתי תלויים)

יהיו X,Y מ"מ ב"ת בדידים בעלי תוחלת סופית על מרחב הסתברות, אז התוחלת של XY קיימת ו-

$$\mathbb{E}\big(X(\omega)Y(\omega)\big)=\mathbb{E}\big(X(\omega)\big)\mathbb{E}\big(Y(\omega)\big)$$

<u>הוכחה</u>:

עבור $supp\ (X)=\{x_1,\dots,x_m\}$, $supp(Y)=\{y_1,\dots,y_k\}$,

נגדיר $A_i=\{\omega|X(\omega)=x_i\}$, $B_j=\{\omega|Y(\omega)=y_j\}$

נזכור כי:

$$X(\omega)=\sum_{i=1}^{m}x_i\cdot\mathbb{1}_{A_i}(\omega),\qquad Y(\omega)=\sum_{j=1}^{k}y_i\cdot\mathbb{1}_{B_j}(\omega)$$

נשים לב כי $\mathbb{1}_{A_i}(\omega)\mathbb{1}_{B_j}(\omega)=\mathbb{1}_{A_i\cap Bj}(\omega)$, ולכן ניתן לרשום:

$$X(\omega)\cdot Y(\omega)=\sum_{i=1}^{n}\sum_{j=1}^{k}x_iy_i\mathbb{1}_{A_i}(\omega)\mathbb{1}_{B_j}(\omega)=\sum_{i=1}^{n}\sum_{j=1}^{k}x_iy_i\mathbb{1}_{A_i\cap B_j}(\omega)$$

נרצה לחשב את התוחלת:

$$\mathbb{E}[X(\omega)\cdot Y(\omega)]\underset{\text{לינאריות התוחלת}}{=}\sum_{i=1}^{n}\sum_{j=1}^{k}x_iy_i\mathbb{E}\left[\mathbb{1}_{A_i}(\omega)\mathbb{1}_{B_j}(\omega)\right]=\sum_{i=1}^{n}\sum_{j=1}^{k}x_iy_i\mathbb{P}\left(\omega\in\left(A_i\cap B_j\right)\right)$$

אבל:

$$\mathbb{P}\left(\omega\in\left(A_i\cap B_j\right)\right)=\mathbb{P}\big(\omega:X(\omega)=x_i,Y(\omega)=y_j\big)\underset{\text{בלתי תלוים}}{=}\mathbb{P}(X(\omega)=x_i)\cdot\mathbb{P}\big(Y(\omega)=y_j\big)$$

לכן מתקיים:

$$\mathbb{E}[X(\omega)\cdot Y(\omega)]=\sum_{i=1}^{n}\sum_{j=1}^{k}x_iy_i\mathbb{P}(X(\omega)=x_i)\cdot\mathbb{P}\big(Y(\omega)=y_j\big)$$

$$=\sum_{i=1}^{m}x_i\mathbb{P}(X(\omega)=x_i)\cdot\sum_{j=1}^{k}y_j\cdot\mathbb{P}\big(Y(\omega)=y_j\big)=\mathbb{E}[X]\cdot\mathbb{E}[Y]$$

באותה צורה ניתן להוכיח עבור סדרה של מספר משתנים מקריים בלתי תלוים

■

תוחלת תחת התנייה

הגדרה: (תוחלת תחת התניה)

יהי X מ״מ במרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$, יהי A מאורע כך ש $\mathbb{P}(A)>0$, אזי:

$$\mathbb{E}[X|A]=\sum_{\omega\in\Omega}X(\omega)\cdot\mathbb{P}_A(\{\omega\})=\sum_{\omega\in\Omega}X(\omega)\cdot\mathbb{P}(\omega|A)=\sum_{s\in Im(X)}s\cdot\mathbb{P}(X=s|A)$$

טענה: (נוסחת התוחלת השלמה)

תהי $(A_n)_{n\in\mathbb{N}}$ חלוקה של מרחב הסתברות $(\Omega,\mathcal{F},\mathbb{P})$. מ״מ X בעל תוחלת סופית מקיים

$$\mathbb{E}(X)=\sum_{i=1}^{\infty}\mathbb{E}(X|A_i)\mathbb{P}(A_i)$$

כאשר אנו מפרשים איברים באגף ימין כשווים ל-0 אם $\mathbb{P}(A_i)=0$.

השונות

שקילות הגדרות השונות

הגדרה: (שונות וסטיית תקן)

יהי X משתנה מקרי, בעל תוחלת סופית. השונות של X מוגדרת כ-

$$Var(X) := \mathbb{E}\left(\left(X-\mathbb{E}(X)\right)^2\right) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

אם ל-X^2 אומנם יש תוחלת. אחרת נאמר של-X יש שונות אינסופית.

את השורש הריבועי של השונות $\sigma = \sqrt{Var(X)}$ נכנס בשם סטיית התקן של X.

<u>אבחנה</u>: נשים לב שהשוויון בין שתי ההגדרות של השונות נובע מלינאריות התוחלת:

$$\mathbb{E}\left(\left(X-\mathbb{E}(X)\right)^2\right) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2)$$
$$= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

טענה: (תכונות השונות)

יהי X מ״מ בעל שונות סופית, ויהי $a \in \mathbb{R}$. אזי

(א) $Var(X) \geq 0$ ושוויון מתקיים רק אם X קבוע.

(ב) $Var(X+a) = Var(X)$

(ג) $Var(aX) = a^2 Var(X)$ (ולכן $\sigma(aX) = |a|\sigma(X)$)

(ד) אם Y משתנה מקרי בעל שונות סופית **ב״ת** ב-X אז $Var(X+Y) = Var(X) + Var(Y)$

שונות ההתפלגויות החשובות

(א) שונות מ״מ בעל **התפלגות אחידה** על $[N]$ היא

$$Var(X) = \sum_{n\in[N]} n^2 \mathbb{P}(X=n) - \left(\frac{N+1}{2}\right)^2 = \frac{N(N+1)(2N+1)}{6N} - \frac{(N+1)^2}{4} = \frac{N^2-1}{12}$$

(ב) שונות מ״מ **ברנולי** $X\sim Ber(p)$ היא

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1-p)$$

(ג) שונות מ״מ **בינומי** $X\sim Bin(N,p)$ תחושב באופן הבא. ניזכר כי X שווה בהתפלגות לסכום $\{Y_n\}_{n\in\mathbb{N}}$ כאשר $Y_n\sim Ber(p)$ ב״ת. לפי סעיף ד׳ של תכונות השונות ולפי סעיף קודם:

$$Var(X) = \sum_{n=1}^{N} Var(Y_n) = Np(1-p)$$

(ד) שונות מ״מ **פואסוני** $X\sim Po(\lambda)$ נחשב באמצעות שימוש עקיף בטור טיילור. נרשום

$$e^{\lambda} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = \sum_{n=2}^{\infty} (n)(n-1)\frac{\lambda^{n-2}}{n!}$$

נכפיל ב-$\frac{\lambda^2}{e^{\lambda}}$ ונקבל:

$$\lambda^2 = \sum_{n=2}^{\infty} (n)(n-1)\frac{\lambda^n}{n!\,e^{\lambda}} = \mathbb{E}\big(X(X-1)\big) = \mathbb{E}(X^2) - \mathbb{E}(X)$$

מכאן ש-$\mathbb{E}(X^2) = \lambda^2 + \lambda$ ולכן $Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda$

(ה) שונות מ״מ **גיאומטרי** $X\sim Geo(p)$ נחשב בעזרת נוסחה לסכום של טור הנדסי עם $|x| < 1$:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

נגזור את שני הצדדים פעמיים ונקבל:

$$\sum_{n=2}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}$$

כאשר הגזירה מוצדקת כיוון שטור הנגזרות (הראשונות והשניות) מתכנס במידה שווה בסביבת x.

נציב כעת $x = 1-p$ ונקבל:

$$\mathbb{E}(X^2) - \mathbb{E}(X) = \mathbb{E}\big(X(X-1)\big) = \sum_{n=2}^{\infty} n(n-1)p(1-p)^{n-1} = p(1-p)\frac{2}{p^3} = \frac{2(1-\mathrm{p})}{\mathrm{p}^2}$$

לכן,

$$Var(X) = \mathbb{E}\big(X(X-1)\big) + \mathbb{E}(X) - \mathbb{E}(X)^2 = \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

משפט פואסון - התכנסות של מ"מ בינומיים להתפלגות פואסון:

יהי $\Omega = \{\omega = (a_1, \dots, a_n) | a_i = 0 \; or \; a_i = 1\}$ מרחב מדגם לניסוי ברנולי. עם $p(\omega) = p^{\sum_{i=1}^{n} a_i} q^{\left(n - \sum_{i=1}^{n} a_i\right)}$ כאשר $p + q = 1$.

$$s_n(\omega) = X_1(\omega) + \dots + X_n(\omega)$$

נניח כי $p = p(n)$, $q = q(n)$ כך שמתקיים:

- $p(n) + q(n) = 1$
- $\lim\limits_{n \to \infty} p(n) = 0$
- $\lim\limits_{n \to \infty} p(n) \cdot n = \lambda > 0$

אזי עבור $k \in \{0,1, \dots\}$ מתקיים:

$$\lim_{n \to \infty} \mathbb{P}\{\omega | S_n(\omega) = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$$

<u>הערה:</u> מכיון ש:

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

נקבל כי סדרת ההסתברויות $\left\{\frac{\lambda^k}{k!} \cdot e^{-\lambda}\right\}_{k=0}^{\infty}$ היא התפלגות ב $\{0,1, \dots\}$

הוכחה:

אנו יודעים כי

$$\mathbb{P}\{\omega | s_n(\omega) = k\} = \binom{n}{k} \cdot \left(p(n)\right)^k \cdot \left(q(n)\right)^{n-k}$$

נרצה לחשב את הגבול. נגדיר פונקציה $n \cdot p(n) - \lambda = \mu(n) \cdot n$ אזי מתקיים:

$$p(n) = \frac{\lambda}{n} + \mu(n) \quad \underset{\lim\limits_{n \to \infty} p(n) \cdot n = \lambda}{\Longrightarrow} \quad \lim_{n \to \infty} n \cdot \mu(n) = 0$$

אזי:

$$\mathbb{P}(s_n = k) = \frac{n!}{(n-k)!\,k!} \cdot \left(\frac{\lambda}{n} + \mu(n)\right)^k \cdot \left(1 - \frac{\lambda}{n} - \mu(n)\right)^{n-k}$$

נתבונן בשני האיברים הראשונים:

$$\frac{n!}{(n-k)!\,k!} \cdot \left(\frac{\lambda}{n} + \mu(n)\right)^k =$$

$$\frac{n(n-1) \dots (n-k+1)}{k!} \cdot \left(\frac{\lambda}{n} + \mu(n)\right)^k \xrightarrow[n \to \infty]{} \frac{n^k}{k!} \cdot \left(\frac{\lambda}{n} + 0\right)^k = \frac{\lambda}{k!}$$

כלומר גבול שני האיברים הראשונים קיים וערכו $\frac{\lambda^k}{k!}$.

כעת נתבונן באיבר הנותר:

$$\left(1-\frac{\lambda}{n}-\mu(n)\right)^{n-k}=e^{(n-k)\cdot\ln\left(1-\frac{\lambda}{n}-\mu(n)\right)}$$

$$=e^{(n-k)\cdot\ln\left(1-\frac{\lambda+n\mu(n)}{n}\right)}\underset{\ln(1+x)\simeq x}{\simeq}e^{-(n-k)\cdot\frac{(\lambda+n\mu(n))}{n}}$$

לכן נקבל:

$$\lim_{n\to\infty}\left(\left(1-\frac{\lambda}{n}-\mu(n)\right)^{n-k}\right)=\lim_{n\to\infty}\left(e^{-(n-k)\cdot\frac{(\lambda+n\mu(n))}{n}}\right)\simeq e^{-\lambda}$$

נחבר הכל ביחד ונקבל:

$$\mathbb{P}(s_n=k)=\frac{n!}{(n-k)!\,k!}\cdot\left(\frac{\lambda}{n}+\mu(n)\right)^{k}\cdot\left(1-\frac{\lambda}{n}-\mu(n)\right)^{n-k}\xrightarrow[n\to\infty]{}\frac{\lambda^k e^{-\lambda}}{k!}$$

אי שיוויונים

משפט: (אי-שוויון צ'בישב)

יהי $(\Omega,\mathcal{F},\mathbb{P})$ מרחב הסתברות. יהי X מ"מ אי-שלילי. אזי לכל $a>0$ מתקיים

$$\mathbb{P}(X\geq a)\leq\frac{\mathbb{E}(X)}{a}$$

<u>הוכחה</u>:

נחלק את Ω לשתי קבוצות:

$$\Omega=\underbrace{\{\omega|X(\omega)\geq a\}}_{A}\cup\underbrace{\{\omega|X(\omega)<a\}}_{A^c}$$

נציג את המשתנה המקרי באופן הבא:

$$\mathrm{X}(\omega)=X(\omega)\cdot\mathbb{I}_A(\omega)+\underbrace{X(\omega)\cdot\mathbb{I}_{A^c}(\omega)}_{\geq 0}>X(\omega)\cdot\mathbb{I}_a(\omega)\geq a\cdot\mathbb{I}_a(\omega)$$

כעת נתבונן בתוחלת:

$$\mathbb{E}(X(\omega))>a\cdot\mathbb{E}\big(\mathbb{I}_a(\omega)\big)=a\cdot\mathbb{P}(X\geq a)$$

וקיבלנו:

$$\mathbb{E}(X)\geq a\cdot\mathbb{P}(X\geq a)\quad\Leftrightarrow\quad\mathbb{P}(X\geq a)\leq\frac{\mathbb{E}(X)}{a}$$

■

מסקנות ממשפט צ'בישב:

$(\Omega,\mathcal{F}\,,\mathbb{P})$ מרחב הסתברות, $X:\Omega\to\mathbb{R}$ משתנה מקרי כלשהו, $\varepsilon>0$.

1. $\mathbb{P}\{\omega:|X(\omega)|\geq a\}\leq\frac{\mathbb{E}\{|X(\omega)|\}}{a}$ הוכחה מידית.
2. $\mathbb{P}\{\omega:|X(\omega)|\geq a\}\leq\frac{\mathbb{E}\{X^2(\omega)\}}{a^2}$
3. $\mathbb{P}\{\omega:|X(\omega)-\mathbb{E}\{X(\omega)\}|\geq a\}\leq\frac{var\{X(\omega)\}}{a^2}$

הוכחה:

1. ברור.
2. $\mathbb{P}\{\omega:|X(\omega)|\geq a\}=\mathbb{P}\{\omega:X^2(\omega)\geq a^2\}\underset{\text{מאי השיוויון צבישב}}{\leq}\frac{\mathbb{E}\{X^2(\omega)\}}{a^2}$
3. $\mathbb{P}\{\omega:|X(\omega)-\mathbb{E}\{X(\omega)\}|\geq a\}\underset{\text{ב}}{\leq}\frac{\mathbb{E}\left\{\left(X(\omega)-\mathbb{E}(X(\omega))\right)^2\right\}}{a^2}=\frac{var\{X(\omega)\}}{a^2}$

החוק החלש של המספרים הגדולים

משפט: (החוק החלש של המספרים הגדולים למשתנים בעלי שונות)

$(\Omega, \mathcal{F}, \mathbb{P})$ מרחב הסתברות כלשהו, $X_1, \dots, X_n: \Omega \to \mathbb{R}$ - משתנים מקריים

$X_1, \dots, X_n$ – בלתי תלויים שווי התפלגות.

$$\mu = \mathbb{E}\{X_1(\omega)\} = \cdots = \mathbb{E}\{X_n(\omega)\}\,; \quad \sigma^2 = Var\{X_1(\omega)\} = \cdots = Var\,\{X_n(\omega)\}$$

אזי:

$$\lim_{n\to\infty}\left\{\mathbb{P}\left\{\omega : \left|\frac{X_1(\omega) + \cdots + X_n(\omega)}{n} - \mu\right| \geq \varepsilon\right\}\right\} = 0$$

הוכחה:

$$s_n(\omega) = X_1(\omega) + \cdots + X_n(\omega)$$

ראשית נשים לב כי: $\mathbb{E}\{S_n(\omega)\} = \mathbb{E}\{X_1(\omega)\} + \cdots + \mathbb{E}\{X_n(\omega)\} = n \cdot \mu$

מהמסקנה של אי שיוויון ציבישב, נקבל כי:

$$\mathbb{P}\left\{\omega: \left|\frac{S_n(\omega)}{n} - \mu\right| \geq \varepsilon\right\} \leq \frac{var\left\{\frac{s_n(\omega)}{n}\right\}}{\varepsilon^2} = \frac{Var\{s_n(\omega)\}}{n^2\varepsilon^2}$$

$$Var\,\{s_n(\omega)\} = \mathbb{E}\{S_n^2(\omega)\} - (\mathbb{E}\{s_n(\omega)\})^2$$

נשים לב כי: $(\mathbb{E}\{s_n(\omega)\})^2 = (n\mu)^2$

$$\begin{aligned}\mathbb{E}\{S_n^2(\omega)\} &= \mathbb{E}\left(\left(\sum_{i=1}^n X_i(\omega)\right)\left(\sum_{j=1}^n X_j(\omega)\right)\right) = \mathbb{E}\left(\sum_{i=1}^n X_i^2(\omega)\right) + 2\sum_{1\leq i<j\leq n}\mathbb{E}\{X_i(\omega)X_j(\omega)\}\\ &= \mathbb{E}\left\{\sum_{i=1}^n X_i^2(\omega)\right\} + 2\sum_{1\leq i<j\leq n}\mu^2 = \sum_{i=1}^n \mathbb{E}\{X_i^2(\omega)\} + 2\binom{n}{2}\mu^2\\ &= \sum_{i=1}^n \mathbb{E}\{X_i^2(\omega)\} + n(n-1)\cdot\mu^2\end{aligned}$$

נציב במה שמצאנו קודם:

$$\begin{aligned}Var\,\{s_n(\omega)\} &= \sum_{i=1}^n \mathbb{E}\{X_i^2(\omega)\} + \mu^2 n(n-1) - n^2\mu^2 = \sum_{i=1}^n \mathbb{E}\{X_i^2(\omega)\} - n\mu^2\\ &= \sum_{i=1}^n\left(\mathbb{E}\{X_i^2(\omega)\} - \mathbb{E}\left(X_i(\omega)\right)^2\right) = \sum_{i=1}^n Var\,\{X_i(\omega)\} = n\sigma^2\end{aligned}$$

וכעת נקבל: $\mathbb{P}\left(\omega: \left|\frac{S_n(\omega)}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n^2\varepsilon^2} \xrightarrow[n\to\infty]{} 0$

כנדרש.

שונות משותפת

הגדרה: (שונות משותפת)

יהיו X, Y משתנים מקריים על אותו מרחב הסתברות, בעלי תוחלת סופית.

השונות המשותפת של X ו-Y, מוגדרת על ידי

$$Cov(X,Y) := \mathbb{E}\Big(\big(X-\mathbb{E}(X)\big)\big(Y-\mathbb{E}(Y)\big)\Big) = \mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$$

כאשר תוחלת זו מוגדרת היטב.

שני משתנים אשר שונותם המשותפת מתאפסת – נקראים בלתי מתואמים.

<u>מסקנה</u>: (אי-תלות גוררת חוסר מותאמות) אם X ו-Y בלתי תלויים אז הם בלתי מתואמים.

טענה: (תכונות השונות המשותפת)

יהיו X, Y, Z מ"מ בעלי שונות סופית, ויהיו $a, b \in \mathbb{R}$.

אזי בכל מקרה בו אגף שמאל מוגדר היטב מתקיים

(א) סימטריות: $Cov(X,Y) = Cov(Y,X)$

(ב) בי-לינאריות: $Cov(aX+bZ,Y) = aCov(X,Y)+bCov(Z,Y)$

(ג) $Cov(X,X) = Var(X)$

שונות של סכום משתנים מקריים

טענה: (שונות של סכום מ"מ)

יהיו X, Y משתנים מקריים. אזי

$$Var(X+Y) = Var(X)+Var(Y)+2Cov(X,Y)$$

בכל מקרה בו אגף ימין של המשוואה מוגדר היטב.

<u>הוכחה</u>: נשתמש בהגדרת השונות ובלינאריות התוחלת

$$Var(X+Y) = \mathbb{E}((X+Y)^2)-\mathbb{E}(X+Y)^2$$
$$= \mathbb{E}(X^2)+2\mathbb{E}(XY)+\mathbb{E}(Y^2)-\mathbb{E}(X)^2-2\mathbb{E}(X)\mathbb{E}(Y)-\mathbb{E}(Y)^2$$
$$= Var(X)+2Cov(X,Y)+Var(Y)$$

■

ניתן להכליל נוסחה זו למספר כלשהו של משתנים מקריים.

טענה: (נוסחת שונות לסכום)

לכל אוסף $(X_n)_{n\in[N]}$ של מ"מ מתקיים

$$Var\left(\sum_{n=1}^{N} X_n\right) = \sum_{n,k\le N} Cov(X_n,X_k) = \sum_{n\le N} Var(X_n) + 2\sum_{n<k\le N} Cov(X_n,X_k)$$

בכל מקרה בו אגף ימין של המשוואה מוגדר היטב.

משפט – $The\ Weierstrass\ Theorem$

תהי $f:[0,1]\to\mathbb{R}$ פונקציה רציפה, יהי $P\in[0,1]$ ונגדיר פולינום (הפולינום של ברנשטיין):

$$B_n(p)\overset{def}{=}\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}$$

נטען כי מתקיים:

$$\lim_{n\to\infty}\left(\max_{p\in[0,1]}\{|f(b)-B_n(p)|\}\right)=0$$

אזי סדרת הפולינומים $\{B_n(p)\}$ מתכנסת במידה שווה לפונקציה f.

הוכחה:

יהי $\varepsilon>0$, אם f רציפה בקטע $[0,1]$ אז היא גם רציפה במידה שווה בקטע $[0,1]$ כלומר קיים δ כך שעבור $x,y\in[0,1]$:

$$|x-y|\le\delta\Longrightarrow|f(x)-f(y)|\le\varepsilon$$

נרצה לחשב את:

$$\begin{aligned}|f(p)-B_n(p)|&=\left|1\cdot f(p)-\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}\right|\\
&\underset{\text{שימוש ב }1=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}}{=}\left|\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}\cdot f(p)-\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}\right|\\
&=\left|\sum_{k=0}^{n}\left(f(p)-f\left(\frac{k}{n}\right)\right)\binom{n}{k}p^k(1-p)^{n-k}\right|\\
&\underset{\text{אי שוויון המשולש}}{\le}\sum_{k=0}^{n}\left|f(p)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}\\
&=\underbrace{\sum_{k:\left|\frac{k}{n}-p\right|\le\delta}\left|f(p)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}}_{=I}\\
&+\underbrace{\sum_{k:\left|\frac{k}{n}-p\right|>\delta}\left|f(p)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}}_{=II}=I+II\end{aligned}$$

ראשית נחסום את הסכום הראשון – I (נפעיל על חלק זה רציפות במידה שווה):

$$I\le\sum_{k:\left|\frac{k}{n}-p\right|\le\delta}\left|f(p)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}p^k(1-p)^{n-k}\underset{\left|f(p)-f\left(\frac{k}{n}\right)\right|<\varepsilon}{\le}\varepsilon\cdot\underbrace{\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}}_{=1}=\varepsilon$$

כעת נחסום את הסכום השני, מכיון ש:

$$|f(x)|\le M,\forall x\in[0,1]\Longrightarrow\left|f(p)-f\left(\frac{n}{k}\right)\right|\le|f(p)|+\left|f\left(\frac{k}{n}\right)\right|\le 2M$$

נקבל:

$$II = \sum_{k:\left|\frac{k}{n}-p\right|>\delta} \left|f(p) - f\left(\frac{k}{n}\right)\right| \binom{n}{k} p^k (1-p)^{n-k} < 2M \sum_{k:\left|\frac{k}{n}-p\right|>\delta} \binom{n}{k} p^k (1-p)^{n-k}$$

כעת, יהיו $X_1, \dots, X_n$ מ״מ בינומים ונגדיר $s_n(\omega) = X_1(\omega) + \cdots + X_n(\omega)$.

לפי החוק החלש של המספרים הגדולים:

$$\mathbb{P}\left(\left\{\omega: \left|\frac{s_n(\omega)}{n} - p\right|\right\} > \delta\right) \leq \frac{n \cdot Var(X_1)}{\delta^2 n^2} = \frac{p(1-p)}{\delta^2 n}$$

או במילים אחרות:

$$\sum_{k:\left|\frac{k}{n}-p\right|>\delta} \binom{n}{k} p^k (1-p)^{n-k} \leq \frac{p(1-p)}{\delta^2 n}$$

$$|f(p) - B_n(p)| \leq \varepsilon + \frac{2M \cdot p(1-p)}{\delta^2 \cdot n} \xrightarrow[n\to\infty]{} \varepsilon \qquad \forall p \in [0,1]$$

$$\lim_{n\to\infty}\left(\max_{p\in[0,1]}\{|f(b) - B_n(p)|\}\right) = 0$$

■

משפט - $\boldsymbol{The\ Gaussian\ approximation\ for\ the\ binomial\ distribution}$

נגדיר:
$$p_n(k)=\binom{n}{k}p^k(1-p)^{n-k} \quad ; \quad 0<p<1$$

$$k=k(n) \quad ; \quad \lim_{n\to\infty} k(n)=\infty \quad ; \quad \lim_{n\to\infty}\big(n-k(n)\big)=\infty$$

$$\lim_{n\to\infty}\frac{(k(n)-np)^3}{n^2}=0$$

אזי:

$$\lim_{n\to\infty}\left[\frac{p_n\big(k(n)\big)}{\frac{1}{\sqrt{2\pi pq}}e^{-\frac{(k(n)-np)^2}{2npq}}}\right]=1$$

כלומר:

$$P_n\big(k(n)\big)\simeq\frac{1}{\sqrt{2\pi pq}}e^{-\frac{(k(n)-np)^2}{2npq}}$$

הוכחה:

נזכר בנוסחת סטרלינג,

$$n!=\sqrt{2\pi n}e^{-n}n^n\big(1+R(n)\big),\quad \lim_{n\to\infty}R(n)=0$$

כעת נציב את נוסחת סטרלינג ב$p_n(k)$:

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

$$=\frac{\sqrt{2\pi n}e^{-n}n^n\cdot p^k(1-p)^{n-k}}{\sqrt{2\pi(n-k)}e^{-(n-k)}(n-k)^{n-k}\cdot\sqrt{2\pi k}e^{-k}k^k}\cdot\frac{\big(1+R(n)\big)}{\big(1+R(n-k)\big)\cdot\big(1+R(k)\big)}$$

כאשר $n\xrightarrow[n\to\infty]{}\infty$,$R\xrightarrow[n\to\infty]{}\infty$,$(n-k)\xrightarrow[n\to\infty]{}\infty$ נקבל:

$$p_n(k)\simeq\left[\frac{n}{2\pi k(n-k)}\right]^{\frac12}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}$$

נגדיר $\delta_k=k-np$, מכך נסיק כי $n-k=nq-\delta_k$.

$$\frac{\delta_k^3}{n^2}\xrightarrow[n\to\infty]{}0$$

אזי ניתן לרשום:

$$p_n(k)=\left[\frac{n}{2\pi(\delta_k+np)(nq-\delta_k)}\right]^{\frac12}\left(\frac{np}{\delta_k+np}\right)^{\delta_k+np}\left(\frac{nq}{nq-\delta_k}\right)^{nq-\delta_k}$$

$$=\left[\frac{n}{2\pi(\delta_k+np)(nq-\delta_k)}\right]^{\frac12}\left(\frac{1}{1+\left(\frac{\delta_k}{np}\right)}\right)^{\delta_k+np}\left(\frac{1}{1-\left(\frac{\delta_k}{nq}\right)}\right)^{nq-\delta_k}$$

ברור כי:

$$\left[\frac{n}{2\pi(\delta_k+np)(nq-\delta_k)}\right]^{\frac12}\simeq\frac{1}{\sqrt{2\pi npq}}$$

כמו כן מתקיים:

$$\left(\frac{1}{1+\left(\frac{\delta_k}{np}\right)}\right)^{\delta_k+np}\left(\frac{1}{1-\left(\frac{\delta_k}{nq}\right)}\right)^{nq-\delta_k} = \mathrm{e}^{-(\delta_k+np)\ln\left(1+\left(\frac{\delta_k}{np}\right)\right)-(nq-\delta_k)\ln 1-\left(\frac{\delta_k}{nq}\right)}$$

$$= \mathrm{e}^{-(\delta_k+np)\left(\frac{\delta_k}{np}-\frac{\delta_k^2}{2n^2p^2}+\cdots\right)-(nq-\delta_k)\left(\frac{\delta_k}{nq}+\frac{\delta_k^2}{2n^2q^2}+\cdots\right)}$$

$$= \exp\left[-\delta_k+\delta_k-\delta_k^2\left(\frac{1}{np}-\frac{1}{2np}\right)-\delta_k^2\left(\frac{1}{nq}-\frac{1}{2nq}\right)+\cdots\right]$$

$$\simeq \exp\left[-\frac{\delta_k^2}{2n}\left(\frac{1}{p}+\frac{1}{q}\right)\right] = \exp\left[-\frac{\delta_k^2}{2npq}\right]$$

אזי:

$$p_n(k) \simeq \frac{1}{\sqrt{2\pi npq}}e^{-\frac{\delta_k^2}{2npq}} = \frac{1}{\sqrt{2\pi npq}}e^{-\frac{(np-k)^2}{2npq}}$$

■

מומנטים

תכונות הפונקציה יוצרת מומנטים

הגדרה: (מומנטים פולינומיאלים)

יהי X משתנה מקרי. המומנט מסדר k של X מוגדר בתור

$$m_k(X) = \mathbb{E}(X^k)$$

כאשר תוחלת זו מוגדרת היטב.

הגדרה: (פונקציה יוצרת מומנטים)

יהי X משתנה מקרי. הפונקציה הממשית $M_X(t)$ הנתונה על ידי

$$M_X(t) \coloneqq \mathbb{E}(e^{tX})$$

לכל t עבורו תוחלת זו מוגדרת היטב, מכונה הפונקציה יוצרת מומנטים של X.

מ"מ בעל פונקציה יוצרת מומנטים בסביבה כלשהי של הראשית נקרא בעל מומנט מעריכי.

טענה: (כפליות פונקציה יוצרת מומנטים)

יהיו X ו-Y משתנים מקריים בלתי תלויים ויהי $Z = X + Y$. אזי

$$M_Z(t) = M_X(t) \cdot M_Y(t)$$

אי שיוויון צ'רנוף

משפט: (אי-שוויון צ'רנוף)

יהי X משתנה מקרי בעל מומנט מעריכי. אזי לכל $t > 0$ עבורו $M_X(t)$ מוגדרת ולכל $a \in \mathbb{R}$ מתקיים

$$\mathbb{P}(X \geq a) \leq M_X(t)e^{-ta}$$

<u>הוכחה</u>: נציב באי שיוויון מרקוב את המ"מ e^{tX} ונקבל

$$\mathbb{P}(X \geq a) = \mathbb{P}(e^{tX} \geq e^{ta}) \leq \frac{\mathbb{E}(e^{tX})}{e^{ta}} = M_X(t)e^{-ta}$$

■

אי שיוויון הופדינג

משפט: (אי-שוויון הופדינג)

יהיו $\{X_i\}_{i\in[N]}$ משתנים מקריים בלתי-תלויים ובעלי תוחלת אפס, המקיימים $|X_i| \leq 1$ לכל $i \in [N]$.

אזי לכל $a > 0$ מתקיים:

$$\mathbb{P}\left(\sum_{i\in[N]} X_i \geq a\right) \leq exp\left(-\frac{a^2}{2N}\right)$$

פונקציה יוצרת מומנטים של ההתפלגויות החשובות

(א) פונקציה יוצרת מומנטים של מ״מ ברנולי $X\sim Ber(p)$ היא

$$\mathbb{E}(e^{tX}) = pe^t + 1 - p$$

(ב) פונקציה יוצרת מומנטים של מ״מ מתפלג $X\sim Uni[n]$, היא:

$$\mathbb{E}[e^{tX}] = \sum_{k=1}^{n} e^{tk} \cdot \mathbb{P}(X=k) = \sum_{k=1}^{n} e^{tk} \cdot \frac{1}{n} = \frac{1}{n}\sum_{k=1}^{n} e^{tk} \underset{\text{סדרה הנדסית}}{=} \frac{1}{n} \cdot \frac{e^t(1-e^{tn})}{1-e^t}$$

(ג) פונקציה יוצרת מומנטים של מ״מ בינומי $X\sim Bin(N,p)$ כאשר קיימים $X_1,\dots,X_n$ מ״מ מתפלגים ברנולי בלתי תלוים עבורם $X=\sum_{i=1}^{n} X_i$, היא:

$$\mathbb{E}(e^{tX}) = \mathbb{E}\left(e^{t\cdot\sum_{k=1}^{n} X_i}\right) \underset{\text{אי תלות}}{=} \prod_{k=1}^{n} \mathbb{E}[e^{tX_i}] = \prod_{k=1}^{n} (pe^t + 1 - p) = (pe^t + 1 - p)^n$$

(ד) פונקציה יוצרת מומנטים של מ״מ פואסוני $X\sim Po(\lambda)$ היא

$$\mathbb{E}(e^{tX}) = \sum_{n=0}^{\infty} e^{tn} \cdot \mathbb{P}(X=n) = \sum_{n=0}^{\infty} e^{-\lambda}\frac{(\lambda e^t)^n}{n!} = e^{-\lambda} \cdot \sum_{n=0}^{\infty} \frac{(\lambda e^t)^n}{n!} \underset{\text{טור טיילור}}{=} e^{\lambda(e^t-1)}$$

(ה) פונקציה יוצרת מומנטים של מ״מ גיאומטרי $X\sim Geo(p)$ היא

$$\mathbb{E}(e^{tX}) = \sum_{n=1}^{\infty} e^{nt} p(1-p)^{n-1} = pe^t \sum_{n=1}^{\infty} e^{t(n-1)}(1-p)^{n-1}$$

$$= \begin{cases} \text{לא מוגדר} & e^{t(n-1)}(1-p)^{n-1} \geq 1 \\ \dfrac{pe^t}{1-(1-p)e^t} & else \end{cases}$$

והיא מוגדרת רק עבור $t < -\ln(1-p)$.

גאוסיאנים, גבול מרכזי ושיט של ייבגני:

הגדרה:

נגדיר את המ״מ X_{Gauss} באופן הבא:

$$\Phi_{Gauss}(x) \overset{def}{=} \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{X} e^{-\frac{t^2}{2}} dt$$

$$\mathbb{P}(\omega : X_{Gauss}(\omega) \leq a) = \Phi_{Gauss}(a)$$

טענה:

$$\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi}$$

תוחלת ושונות:

$$\mathbb{E}[X_{Gauss}] = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{-\infty} t \cdot e^{-\frac{t^2}{2}} dt = 0$$

$$Var[X_{Gauss}] = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} t^2 \cdot e^{-\frac{t^2}{2}} dt = 1$$

משפט הגבול המרכזי:

יהיו $X_1, \dots, X_n$ משתנים מקריים בלתי תלויים עם אותה התפלגות.

$$\mathbb{E}[X_1] = \cdots = \mathbb{E}[X_n] = \mu\,, \qquad Var(X_1) = \cdots = Var(X_n) = \sigma^2$$

אז מתקיים:

$$\lim_{n\to\infty} \left[\mathbb{P}\left\{\omega : a < \frac{X_1(\omega) + \cdots + X_n(\omega) - n\mu}{\sigma\sqrt{n}} < b\right\}\right] = \frac{1}{\sqrt{2\pi}} \cdot \int_{a}^{b} e^{-\frac{t^2}{2}}\, dt$$

או:

$$X_1(\omega) + \cdots + X_n(\omega) \simeq n\mu + \sigma\sqrt{n} \cdot X_{Gauss}$$

פרק 8 – משתנים מקריים רציפים בהחלט

הגדרה: (משתנה מקרי רציף בהחלט)

יהיה X מ"מ על מרחב הסתברות. נאמר ש-X רציף בהחלט אם קיימת פונקציה אינטגרבילית $f: \mathbb{R} \to \mathbb{R}_+$ כך שמתקיים

$$\mathbb{P}(X \leq a) = F(a) = \int_{-\infty}^{a} f(x)\, dx$$

הפונקציה f נקראת הצפיפות של X.

בפרט לכל $a, b \in \mathbb{R}$ מתקיים ש:

$$\mathbb{P}(X \in (a,b]) = \mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(x \leq a) = F_X(b) - F_X(a) = \int_{a}^{b} f(x)\, dx$$

תכונות של מ"מ רציף:

3. נרמול:

$$\int_{\mathbb{R}} f_X(t)dt = 1$$

4. רציפות: פונקציית ההתפלגות המצטברת F_X היא רציפה. בפרט $P(X = a) = 0$ לכל $a \in \mathbb{R}$.

התפלגויות חשובות: אחידה, מעריכית ונורמלית

הגדרה: (התפלגות אחידה)

יהי $[a,b]\subset\mathbb{R}$ קטע. נאמר שלמ״מ X התפלגות אחידה על $[a,b]$ ונכתוב $X\sim Unif([a,b])$ אם צפיפותו היא

$$f(x)=\frac{\mathbb{I}_{([a,b])}}{b-a}$$

אבחנה: (תכונות התפלגות אחידה)

יהי X מ״מ $X\sim Unif([a,b])$ אזי מתקיים

$$\mathbb{E}(X)=\frac{a+b}{2}$$

$$F_X(t)=\begin{cases}0, & t<a\\ \frac{t-a}{b-a}, & a\le t\le b\\ 1, & t>b\end{cases}$$

$$Var(X)=\frac{(b-a)^2}{12}$$

$$M_X(t)=\frac{\left(e^{tb}-e^{ta}\right)}{t(b-a)}$$

<u>הוכחה</u>:

$$F_X(t)=\int_{-\infty}^{t}\frac{\mathbb{I}_{([a,b])}}{b-a}dx=\begin{cases}0, & t<a\\ \frac{t-a}{b-a}, & a\le t\le b\\ 1, & t>b\end{cases}$$

$$\mathbb{E}(X)=\int_a^b\frac{xdx}{b-a}=\frac{x^2}{2(b-a)}\bigg|_a^b=\frac{a+b}{2}$$

$$Var(X)=\mathbb{E}(X^2)-\mathbb{E}(X)^2=\int_a^b\frac{x^2dx}{b-a}-\left(\frac{a+b}{2}\right)^2=\frac{b^3-a^3}{3(b-a)}-\frac{(a+b)^2}{4}$$

$$=\frac{a^2+ab+b^2}{3}-\frac{(a+b)^2}{4}=\frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12}=\frac{(b-a)^2}{12}$$

$$M_X(t)=\mathbb{E}(e^{Xt})=\int_a^b\frac{e^{tX}dx}{b-a}=\frac{e^{tb}-e^{ta}}{t(b-a)}$$

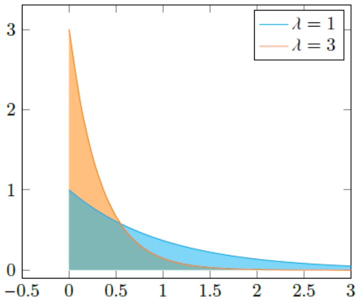
■

הגדרה: (התפלגות מעריכית)

נאמר שלמ"מ X התפלגות מעריכית עם פרמטר λ,

ונכתוב $X\sim \exp(\lambda)$ אם צפיפותו היא

$$f(x) = \lambda e^{-\lambda x}$$



צפיפות התפלגות מעריכית עבור שני פרמטרים שונים.

אבחנה: (תכונות התפלגות מעריכית)

יהי X מ"מ $X\sim Exp(\lambda)$ אזי מתקיים

$$\mathbb{E}(X) = \frac{1}{\lambda} \qquad F_X(t) = max(1 - e^{-\lambda t}, 0)$$

$$Var(X) = \frac{1}{\lambda^2} \qquad M_X(t) = \frac{\lambda}{\lambda - t} \quad \text{עבור } t < \lambda$$

נזכור כי אם $X\sim Exp(\lambda)$ אזי X חסר זכרון.

<u>הוכחה</u>:

$$F_X(t) = \int_0^s \lambda e^{-\lambda s} ds = max(1 - e^{-\lambda t}, 0)$$

$$\mathbb{E}(X) = \int_0^\infty \lambda s e^{-\lambda s} ds = \lim_{s\to\infty} \left[-xe^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x} \right]_{x=0}^s = \frac{1}{\lambda}$$

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \int_0^\infty \lambda s^2 e^{-\lambda s} ds - \frac{1}{\lambda^2}$$

לפי אינטגרציה בהצבה כאשר

$$\begin{vmatrix} f(x) = x^2 & f'(x) = 2x \\ g(x) = e^{-\lambda x} & g'(x) = -\lambda e^{-\lambda x} \end{vmatrix}$$

מתקיים:

$$\lim_{s\to\infty} \left(\left[-x^2 e^{-\lambda x} \right]_0^s + 2 \cdot \int_0^r x e^{-kx} dx \right) = \lim_{s\to\infty} \left[-x^2 e^{-\lambda x} - \frac{2}{\lambda} x e^{-\lambda x} - \frac{2}{\lambda^2} e^{-\lambda} \right]_{x=0}^s - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

כמו כן, עבור $\lambda < t$ מתקיים:

$$M_X(t) = \int_0^\infty \lambda e^{(t-\lambda)s} ds = \lim_{s\to\infty} \left[\frac{\lambda}{\lambda - t} e^{(t-\lambda)x} \right]_{x=0}^s = \frac{\lambda}{\lambda - t'}$$

וכן $F_{\alpha X}(t) = F_X\left(\frac{t}{\alpha}\right) = max\left(1 - e^{-\frac{\lambda t}{\alpha}}, 0\right)$

■

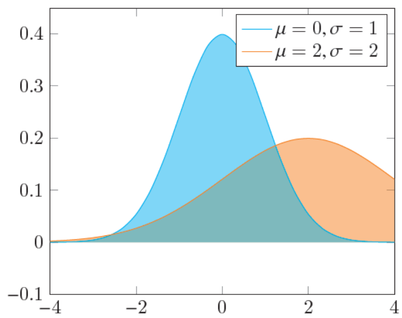
הגדרה: (התפלגות נורמלית)

נאמר שלמ"מ X התפלגות נורמלי עם תוחלת μ ושונות σ^2

ונכתוב $X \sim N(\mu, \sigma^2)$, אם צפיפותו היא

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

אם $\mu = 0$ ו-$\sigma = 1$ נאמר ש-X נורמלי סטנדרטי.



צפיפות התפלגות נורמלית עבור שני זוגות פרמטרים שונים.

הגדרה: (פונקציית התפלגות מצטברת של התפלגות נורמלית)

נסמן מעתה את פונקציית ההתפלגות המצטברת של משתנה מקרי $X \sim N(0,1)$ ב-

$$\Phi(x) = \mathbb{P}(X \leq x)$$

אבחנה: (תכונות התפלגות נורמאלית)

יהי X מ"מ $X \sim Exp(\lambda)$ אזי מתקיים

$$\mathbb{E}(X) = \mu \qquad F_X(t) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$Var(X) = \sigma^2 \qquad M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

וכן לכל $\alpha > 0, \beta \in \mathbb{R}$ מתקיים $\alpha X + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2)$

סכום של שני משתנים מקריים בלתי תלוים:

יהיו X_1, X_2 מ״מ רציפים בלתי תלויים.

נסמן $f_{X_1}(x)$ פונקצית הצפיפות של X_1, $f_{X_2}(x)$ פונקציית הצפיפות של X_2. אזי מתקיים:

$$f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(x-t)dt$$

<u>הוכחה:</u>

ראשית נשים לב שמכיון ש X_1, X_2 מ״מ ב״ת ניתן לרשום את פונקצית הצפיפות המשותפת שלהם באופן הבא:

$$f_{X_1,X_2}(a,b) = f_{X_1}(a) \cdot f_{X_2}(b)$$

כעת:

$$\mathbb{P}\{\omega: X_1(\omega) + X_2(\omega) \le z\} = \iint_{(a,b):a+b\le z} f_{X_1,X_2}(a,b)\, dx\, dy$$

נסתכל על השטח שאנו רוצים לחשב (הפונקציה $a + b \le z$ שקולה ל $b \le z - a$, אז למעשה כל השטח מתחת לישר $b = z - a$)

נסמן פונקציות: $[-\infty, z] \ni u \Leftarrow u = a + b$ ונשאיר את $a = a$.

כאשר הסימון $\frac{\partial y}{\partial x}$ – מייצג את הנגזרת של y לפי x . לבסוף חישבנו דטרמיננטה של המטריצה שהתקבלה.

$$\begin{bmatrix} \frac{\partial u}{\partial a} & \frac{\partial u}{\partial b} \\ \frac{\partial a}{\partial a} & \frac{\partial a}{\partial b} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = 1$$

לכן ניתן לרשום:

$$\iint_{(a,b):a+b\le z} f_{X_1,X_2}(a,b)\, da\, db \underset{\text{נציב } u \text{ במקום } y}{=} \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f_{X_1,X_2}(a, u-a)\, db\right) da$$

$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f_{X_1}(a) \cdot f_{X_2}(u-a)\ du\right) da$$

$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f_{X_1}(a) \cdot f_{X_2}(u-a)\ da\right) du$$

ולכן נקבל (אם נסתכל על פונקציית הצפיפות של המ״מ החדש):

$$f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_{X_1}(t) f_{X_2}(x-t)dt = \int_{-\infty}^{\infty} f_{X_1}(a) \cdot f_{X_2}(u-a)\ da$$

תוחלת, שונות ופונקציה יוצרת מומנטים של ההתפלגויות החשובות

הגדרה: (תוחלת משתנה מקרי רציף)

יהי X מ״מ רציף בהחלט בעל פונקציית צפיפות $f_X(x)$ אזי

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)\,dx$$

בכל מקרה בו אינטגרל זה מתכנס בהחלט. אחרת ל-X אין תוחלת.

טענה: (תוחלת של פונקציה של מ״מ – סטטיסטיקאי חסר הכרה)

יהי X מ״מ בעל צפיפות f_X ותהי $g: \mathbb{R} \to \mathbb{R}$ פונקציה מדידה.

אז $Y = g(X)$ הוא מ״מ המקיים

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

אם ורק אם האינטגרל מתכנס בהחלט.

נשים לב שלא דרשנו בטענה זו כי המשתנה Y יהיה רציף, ואומנם הטענה נכונה באופן כללי.

על סמך הטענה לעיל נוכל לחשב מומנטים של מ״מ רציפים בהחלט.

כך למשל, שונות של מ״מ רציף X תתקבל מן החשבון

$$Var(X) = \int_{-\infty}^{\infty} x^2 f_X(x)dx - \left(\int_{-\infty}^{\infty} xf_X(x)dx\right)^2$$

צפיפות משותפת

הגדרה: (צפיפות משותפת של שני משתנים מקריים)

נאמר כי לשני משתנים מקריים X, Y מעל מרחב הסתברות משותף $(\Omega, \mathcal{F}, \mathbb{P})$

יש צפיפות משותפת אם קיימת פונקציה אינטגרבילית $f_{X,Y}(x,y): \mathbb{R}^2 \to [0, \infty)$

המקיימת לכל קבוצה מהטיפוס $A = (-\infty, a] \times (-\infty, b]$

$$\mathbb{P}_{X,Y}(A) = \mathbb{P}(X \leq a, Y \leq b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x,y)\, dx\, dy$$

הפונקציה $f_{X,Y}(x,y)$ נקראת צפיפות משותפת של X ו-Y.

התפלגות שולית מהתפלגות מצפיפות משותפת

<u>אבחנה</u>: יהיו X, Y מ"מ בעלי צפיפות משותפת $f_{X,Y}$ אז X ו-Y רציפים בהחלט ומתקיים כי

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$$

הן צפיפויות של X ו-Y בהתאמה. התפלגויותיהם של X ו-Y נקראות ההתפלגויות השוליות של X ו-Y.

<u>אבחנה</u>: (ב"ת אם"ם קיימת צפיפות משותפת)

יהיו X, Y רציפים בהחלט בעלי צפיפות f_X, f_Y בהתאמה,

אזי X ו-Y ב"ת אם ורק אם קיימת להם צפיפות משותפת המקיימת

$$f_{X,Y} = f_X f_Y$$

צפיפות מותנית

הגדרה: (צפיפות מותנית)

יהיו X,Y מ"מ רציפים בהחלט בעלי התפלגות משותפת. נסמן את הצפיפות של X בהינתן Y על ידי

$$f_{X|Y=y}(x)=\frac{f_{X,Y}(x,y)}{f_Y(y)}$$

עבור כל $y\in\mathbb{R}$ בו $f_Y(y)$ רציפה אשר מקיים $f_Y(y)\neq 0$.

טענה: (נוסחת הסתברות שלמה רציפה)

לכל X,Y בעלי צפיפות משותפת מתקיים

$$f_X(x)=\int_{y\in\mathbb{R}} f_Y(y)\, f_{X|Y=y}(x)\, dy$$

כאשר נפרש את הביטוי באינטגרל כשווה לאפס אם $f_Y(y)=0$.

הוכחה: נרשום את הנוסחה ל-$f_X(x)$ כהתפלגות שולית ונציב את הגדרת הצפיפות המותנית

$$f_X(x)=\int_{y\in\mathbb{R}} f_{X,Y}(x,y)\ dy=\int_{y\in\mathbb{R}}\frac{f_{X,Y}(x,y)}{f_Y(y)} f_Y(y)\, dy=\int_{y\in\mathbb{R}} f_Y(y) f_{X|Y=y}(x)\, dy$$

■

אבחנה: (כלל בייס רציף)

לכל X,Y בעלי צפיפות משותפת מתקיים

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

כאשר נתייחס לאגף שאינו מוגדר מפאת אי קיומה של צפיפות מותנית כשווה ל-0.

אבחנה: (תנאי לאי-תלות)

מ"מ X ו-Y רציפים בהחלט הנם בלתי תלויים אם ורק אם

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

לכל y עבורו $f_Y(y)\neq 0$

הוכחה: העובדה ש-X ו-Y ב"ת שקולה לפי האבחנה הזו לקיום פונקציית צפיפות משותפת

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

נצמצם ונקבל $f_{X|Y=y}(x)=f_X(x)$ כנדרש.

■

תוחלת תחת התנייה:

יהיו X,Y מ"מ בדידים אזי $\mathbb{E}[Y] = \mathbb{E}\big[\mathbb{E}[Y|X]\big]$.

<u>הוכחה:</u>

נסמן $\mathbb{E}[Y|X] = \varphi(X)$ אז ממשפט הסטטיסטיקאי נקבל:

$$\mathbb{E}[\varphi(X)] = \sum_{Im(X)} \varphi(X) \cdot \mathbb{P}\{\omega | X(\omega) = X\}$$

מכיון שאנו עובדים בקבוצות סופיות, נגדיר $Im(X) = \{X_1, \dots, X_n\}$.

נגדיר:

$$A_1 = \{\omega | X(\omega) = X_1\}$$

$$\vdots$$

$$\mathrm{A}_n = \{\omega | X(\omega) = X_n\}$$

מתקיים כי $\Omega = \{A_1 \cup \dots \cup A_n\}$ כאשר $A_i \cap A_j = \emptyset$. כעת נוכל לחשב:

$$\mathbb{E}[\varphi(X)] = \sum_{i=1}^{n} \varphi(X_i) \cdot \mathbb{P}\{\omega | X(\omega) = X_i\} = \sum_{i=1}^{n} \varphi(X_i) \cdot \mathbb{P}\{A_i\} = \sum_{i=1}^{n} \mathbb{E}[Y | X = X_i] \cdot \mathbb{P}\{A_i\}$$

$$= \sum_{i=1}^{n} \mathbb{E}[Y|A_i] \cdot \mathbb{P}(A_i) \underset{\text{תוחלת שלמה}}{=} \mathbb{E}[Y]$$

■

שרשראות מרקוב:

הגדרה –

תהי $(X_n)_{n=0}^{\infty}$ שרשרת מרקוב עם $X = [0, \dots, m]$ או $X = [0,1, \dots]$ אזי הסתברות מעבר מסדר k הינה:

$$\mathbb{P}(X_{n+k} = j | X_n = i)$$

נסמן $P_{i,j}^{(k)}$ (i הינו המיקום הקודם, j הינו לאן רוצים לעבור)

משפט צפמן קולמוגרוב:

נוסח: תהי $(X_n)_{n=0}^{\infty}$ שרשרת מרקוב עם קבוצת מצבים $\mathrm{X} = [0, \dots, M]$ אזי

$\mathrm{P}_{\mathrm{i,j}}^{(n)} = \sum_{l=0}^{M} P_{i,l}^{(R)} \cdot P_{j,l}^{(n-R)}$ כאשר $i, j \in [0, \dots, M]$ ו- $R \in [0, \dots, n-1]$

<u>הוכחה:</u>

$$P_{i,j}^{(n)} = \mathbb{P}(X_{n+k} = j | X_k = i) = \mathbb{P}(X_n = j | X_0 = i) = \sum_{l=0}^{M} \mathbb{P}(X_n = j, X_R = l | X_0 = i)$$

$$= \sum_{l=0}^{M} \frac{\mathbb{P}(X_n = j \cap X_R = l \cap X_0 = i)}{\mathbb{P}(X_R = l \cap X_0 = i)} \cdot \frac{\mathbb{P}(X_R = l \cap X_0 = i)}{\mathbb{P}(X_0 = i)}$$

$$= \sum_{l=0}^{M} \mathbb{P}(X_n = j | X_R = l, X_0 = i)$$

$$\cdot \mathbb{P}(X_R = l | X_0 = i) \underset{\text{מרקוב}}{=} \sum_{l=0}^{M} \mathbb{P}(X_n = j | X_R = l) \cdot \mathbb{P}(X_R = l | X_0 = i)$$

$$= \sum_{l=0}^{M} P_{l,j}^{(n-R)} \cdot P_{i,l}^{(R)}$$

כנדרש.

דוגמאות נפוצות:

The matching problem:

n אנשים משליכים איש את כובעו לכד. אחר כך מחלקים באקראי כובע לכל אחד.

נגדיר מרחב הסתברות:

$$\Omega = \{(l_1, \dots, l_n) | (1, \dots, n) \text{ תמורה של } (l_1, \dots, l_n)\}, \qquad |\Omega| = n!$$

מרחב ההסתברות אחיד ולכן $\mathbb{P}(A) = \frac{|A|}{n!}$. נגדיר:

$A_1 =$ איש 1 מקבל את כובע 1.

$A_2 =$ איש 2 מקבל את כובע 2.

$\vdots$

$\mathrm{A}_n =$ איש n מקבל את כובע n.

<u>מהי ההסתברות שאף אחד לא מקבל את כובעו?</u>

מכיוון שההסתברות אחידה מתקיים $\mathbb{P}(l_1 \cap \dots \cap l_n) = \frac{1}{n!}$

נשים לב כי:

$(A_1 \cup \dots \cup A_n)$ – לפחות אחד קיבל את הכובע המתאים לו.

$(A_1 \cap \dots \cap A_n)$ – כולם קיבלו את הכובע המתאים להם.

$(A_1 \cup \dots \cup A_n)^c$ – אף אחד לא מקבל את הכובע שלו. – מה שאנחנו מחפשים.

נשים לב כי $\mathbb{P}((A_1 \cup \dots \cup A_n)^c) = 1 - \mathbb{P}(A_1 \cup \dots \cup A_n)$

מעקרון הכלה והדרה מתקיים:

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = s_1 - s_2 + \dots + (-1)^{n-1} s_n$$

כאשר $s_k = \sum_{1 \le l_1 < .. < l_k} \mathbb{P}\left(A_{l_1} \cap \dots \cap A_{l_k}\right)$

$$\mathbb{P}\left(A_{l_1} \cap \dots \cap A_{l_k}\right) = \frac{\left|A_{l_1} \cap \dots \cap A_{l_k}\right|}{|\Omega|}$$

המאורע $\left(A_{l_1} \cap \dots \cap A_{l_k}\right)$ הינו המאורע שבו האנשים $l_1, \dots, l_k$ קיבלו את הכובע שלהם ואילו השאר לא מעניינים אותנו, ולכן $\left|A_{l_1} \cap \dots \cap A_{l_k}\right| = (n-k)!$.

מכאן נקבל $\mathbb{P}\left(A_{l_1} \cap \dots \cap A_{l_k}\right) = \frac{(n-k)!}{n!}$, ולכן:

$$s_k = \sum_{1 \le l_1 < \dots < l_k \le n} \mathbb{P}\left(A_{l_1} \cap \dots \cap A_{l_k}\right) = \binom{n}{k} \cdot \frac{(n-k)!}{n!} = \frac{n!}{(n-k)! \cdot k!} \cdot \frac{(n-k)!}{n!} = \frac{1}{k!}$$

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \dots + \frac{(-1)^{n-1}}{n!}$$

נזכר בפיתוח טור טיילור של $1-e^{x}$:

$$1-\left(1-\frac{x}{1!}+\frac{x^2}{2!}-\cdots\right)=\frac{x}{1!}-\frac{x^2}{2!}+\frac{x^3}{3!}-\cdots$$

כלומר:

$$1-e^{-1}=1-\frac{1}{2!}+\frac{1}{3!}-\cdots$$

$$\lim_{n\to\infty}\mathbb{P}(A_1\cup\ldots\cup A_n)=1-e^{-1}\simeq 0.632$$

ולכן ההסתברות המבוקשת הינה 0.368.

<u>מהי ההסתברות שבדיוק m אנשים יקבלו את כבעם?</u>

נחפש את p_m = ההסתברות שבדיוק m מאורעות יקרו. מהמשפט נקבל:

$$\begin{aligned}p_m=s_m-\binom{m+1}{m}s_{m+1}+\cdots+(-1)^{n-m}\binom{n}{m}s_n\\&=\frac{1}{m!}-\binom{m+1}{m}\cdot\frac{1}{(m+1)!}+\cdots+(-1)^{n-m}\cdot\frac{1}{n!}\\&=\frac{1}{m!}-\frac{1}{1!\,m!}+\cdots+(-1)^{n-m}\cdot\frac{1}{m(n-m)!}\\&=\frac{1}{m!}\left(1-\frac{1}{1!}+\frac{1}{2!}-\frac{1}{3!}+\cdots+\frac{(-1)^{n-m}}{(n-m)!}\right)\end{aligned}$$

במקרה שבו $n\to\infty$ מתקיים:

$$p_m=\frac{1}{m!}\underbrace{\left(1-\frac{1}{1!}+\frac{1}{2!}-\frac{1}{3!}+\cdots+\frac{(-1)^{n-m}}{(n-m)!}\right)}_{\text{טור טיילור של } e^{-1}}$$

The laplace problem:

בהנתן n כדים, עבור $0 \le k \le n$ מתקיים שבכד ה- k יש k כדורים אדומים, וגם $n-k$ כדורים לבנים.

נבחר כד באקראי, ואז שולפים ומחזירים בזה אחר זה $N+1$ כדורים.

מהי ההסתברות שהכדור ה- $N+1$ הוא אדום בהנתן שכל N הכדורים הראשונים גם אדומים?

פתרון:

נגדיר:

A – N כדורים ראשונים הם אדומים.

B – כדור $N+1$ הוא אדום.

j - לוקחים כדור מכד j, אזי מההסתברות השלמה:

$$\mathbb{P}(A)=\sum_{j=0}^{n}\mathbb{P}(A|j)\cdot\mathbb{P}(j)$$

נשים לב כי $\mathbb{P}(j)=\frac{1}{n+1}$ – בחירת הכד ה j מתוך $n+1$ כדים.

בכל כד יש j כדורים אדומים ו- $n-j$ כדורים לבנים, ומכיון שבכל שליפה ההסתברות להוציא כדור אדום היא $\frac{j}{n}$, ואנו מבצעים אותה N פעמים, נקבל: $\mathbb{P}(A|j)=\frac{j^N}{n^N}$

לכן:

$$\mathbb{P}(A)=\sum_{j=0}^{n}\frac{j^N}{n^N}\cdot\frac{1}{n+1}=\frac{1}{n+1}\cdot\sum_{j=0}^{n}\frac{j^N}{n^N},\quad \mathbb{P}(B)=\frac{1}{n+1}\cdot\sum_{j=0}^{n}\frac{j^{N+1}}{n^{N+1}}$$

אנו רוצים למצוא את $\mathbb{P}(B|A)=\frac{\mathbb{P}(B\cap A)}{\mathbb{P}(A)}$

נשים לב ש- $\mathbb{P}(B\cap A)$ זה ההסתברות שיצאו גם N אדומים וגם $N+1$ אדומים,

כלומר $\mathbb{P}(B\cap A)=\mathbb{P}(B)$. לכן:

$$\mathbb{P}(B|A)=\frac{\frac{1}{n+1}\cdot\underbrace{\sum_{j=0}^{n}\left(\frac{j}{n}\right)^{N+1}}_{\text{סכום רימן}}}{\frac{1}{n+1}\cdot\underbrace{\sum_{j=0}^{n}\left(\frac{j}{n}\right)^{N}}_{\text{סכום רימן}}}\xrightarrow[n\to\infty]{}\frac{\int_0^1 x^{N+1}dx}{\int_0^1 x^N dx}=\frac{N+1}{N+2}$$

The Gambler's Ruin Problem:

בידי שחקן יש $x\$$. כללי המשחק הם:

מטבע הוטל. אם יצא H השחקן מקבל $1\$$, אם יצא T השחקן מפסיד $1\$$.

המשחק מסתיים כאשר:

1. השחקן מפסיד הכל.
2. לשחקן יש $m\$$ כאשר $m > x$.

<u>מה ההסתברות שהשחקן יפסיד הכל?</u>

נגדיר:

$p(x)$ – ההסתברות שהשחקן יפסיד הכל, כאשר הוא מתחיל עם $x\$$.

A – המאורע "השחקן הפסיד הכל כאשר התחיל עם $x\$$". $p(x) = \mathbb{P}(A)$.

B_1 – המאורע "השחקן קיבל H בהטלה הראשונה של המטבע", כל המסלולים שעוברים דרך $x+1$.

B_2 - המאורע "השחקן קיבל T בהטלה הראשונה של המטבע", כל המסלולים שעוברים דרך $x-1$.

נשים לב כי $\Omega = B_1 \cup B_2$, עבור $B_1 \cap B_2 = \emptyset$. ולכן מנוסחת ההסתברות השלמה:

$$\mathbb{P}(A) = \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\mathbb{P}(B_2)$$

כעת נוכל להציב:

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

אנו יודעים כי $p(o) = 1$, $p(m) = 0$, כעת נקבל:

$$p(x) = c_1 + c_2x = \frac{1}{2}\big(c_1 + c_2(x+1) + c_1 + c_2(x-1)\big)$$

$$1 = p(0) = c_1 \;\;\Rightarrow\;\; 0 = p(m) = \frac{1}{2}\big(1 + c_2(m+1) + 1 + c_2(m-1)\big) = \frac{1}{2}(2 + 2c_2m)$$

$$= 1 + c_2\cdot m = 0 \;\;\Rightarrow\;\; c_2 = -\frac{1}{m}$$

נציב בנוסחה ונקבל:

$$p(x) = 1 - \frac{x}{m} \;\;\xrightarrow[m\to\infty]{}\;\; 1$$

Polya Urn:

בכד ישנם b כדורים שחורים ו- r כדורים אדומים.

כדור נשלף באקראי. אם הכדור הנשלף הוא שחור אז מחזירים אותו ומוסיפים עוד c כדורים שחורים, אם הכדור הנשלף הוא אדום אז מחזירים אותו ומוסיפים עוד c כדורים אדומים. נניח כי מבצעים שליפה n פעמים.

<u>מה ההסתברות שכל n הכדורים יהיו שחורים?</u>

נגדיר את המאורעות B_i – בשליפה ה i יצא כדור שחור.

אם $n = 1$, אז $\mathbb{P}(B_1) = \frac{b}{b+r}$.

אם $n = 2$, אז:

$$\mathbb{P}(B_1 \cap B_2) = \mathbb{P}(B_2|B_1) \cdot \mathbb{P}(B_1) = \frac{b+c}{b+r+c} \cdot \frac{b}{b+r}$$

אם $n = 3$, אז:

$$\mathbb{P}(B_1 \cap B_2 \cap B_3) \underset{\text{כלל המכפלה}}{=} \mathbb{P}(B_3|B_2 \cap B_1) \cdot \mathbb{P}(B_2|B_1) \cdot \mathbb{P}(B_1)$$

$$= \frac{b+2c}{b+r+2c} \cdot \frac{b+c}{b+r+c} \cdot \frac{b}{b+r}$$

ובאופן כללי נקבל:

$$\mathbb{P}(B_n \cap \dots \cap B_1) = \frac{b+(n-1)c}{b+r+(n-1)c} \cdot \frac{b+(n-2)c}{b+r+(n-2)c} \cdot \dots \cdot \frac{b}{b+r}$$

דוגמא 5 מבצעים את הניסוי הבא: זורקים מטבע הוגן, הנופל על ראש בהסתברות p, בזה אחר זה באופן ב"ת. יהיו $n,m \in \mathbb{N}$. מה ההסתברות שנקבל n ראשים לפני שנקבל m פעמים פאלי?

הפיתרון של פרמה נבחין שתנאי מספיק והכרחי על מנת שיהיו n ראשים לפני שהיו m פאלי, הוא שהיו לפחות n ראשים מתוך ה $n+m-1$ הנסיונות הראשונים. מדוע זה נכון? מכיוון שאם היו לפחות n ראשים מתוך ה $n+m-1$ נסיונות הראשונים, אז היו לכל היותר $m-1$ פאלי בנסיונות הללו. מכאן שהיו n ראשים לפני שהיו m פאלי. מאידך, אם היו פחות מ n ראשים ב $n+m-1$ נסיונות הראשונים, אז היו לפחות m פאלי שיצאו שם, ולכן לא קיבלנו n ראשים לפני שקיבלנו m פאלי.

כעת, יהא (Ω, P) מ"ה המתאים לשאלה (אפשר למשל לקחת מרחב מכפלה כמו בשאלה הקודמת עבור $n+m-1$ נסיונות). יהא $A_{n,m}$ המאורע "קיבלנו n ראשים לפני שקיבלנו m פאלי". יהא A_k המאורע "קיבלנו בדיוק k ראשים מתוך ה $n+m-1$ נסיונות הראשונים". אז לפי הדוגמא הקודמת,

$$P(A_k) = \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}$$

וכן לפי הדיון הקודם,

$$A_{n,m} = \bigcup_{k=n}^{m+n-1} A_k$$

מכיוון שמדובר באיחוד זר (מדוע זה איחוד זר?) נקבל

$$P(A_{n,m}) = \sum_{k=n}^{m+n-1} P(A_k) = \sum_{k=n}^{m+n-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}$$

:Random Walk

יש חלקיק על גרף כאשר $a_i = \pm 1$.

$$\mathbb{P}(\{\omega\}) \stackrel{def}{=} p^{\sum_{i=1}^{n} a_i} q^{\left(n - \sum_{i=1}^{n} a_i\right)}$$

נראה כי $\sum_{\omega\in\Omega} p(\omega) = 1$:

$$\sum_{\omega\in\Omega} p(\omega) = \sum_{\substack{k=-n \\ a_1+\cdots+a_n=k}}^{n} \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}} = \sum_{m=0}^{n} \binom{n}{k} p^m q^{n-m} = (p+q)^n = 1$$

כעת נגדיר מאורע A_0 – החלקיק חוזר לנקודה 0 אחרי n שלבים.

נניח כי $n = 2N$, $p = q = \frac{1}{2}$, ונחשב את $\mathbb{P}(A)$ כאשר $n \to \infty$:

נזכר בנוסחת סטרלינג:

$$N! \simeq \sqrt{2\pi N} \cdot e^{-N} \cdot N^N$$

כעת נוכל לחשב:

$$\mathbb{P}(A_0) = \underbrace{\binom{2N}{n} \left(\frac{1}{2}\right)^N \cdot \left(\frac{1}{2}\right)^N}_{\text{נוסחת ברנולי}} = \frac{(2N)!}{N! \cdot N!} \cdot \frac{1}{2^{2N}} \simeq \frac{\sqrt{2\pi 2N} \cdot e^{-2N} \cdot (2N)^{2N}}{\sqrt{2\pi N} \cdot e^{-N} \cdot N^N \cdot \sqrt{2\pi N} \cdot e^{-N} \cdot N^N} = \frac{1}{\sqrt{\pi N}}$$

נסיק כי עבור $N \to \infty$ נקבל כי $\mathbb{P}(A_0) \simeq \frac{1}{\sqrt{\pi N}}$.

:Waiting time

נגדיר את הניסוי הבא, הטלת מטבע n פעמים. מרחב המדגם הינו $\Omega = \{\omega = (a_1, \dots, a_n) | a_i \in \{0,1\}\}$,

$$\mathbb{P}(\{\omega\}) = p^{\sum_{i=1}^{n} a_i} \cdot q^{\left(n-\sum_{i=1}^{n} a_i\right)}, \qquad p+q=1, \qquad p,q \geq 0$$

זמן ההמתנה הינו משתנה מקרי: מספר האפסים שקיבלנו 0 עד שקיבלנו בפעם הראשונה 1.

$$supp(X) = \{0,1 \dots, n\}$$

נרצה לחשב את ההתפלגות. כלומר $p_x(k) = \mathbb{P}(\{\omega | X(\omega) = k\})$.

מקרה ראשון - $0 \leq k \leq n-2$:

$$\mathbb{P}\{\omega : X(\omega) = k\} = \mathbb{P}\left\{\omega : \omega = \left(\underbrace{0, \dots, 0}_{k\ times}, \underbrace{1}_{k+1}, \underbrace{*, \dots, *}_{n-k+1}\right)\right\} =$$

$$= \underbrace{\sum_{a_{k+2}=0}^{1} \dots \sum_{a_n=0}^{1} p^{1+\sum_{j=k+2}^{n} a_j} q^{n-1-\sum_{j=k+2}^{n} a_j}}_{(*)} =$$

(*) - אנו מחשבים לכל $k+2 \leq i \leq n$ את האפשרויות של a_i. a_i יכול להיות או 1 או 0, כלומר סה"כ נחשב את מספר האפשרויות לקבלת הביטוי $\underbrace{*, \dots, *}_{n-k+1}$, נשים לב שבמקום ה $k+1$ אנו יודעים בוודאות שיש 1, ולפניו יש רק אפסים, לכן נחשב את החזקה של p לפי מספר ה 1 שהופיעו אחרי המקום ה $k+1$ ועוד 1, ובהתאם את q.

נסמן את מספר הפעמים שקיבלנו 1 אחרי ההטלה ה $k+1$. נשים לב שהבעיה שקולה לפיזור l כדורים מתוך קבוצה של $n-k-1$ כדורים, עבור $0 \leq l \leq n-k-1$, ולכן:

$$= \sum_{l=0}^{n-k-1} p^{1+l} q^{n-1-l} \binom{n-k-1}{l} = pq^k \underbrace{\sum_{l=0}^{n-k-1} \binom{n-k-1}{l} p^l q^{n-k-l-1}}_{(p+q)^{n-k-1}=1} = pq^k$$

מקרה שני - $k = n-1$:

$$\mathbb{P}\{\omega : X(\omega) = n-1\} = \mathbb{P}\left\{\omega : \omega = \left(\underbrace{0, \dots, 0}_{n-1}, 1\right)\right\} = pq^{n-1}$$

מקרה שלישי - $k = n$:

$$\mathbb{P}\{\omega : X(\omega) = n\} = \mathbb{P}\{\omega : \omega = (0, \dots, 0)\} = q^n$$

סה"כ נקבל:

$$\mathbb{P}_X(k) = \begin{cases} pq^k & 0 \leq k \leq n-1 \\ q^n & k = n \end{cases}$$

לא לשכוח לבדוק שמתקיים $\sum_{k=0}^{n} p_X(k) = 1$:

$$\sum_{k=0}^{n} p_X(k) = p \cdot \frac{q^n - 1}{q-1} + q^n = \frac{pq^n - p + q^{n+1} - q^n}{q-1} = \frac{q^n \underbrace{(p+q)}_{=1} - p - q^n}{q-1}$$

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

נחשב את התוחלת של X:

$$\mathbb{E}[X] = \sum_{k=0}^{n} k \cdot \mathbb{P}_X(k) = \sum_{k=0}^{n-1}\big(k \cdot \mathbb{P}_X(k)\big) + k \cdot \mathbb{P}_X(n) = \sum_{k=0}^{n-1} k \cdot pq^k + nq^n$$
$$= pq\sum_{k=0}^{n-1}\left(k \cdot q^{k-1}\right) + nq^n = (*)$$

כעת נוכיח את הטענה הבאה:

$$\sum_{k=0}^{n-1}\left(k \cdot q^{k-1}\right) = \frac{(n-1)q^n - nq^{n-1} + 1}{(q-1)^2}$$

הוכחה:

$$\sum_{k=0}^{n-1}\left(q^k\right) = 1 + q + q^2 + \cdots + q^{n-1} = \frac{q^n - 1}{q - 1}$$
$$\frac{d}{dq}\sum_{k=0}^{n-1} q^n = \sum_{k=0}^{n-1}\left(k \cdot q^{k-1}\right)$$
$$\frac{d}{dq}\left(\frac{q^n - 1}{q - 1}\right) = \frac{nq^{n-1}(q-1) - (q^n - 1)}{(q-1)^2} = \frac{(n-1)q^n - nq^{n-1} + 1}{(q-1)^2}$$

מכאן נקבל את הטענה, כנדרש. כעת נקבל כי התוחלת המבוקשת הינה:

$$(*) = pq \cdot \frac{(n-1)q^n - nq^{n-1} + 1}{(q-1)^2} + nq^n$$

נרצה לחשב את $\lim_{n\to\infty} \mathbb{E}[X]$:

נשים לב כי:

$\lim_{n\to\infty} q^n = 0$, $\lim_{n\to\infty} nq^n = 0$ כאשר $0 \le q \le 1$. ולכן נקבל:

$$\lim_{n\to\infty} \mathbb{E}[X] = \lim_{n\to\infty}\left(pq \cdot \frac{(n-1)q^n - nq^{n-1} + 1}{(q-1)^2} + nq^n\right) = pq\,\frac{(n-1)\cdot 0 - 0 + 1}{(q-1)^2} + 0$$
$$= \frac{pq}{(q-1)^2}$$