דף נוסחאות

דניאל גרימלנד (מורחב ל2025)

קומבינטוריקה

עקרון הסכום: מתקיים:

$$|A| = |A_1| + \cdots + |A_n|, \text{ for } A = \bigsqcup_{i=1}^{n} A_i$$

עקרון המכפלה: מתקיים:

$$|A_1 \times \cdots \times A_n| = \prod_{i=1}^{n} |A_i|$$

זו טענה שימושית בשביל להראות שהעוצמה של קבוצה היא מכפלה: מנחסים את הקבוצה במונחים של מכפלה קרטזית.

מקדם מולטינומי: מתאר כמה דרכים יש לחלק קבוצה בגודל n לתת קבוצות בעל גדלים m_i

$$\binom{n}{m_1, \ldots, m_k} = \frac{n!}{m_1! \cdot \cdots \cdot m_k!}$$

נוסחאות בחירה: מספר הדרכים לבחור k פריטים מתוך n פריטים תלוי בנסיבות של הבחירה, אך מתחלק לארבעת המקרים הבאים:

פרמטרים	עם חזרה	בלי חזרה
עם סדר	n^k	$\frac{n!}{(n-k)!}$
בלי סדר	$\binom{n+k-1}{k}$	$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

הבינום של ניוטון וזהויות בינומיות: מתקיימות הזהויות הבאות:

$$(x+y)^n = \sum_{i=1}^{n} \binom{n}{i} x^{n-i} y^i$$

$$\sum_{i=1}^{n} \binom{n}{i} = 2^i$$

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

$$\binom{n}{k} = \binom{n}{n-k}$$

עקרון ההכלה וההדחה: העוצמה של איחוד סופי כללי היא:

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^{n} |A_k|$$

$$- \sum_{1 \le k < j \le n} |A_k \cap A_j| + \sum_{1 \le k < h < j \le n} |A_k \cap A_h \cap A_j| \ldots$$

ובמקרה הסימטרי (גודל החיתוך תלוי רק בכמה מחתכים) מתקיים:

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} |A_1 \cap \cdots \cap A_k|$$

מרחבי הסתברות בדידים וכלליים

מרחב הסתברות כללי: מרחב הסתברות הוא שלשה $(\Omega, \mathcal{F}, P)$ כך ש Ω היא קבוצת התוצאות האפשריות, $\mathcal{F}$ היא סיגמא-אלגברת המאורעות (קבוצה של תתי-קבוצות של Ω שמכילות את הקבוצה הריקה, סגורה למשלים, וסגורה לאיחוד וחיתוך בני מנייה), ו P היא פונקציה מ $\mathcal{F}$ לממשיים החיוביים כך שמתקיים:

$$P(A) = \sum_{i=1}^{\infty} P(A_i), \text{ for } A = \bigsqcup_{i=1}^{\infty} A_i \qquad P(\Omega) = 1$$

מרחב הסתברות בדיד: במקרה הבדיד, $\mathcal{F} = \mathcal{P}(\Omega)$. נסמן (Ω, P).

תכונות של מרחבי הסתברות:

1. ההסתברות של מאורע היא תמיד בין 0 ל 1.
2. אם $A \subseteq B \in \mathcal{F}$, אז $P(A) \le P(B)$.
3. אם $A \subseteq B \in \mathcal{F}$, אז $P(B \setminus A) = P(B) - P(A)$.
4. חסם האיחוד: עבור $A_n \in \mathcal{F}$:

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

5. רציפות ההסתברות: אם $A_n \in \mathcal{F}$ סדרת מאורעות עולה או יורדת (בהכלה) אז מתקיים בהתאמה:

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

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

הכלה והדחה :עבור $A_1, \ldots, A_n \in \mathcal{F}$, מתקיים:

$$P(A_1 \cup \cdots \cup A_n) = \sum_{I \subseteq \{1, \ldots, n\}} (-1)^{|I|+1} P\left(\bigcap_{i \in I} A_i\right)$$

הסתברות מותנית: עבור $A, B_1, \ldots, B_n \in \mathcal{F}$:

1. מגדירים, בהנחה שלא מחלקים באפס:

$$P(A \mid B_1, \ldots, B_n) = \frac{P(A \cap B_1 \cap \cdots \cap B_n)}{P(B_1 \cap \cdots \cap B_n)}$$

2. ההסתברות של חיתוך כללי של מאורעות יכולה להיות מחושבת ע"י הסתברות מותנית (לא לצטט בלי הוכחה!):

$$P(B_1 \cap \cdots \cap B_n) =$$
$$P((B_1) P(B_2 \mid B_1) \cdot \cdots \cdot P(B_n \mid B_{n-1}, \ldots, B_1))$$

נוסחת ההסתברות השלמה: עבור חלוקה של Ω לסדרת מאורעות $A_n \in \mathcal{F}$, ומאורע אחר $B \in \mathcal{F}$, מתקיים:

$$P(B) = \sum_{i=1}^{\infty} P(B \mid A_i) P(A_i)$$

נוסחת ההיפוך (חוק בייס): עבור $A, B \in \mathcal{F}$:

$$P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$$

תלות ואי-תלות של מאורעות: עבור $A, B \in \mathcal{F}$:

1. מגדירים שהם ב"ת אם $P(A \cap B) = P(A) P(B)$.

2. שקילות לאי-תלות:

$$P\left(A\,|\,B\right)=P\left(A\right),\text{ or }P\left(B\,|\,A\right)=P\left(B\right)$$

3. מגדירים שהם ב"ת בתנאי $C\in\mathcal{F}$ אם:

$$P\left(A\cap B\,|\,C\right)=P\left(A\,|\,C\right)P\left(B\,|\,C\right)$$

4. שקילות לאי-תלות בתנאי:

$$P\left(A\,|\,B,C\right)=P\left(A\,|\,C\right)P,\text{ or }P\left(B\,|\,A,C\right)=P\left(B\,|\,C\right)$$

תלות ואי-תלות במשותף: עבור $A_1,\dots,A_n\in\mathcal{F}$, נגדיר שהם בלתי תלוים במשותף אם:

$$\forall\alpha_1,\dots,\alpha_n\in\{1,c\}:P\left(A_1^{\alpha_1}\cap\dots\cap A_n^{\alpha_n}\right)=P\left(A_1^{\alpha_1}\right)\cdot\dots\cdot P\left(A_n^{\alpha_n}\right)$$

כל תת-קבוצה של מאורעות ב"ת במשותף היא גם ב"ת במשותף.

משתנים מקריים בדידים

משתנה מקרי בדיד: משתנה מקרי בדיד היא פונקציה $X:\Omega\to\mathbb{R}$ עבור מרחב הסתברות בדיד ברקע (Ω,P).

התפלגות של מ"מ: ההתפלגות של מ"מ X היא פונקציה ממשית המוגדרת לפי הכלל $a\longmapsto P\left(X=a\right)$. מדובר במדד החשוב ביותר של מ"מ, עד כדי שכמעט כל ניתוח שעושים עליהם מתבצע ברמת ההתפלגות. מסיבה זו פעמים רבות נקבל מ"מ רק בתור התפלגות, ולא פונקציות.

מ"מ בדיד מותנה במאורע: עבור מאורע A עם הסתברות חיובית, $X\,|\,A$ מציין את המ"מ X תחת התנייה ב A. בכך משתנה ההתפלגות שלו לכדי $a\longmapsto P\left(X=a\,|\,A\right)$.

מ"מ בדיד מותנה במ"מ בדיד אחר: עבור שני מ"מ X,Y מעל אותו מרחב, $X\,|\,Y$ מציין את המ"מ X כאשר הוא מקבל את ההפלגות שלו לכל ערך של Y. כלומר, לכל ערך של Y המשתנה $X\,|\,Y$ מקבל התפלגות, ההתפלגות המותנית במאורע $Y=b$.

אלגברת מ"מ: כל פונקציה של כל מספר של מ"מ היא גם כן מ"מ. אפשר גם לסכום מ"מ מותנים במאורע ומותנים במ"מ אחר באותה צורה בתנאי שההתנייה היא באותו דבר. כמה דוגמאות לכך:

$$\begin{aligned}&X+Y+\cos^{W}\left(Z\right)\\&\left(X\,|\,A\right)^2=X^2\,\big|\,A\\&X\,|\,Y+Z\,|\,Y=X+Z\,|\,Y\end{aligned}$$

אי-תלות של מ"מ: עבור סדרת מ"מ $X_1\dots,X_n$(עבור התנייה במאורע זה דומה, ועבור התנייה במשתנה דורשים אי-תלות לכל ערך של המ"מ המתנה) מעל אותו מרחב:

1. הם נקראים ב"ת במשותף אם:

$$P\left(X_1=a_1,\dots,X_n=a_n\right)=P\left(X_1=a_1\right)\cdot\dots\cdot P\left(X_n=a_n\right)$$

2. הם נקראים ב"ת בזוגות אם כל זוג שונה של מ"מ בסדרה הוא ב"ת במשותף.

אם $X_1\dots,X_n$ ב"ת במשותף, אז כל פונקציות שלהם הן ב"ת במשותף, בתנאי שכל משתנה מופיע לכל היותר בפונקציה אחת. למשל $X_1X_2+X_3,\ X_4^{X_5}$ הם ב"ת, וגם $f_1\left(X_1\right),\dots,f_n\left(X_n\right)$ הם ב"ת במשותף.

תוחלת של מ"מ:

1. התוחלת של מ"מ בדיד X מוגדרת בתור:

$$\mathbf{E}\left(X\right)=\mathbb{E}\left(X\right)=\sum_{a\in\mathbb{R}}aP\left(X=a\right)=\sum_{\omega\in\Omega}\omega P\left(X=\omega\right)$$

2. התוחלת של $X\,|\,A$ מוגדרת באופן זהה רק עם הסתברות מותנית ב A.

3. התוחלת של $X\,|\,Y$ היא מ"מ שהוא פונקציה של Y, כי לכל ערך של Y ההתפלגות נקבעת ויש תוחלת.

תכונות של התוחלת של מ"מ: עבור מ"מ $X_1,\dots,X_n$ מעל אותו המרחב (דומה עבור משתנים מקריים מותנים):

1. $\mathbf{E}\left(c\right)=c$

2. מתקיים:

$$\mathbf{E}\left(f\left(X_1,\dots,X_n\right)\right)=\sum_{a_1,\dots,a_n\in\mathbb{R}}f\left(a_1,\dots,a_n\right)P\left(X_1=a_1,\dots,X_n=a_n\right)$$

3. $\mathbf{E}\left(a_1X_1+\dots+a_nX_n\right)=a_1\mathbf{E}\left(X_1\right)+\dots+a_n\mathbf{E}\left(X_n\right)$

4. אם המשתנים ב"ת במשותף מתקיים:

$$\mathbf{E}\left(X_1\cdot\dots\cdot X_n\right)=\mathbf{E}\left(X_1\right)\cdot\dots\cdot\mathbf{E}\left(X_n\right)$$

חוק התוחלת החוזרת: עבור מ"מ X,Y מעל אותו מרחב:

$$\mathbf{E}\left(X\right)=\mathbf{E}\left(\mathbf{E}\left(X\,|\,Y\right)\right)$$

זה הגיוני, כי מימין מדובר בתוחלת של מ"מ.

חוק התוחלת השלמה: עבור מ"מ X וחלוקה של המרחב למאורעות $A_n\in\mathcal{F}$, מתקיים:

$$\mathbf{E}\left(X\right)=\sum_{i=1}^{\infty}\mathbf{E}\left(X\,|\,A_i\right)P\left(A_i\right)$$

שונות של מ"מ: השונות של מ"מ בדיד X (בדומה עבור מ"מ מותנה לפי ההגדרה עבור התוחלת) מוגדרת בתור:

$$\mathbf{V}\left(X\right)=\mathrm{Var}\left(X\right)=\mathbf{E}\left(\left(X-\mathbf{E}\left(X\right)\right)^2\right)=\mathbf{E}\left(X^2\right)-\mathbf{E}^2\left(X\right)$$

תכונות של השונות של מ"מ: עבור מ"מ X,Y מעל אותו המרחב (דומה עבור משתנים מקריים מותנים):

1. השונות תמיד אי-שלילית, ושווה אפס אם ורק אם המ"מ הוא קבוע בהסתברות 1.

2. $\mathbf{V}\left(X+a\right)=\mathbf{V}\left(X\right)$

3. $\mathbf{V}\left(aX\right)=a^2\mathbf{V}\left(X\right)$

4. $\mathbf{V}\left(X+Y\right)=\mathbf{V}\left(X\right)+\mathbf{V}\left(Y\right)+2\mathrm{Cov}\left(X,Y\right)$

5. נוסחאת פירוק השונות:

$$\mathbf{V}\left(X\right)=\mathbf{V}\left(\mathbf{E}\left(X\,|\,Y\right)\right)+\mathbf{E}\left(\mathbf{V}\left(X\,|\,Y\right)\right)$$

שונות משותפת ואי-תיאום: עבור מ"מ X,Y מאותו מרחב, נגדיר:

$$\begin{aligned}&\mathrm{Cov}\left(X,Y\right)=\mathbf{E}\left(XY\right)-\mathbf{E}\left(X\right)\mathbf{E}\left(Y\right)=\\&\mathbf{E}\left(\left(X-\mathbf{E}\left(X\right)\right)\left(Y-\mathbf{E}\left(Y\right)\right)\right)\end{aligned}$$

השונות המשותפת היא מכפלה פנימית על מרחב המנה של המ"מ תחת יחס השקילות: "כל שני מ"מ במרחק הזה בקבוע זה מזה שקולים". זו מסקנה שימושית מאוד להוכחת טענות מעניינות לגבי השונות המשותפת.

נגדיר שמ"מ $X_1,\dots X_n$ הם בלתי מתואמים אם השונות המשותפת של כל זוג מ"מ שונים מהם היא 0.

מתקיים שכל $X_1,\dots X_n$ ב"ת הם גם בלתי מתואמים.

תכונות של השונות המשותפת: עבור מ"מ X,Y מאותו מרחב:

1. $\mathrm{Cov}(\alpha_1+\beta_1 X,\alpha_2+\beta_2 Y)=\beta_1\beta_2\mathrm{Cov}(X,Y)$
2. $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
3. $\mathrm{Cov}(X,X)=\mathbf{V}(X)$

מקדם המתאם: עבור מ"מ X,Y מאותו מרחב, נגדיר:

$$\rho(X,Y)=\rho_{X,Y}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbf{V}(X)\mathbf{V}(Y)}}$$

תכונות של מקדם המתאם: עבור מ"מ X,Y מאותו מרחב:

1. $|\rho(X,Y)|\leq 1$ ושווה 1 אם ורק אם אחד הוא פונקציה לינארית של השני עם שיפוע חיובי, ושווה 1- אם ורק אם אחד הוא פונקציה של השני עם שיפוע שלילי.
2. $\rho(\alpha_1+\beta_1 X,\alpha_2+\beta_2 Y)=\frac{\beta_1\beta_2}{|\beta_1\beta_2|}\rho(X,Y)$

התפלגויות בדידות

התפלגות אחידה בדידה: התפלגות של מ"מ שמקבל ערכים בקבוצה $\{n,n+1,\ldots,m-1,m\}$ בהסתברות אחידה. משתנה מקרי בהתפלגות זו ($X\sim U[n,m]$) אם ורק אם:

$$P(X=a)=\begin{cases}\frac{1}{m-n+1} & x\in\{n,\ldots,m\}\\ 0 & \text{Otherwise}\end{cases}$$

התפלגות ברנולית: התפלגות של מ"מ שמקבל 1 בהסתברות p אם קרה משהו ו 0 אחרת. משתנה מקרי מתפלג כך ($X\sim b(p)$) אם ורק אם:

$$P(X=a)=\begin{cases}p & a=1\\ 1-p & a=0\\ 0 & \text{Otherwise}\end{cases}$$

התפלגות בינומית: התפלגות של מ"מ המציין את מספר ההצלחות ב n ניסויי ברנולי ב"ת עם הסתברות הצלחה של p. משתנה מקרי מתפלג כך ($X\sim Bin(n,p)$) אם ורק אם:

$$P(X=a)=\begin{cases}\binom{n}{a}p^aq^{n-a} & a\in\{0,\ldots,n\}\\ 0 & \text{Otherwise}\end{cases}$$

התפלגות פואסונית: התפלגות המתארת התפלגות בינומית בגבול, כאשר כמות הניסויים שואפת לאינסוף, ההסתברות להצלחה שואפת לאפס, אך הם עושים זאת תוך שמירה על תוחלת קבועה. ההתפלגות שימושית בהרבה מקרים שבהם מתרחשים מאורעות ב"ת בפרק זמן שאפשר לחלק לחלקים קטנים כרצוננו כאשר ידועה תוחלת מספר ההתרחשויות. משתנה מקרי מתפלג כך ($X\sim Poi(\lambda)$) אם ורק אם:

$$P(X=a)=\begin{cases}e^{-\lambda}\frac{\lambda^a}{a!} & a\in\mathbb{N}_0\\ 0 & \text{Otherwise}\end{cases}$$

התפלגות גיאומטרית: התפלגות של מ"מ המציין את מספר הניסויים עד להצלחה הראשונה בסדרת ניסויי ברנולי ב"ת פוטנציאלית אינסופית. משתנה מקרי מתפלג כך ($X\sim G(p)$) אם ורק אם:

$$P(X=a)=\begin{cases}pq^{a-1} & a\in\mathbb{N}\\ 0 & \text{Otherwise}\end{cases}$$

מ"מ בדיד X המקבל ערכים אי-שליליים נקרא חסר זיכרון אם לכל $n,m\in\mathbb{N}_0$ $P(X>n+m\mid X>n)=P(X>m)$. מ"מ בדיד מתפלג גיאומטרית עם פרמטר כלשהו אם ורק אם הוא חסר זיכרון.

התפלגות בינומית שלילית: מודד מספר הצלחות בסדרה של נסיונות ברנולי עד הכישלון מספר הז, מתקיים כי הפונקצית הסתברות היא:

$$P(X=k)=\binom{k+r-1}{k}p^kq^r$$

נשים לב שזה עבור כל k מ0 עד אינסוף.

התפלגות מולטינומית: מתאר *סדרה* של מ"מ ומסמנים: $(X_1,\ldots,X_k)\sim M(n;p_1,\ldots,p_k)$

כאשר יש את הסיטואציה הבאה, יש מכונה שמייצרת n כדורים וההסתברות שיצא כדור מסוג i הוא p_i , כל מ"מ מתאר כמה כדורים מהסוג נוצרו נשים לב:

$$P(X_1=a_1,\ldots,X_k=a_k)=\binom{n}{a_1,\ldots,a_k}p_1^{a_1}\cdot\cdots\cdot p_k^{a_k}$$

כאשר הa_i בין 0 לn וסכומם הוא n, ההסתברות לכל אחד יחיד זה פשוט בינומי.

$$cov(X_i,X_j)=-np_ip_j$$

סיכום מדדים של התפלגויות בדידות:

	תוחלת	שונות	פונקציה יוצרת מומנטים
אחידה $U[n,m]$	$\frac{n+m}{2}$	$\frac{(m-n+1)^2-1}{12}$	$\frac{e^{nt}-e^{(m+1)t}}{(m-n+1)(1-e^t)}$
ברנולית $b(p)$	p	pq	$q+pe^t$
בינומית $Bin(n,p)$	np	npq	$(q+pe^t)^n$
פואסונית $Poi(\lambda)$	λ	λ	$e^{\lambda(e^t-1)}$
גיאומטרית $G(p)$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^t}{1-qe^t}$
בינומית שלילית NB(r, p)	$\frac{rp}{q}$	$\frac{rp}{q^2}$	$\left(\frac{p}{1-qe^t}\right)^r$
מולטינומית $M(n;p_1,\ldots,p_k)$	np_i	np_iq_i	

משתנים מקריים רציפים

הגדרה פורמלית של מ"מ רציף: עבור מרחב הסתברות $(\Omega,\mathcal{F},P)$, משתנה מקרי עליו הוא פונקציה $X:\Omega\to\mathbb{R}$ כך שלכל קבוצה $B\in\mathcal{B}$, עבור $\mathcal{B}$ הסיגמא-אלגברה של בורל (כל האיחודים, המשלימים, והחיתוכים של קטעים ממשיים), מתקיים $(X\in B)\in\mathcal{F}$. נקרא למ"מ רציף אם הוא מקבל ערכים על קבוצה שאינה בת-מנייה של ערכים.

פונקציות הצטברות ופונקציות צפיפות: פונקציה F היא פונקציית אם הגבול שלה במינוס אינסוף הוא 0, באינסוף הוא 1, היא רציפה מימין,

ועולה. פונקציית צפיפות f היא פונקציה שהאינטגרל הכללי עליה הוא 1, והיא אי-שלילית.

פונקציות הצטברות וצפיפות מהוות דרכים לתיאור של מ"מ ע"י תיאור ההסתברויות שלהם. אפשר לחשוב על פונקציית הצטברות $F(t)$ כמתארת את ההסתברות של מ"מ להיות קטן-שווה t, ואפשר לחשוב על פונקציית צפיפות f כמתארת את ההסתברות של מ"מ להיות בקטעים מוכללים ממשיים ע"י האינטגרלים עליה.

אפשר גם להוכיח שכל מ"מ יכול להיות מתואר ע"י הצטברות, ושכל הצטברות מגדירה מ"מ, וכן שכל מ"מ עם הצטברות גזירה (נקרא גם גזיר) יכול להיות מתואר ע"י פונקציית צפיפות, ושכל פונקציית צפיפות מתארת מ"מ שכזה.

פונקציות ההצטברות והצפיפות של מ"מ: עבור מ"מ X, פונקציית ההצטברות שלו מוגדרת בתור $F_X(t) = P(X \leq t)$. אם הפונקציה גזירה, מגדירים את פונקציית הצפיפות של X, $f_X(t)$, בתור הנגזרת שלה. בכיוון ההפוך, אם נתונה לנו פונקציית הצפיפות $f_X(t)$, מתקיים הקשר:

$$\int_{-\infty}^{t} f_X(x)dx = F_X(t)$$

יש לציין שבקורס אנחנו מניחים אלא אם צויין אחרת שלכל מ"מ רציף יש פונקציית צפיפות. כמו כן:

1. לכל מ"מ שקיימת לו פונקציית צפיפות (יותר נכון שההצטברות שלו גזירה) ההסתברות שלו להיות בנקודה מסוימת היא 0.
2. לעיתים מתייחסים לצפיפות של מ"מ בתור ההתפלגות שלו.

צפיפות והצטברות משותפת, וצפיפויות שוליות של זוגות מ"מ: עבור זוג מ"מ רציפים X, Y, פונקציית ההצטברות המשותפת שלהם מוגדרת בתור:

$$F_{X,Y}(a,b) = P(X \leq a, Y \leq b)$$

פונקציית הצפיפות המשותפת שלהם מוגדרת בתור הפונקציה היחידה $f_{X,Y}(a,b)$ שהיא אי-שלילית, בעלת אינטגרל כללי של 1, ומקיימת:

$$\int_a^b \int_c^d f_{X,Y}(x,y)dxdy = P(c \leq X \leq d, a \leq Y \leq b)$$

הצפיפות השולית של כל אחד מהמשתנים הן פונקציית הצפיפות כפי שהן נובעות מפונקציית הצפיפות, כלומר:

$$f_X(b) := \int_{-\infty}^{\infty} f_{X,Y}(b,y)dy$$

$$f_Y(a) := \int_{-\infty}^{\infty} f_{X,Y}(x,a)dx$$

כדאי לציין שלעיתים גם נקבל מ"מ ע"י פונקציית הצטברות או צפיפות משותפת שלהם.

התנייה במאורעות ובמ"מ במקרה הרציף:

1. עבור מ"מ X רציף, המ"מ המותנה $X \mid A$ מוגדר בתור משתנה עם פונקציית ההצטברות $F_{X \mid A}(t) = P(X \leq t \mid A)$, ופונקציית הצפיפות שלו מוגדרת בתור $f_{X \mid A}(x) = F'_{X \mid A}(x)$.
2. עבור מ"מ רציף נוסף Y, כאשר רוצים להתנות במאורע $Y = y$, פונקציית הצפיפות של $X \mid Y = y$ מוגדרת בתור:

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

פונקציית ההצטברות מוגדרת באופן דומה:

$$F_{X \mid Y=y}(x) = \int_{-\infty}^{x} f_{X \mid Y=y}(t)dt$$

3. המ"מ המותנה $X \mid Y$ מקבל את הצפיפות וההצטברות של $X \mid Y = y$ לכל ערך y של Y.
4. למרות שעדיף להימנע מכך, אפשר לחשב הסתברות מותנית של מאורע A בהינתן המאורע $X = x$ למרות שזה מאורע עם הסתברות אפס. במקרה זה נבצע:

$$P(A \mid X = x) = \frac{f_{X \mid A}(x)P(A)}{f_X(x)}$$

תלות ואי-תלות בין מ"מ רציפים: עבור מ"מ X, Y, הם מוגדרים להיות ב"ת אם $f_{X,Y}(x,y) = f_X(x)f_Y(y)$. לא הגדרנו בקורס הגדרה פורמלית עבור אי-תלות במשותף של כמה מ"מ רציפים, על אף שקיימת אחת.

המקרה המעורב: עבור מ"מ בדיד D, ומ"מ רציף C, המ"מ $C \mid D$ הוא מ"מ שמקבל את ההתפלגות של $C \mid D = d$ לכל ערך d של D.

המ"מ המותנה $D \mid C$ מוגדר להיות בעל הצפיפות הבאה:

$$f_{D \mid C}(d \mid c) = \frac{f_{C \mid D}(c \mid d)P(D = d)}{f_C(c)}$$

והצטברות שהיא האינטגרל ממינוס אינסוף עד d של הצפיפות:

$$F_{D \mid C}(d \mid c) = \int_{-\infty}^{d} f_{D \mid C}(t \mid c)dt$$

אפשר לקבל מנוסחאות אלה את הצפיפות וההצטברות המשותפת של המשתנים ע"י:

$$\begin{aligned} f_{C,D}(c,d) &= f_{D \mid C}(d,c)f_C(c) \\ f_{C,D}(c,d) &= f_{C \mid D}(c,d)P(D = d) \\ F_{C,D}(c,d) &= \int_{-\infty}^{c}\int_{-\infty}^{d} f_{C,D}(x,y)dydx \end{aligned}$$

תוחלת של מ"מ רציף: עבור מ"מ X, Y רציפים, ופונקציה g, מגדירים:

$$\mathbf{E}(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t_1,t_2)f_{X,Y}(t_1,t_2)dt_1dt_2$$

1. ההגדרה עבור מ"מ מותנים מתקבלת ע"י החלפת הצפיפות בצפיפות מותנית.
2. כל התכונות של התוחלת מיתרגמות בצורה זהה עבור מ"מ רציפים.

מדדים אחרים של מ"מ רציפים: השונות, השונות המשותפת, ומקדם המתאם של מ"מ רציפים מוגדרות בצורה זהה למ"מ בדידים, כנ"ל לגבי הגדרות אחרות שמעורבות איתם ותכונותיהם.

נוסחאות בסגנון נוסחת ההסתברות השלמה במקרה הרציף :

1. מתקיים $P(A) = \int_{-\infty}^{\infty} P(A \mid X = x)f_X(x)dx$
2. מתקיים $f_X(t) = \int_{-\infty}^{\infty} f_{X \mid Y}(t \mid y)f_Y(y)dy$. זו מסקנה מההגדרה של צפיפות מותנית במ"מ אחר.
3. מתקיים $F_X(t) = \int_{-\infty}^{t}\int_{-\infty}^{\infty} f_{X \mid Y}(t \mid y)f_Y(y)dy$
4. חוק התוחלת החוזרת וחוק פירוק השונות מאוד שימושיים בנוסחאות בסגנון זה. למשל:

$$\mathbf{E}(X)=\int_{-\infty}^{\infty}\mathbf{E}(X\mid Y=y)f_Y(y)dy$$

(מסקנה מחוק התוחלת החוזרת)

התפלגויות רציפות

התפלגות אחידה רציפה: התפלגות של מ״מ רציף שמקבל ערכים בקטע סגור בצפיפות שווה, כלומר לקטעים בגדלים שווים בקטע אותה הסתברות למשתנה להיות בהם. משתנה מקרי רציף מתפלג כך ($X \sim U(a,b)$) אם ורק אם:

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

בכך ההצטברות של מ״מ המתפלג אחידה בקטע היא:

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

התפלגות מעריכית: התפלגות של מ״מ רציף המתאר את הזמן שלוקח את שקורה מאורע, כאשר משך הזמן מתאפיין בחוסר זיכרון: העובדה שעבר זמן לא רלוונטית לכמה זמן עוד יעבור. משתנה מקרי רציף מתפלג כך ($X \sim Exp(\lambda)$) אם ורק אם:

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

בכך ההצטברות של מ״מ המתפלג מעריכית בקטע היא:

$$F_X(t)=1-e^{-\lambda t}$$

תכונות מיוחדות של ההתפלגות מעריכית:

1. חוסר זיכרון: מ״מ רציף X המקבל ערכים אי-שליליים נקרא חסר זיכרון אם ורק אם לכל $a,b\ge 0$ מתקיים $P(X\ge a+b\mid X>a)=P(X\ge b)$.

   מ״מ רציף המקבל מתפלג מעריכית עם פרמטר כלשהו אם ורק אם הוא חסר זיכרון.

2. קשר בין התפלגות מעריכית להתפלגות פואסונית: עבור סדרת מ״מ ב״ת בזוגות $X_1,\dots,X_n \sim Exp(\lambda)$, נגדיר:

$$N_t=\max\left(\{k\in\mathbb{N}\mid X_1+\cdots+X_k\le t\}\right)$$

אזי מתקיים: $N_t \sim Poi(\lambda t)$. גם הכיוון ההפוך נכון (אם ההתפלגות פואסונית כולם מתפלגים אקספוננציאלית).

התפלגות נורמלית: התפלגות נפוצה מאוד בטבע המתארת מ״מ הבנוי מסכום או ממוצע של הרבה גורמים ב״ת. משתנה מקרי רציף מתפלג כך ($X \sim N(\mu,\sigma^2)$) אם ורק אם:

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

ההצטברות של מ״מ נורמלי אינה אלמנטרית.

תכונות מיוחדות של ההתפלגות הנורמלית:

1. אם $X \sim N(\mu,\sigma^2)$, אז:

$$\alpha X+\beta \sim N\left(\alpha\mu+\beta,(\alpha\sigma)^2\right)$$

2. אם $X \sim N(\mu,\sigma^2)$, אז: $\frac{X-\mu}{\sigma}\sim N(0,1)$

3. אם $X\sim N(0,1)$ אז $\sigma X+\mu \sim N(\mu,\sigma^2)$

4. אם X,Y ב״ת ומתפלגים נורמלית, גם הסכום שלהם מתפלג נורמלית.

5. אם שני מ״מ X,Y שווי התפלגות וב״ת, וגם $X+Y, X-Y$ הם בלתי-תלויים, אז X,Y מתפלגים נורמלית.

6. אם שני מ״מ X,Y שווי התפלגות וב״ת, עם תוחלת 0, ולכל זווית θ מתקיים ש $X\cos(\theta)+Y\sin(\theta)$ מתפלג כמוהם, אז בהכרח X,Y מתפלגים נורמלית.

התפלגות גאמא: חבר של התפלגות המעריכית, בעצם מתאר סכום של הפתלגויות מעריכיות, ומתאר כמה זמן עד שמשהו מופיע, נשים לב כי עבור מספרים שלמים: $\Gamma(n)=(n-1)!$

$$f_X(t)=\begin{cases}\frac{t^{n-1}}{\Gamma(n)}\cdot\lambda^n\cdot e^{-\frac{t}{\lambda}} & t\in(0,\infty)\\ 0 & O.W\end{cases}$$

ההצטברות לא אלמנטרית. הקשר בין מעריכית לגאמא (המ"מ ב"ת):

$$X_i\sim Exp(\lambda)\Rightarrow\sum_{i=1}^{n}X_i\sim\Gamma(n,\lambda)$$

סיכום מדדים של התפלגויות רציפות:

	תוחלת	שונות	פונקציה יוצרת מומנטים
אחידה $U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{t(b-a)}$
מעריכית $Exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{1}{\left(1-\frac{t}{\mu}\right)}$
נורמלית $N(\mu,\sigma^2)$	μ	σ^2	$e^{\mu t+\frac{\sigma^2t^2}{2}}$
גאמא $\boldsymbol{\Gamma(n,\lambda)}$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$\left(1-\frac{t}{\lambda}\right)^{-n}$

אי-שוויונים

אי-שוויון מרקוב: אם X מ״מ חיובי, ו $a\in\mathbb{R}$, אז:

$$P(X\ge a\mathbf{E}(X))\le\frac{1}{a}\;P(X\ge a)\le\frac{\mathbf{E}(X)}{a}$$

אי-שוויון צ׳בישב: אם X מ״מ, ו $a\in\mathbb{R}^+$, אז:

$$P\left(|X-\mathbf{E}(X)|\ge a\sqrt{\mathbf{V}(X)}\right)\le\frac{1}{a^2}$$

$$P(|X-\mathbf{E}(X)|\ge a)\le\frac{\mathbf{V}(X)}{a^2}$$

מומנטים

מומנטים ומומנטים מרכזיים: המומנט ה-k של מ"מ X מוגדר בתור $\mathbf{E}\left(X^k\right)$. המומנט המרכזי ה-k מוגדר בתור $\mathbf{E}\left(\left(X-\mathbf{E}\left(X\right)\right)^k\right)$.

גבנוניות וצידוד:

1. הצידוד של מ"מ X היא המומנט המרכזי השלישי שלו. משמעותו היא שככל שהוא גדול יותר כך המ"מ מתפלג פחות סימטרית סביב התוחלת. צידוד חיובי אומר שסביר יותר ליפול מתחת לתוחלת, צידוד שלילי אומר שסביר יותר ליפול מעל התוחלת, וצידוד 0 אומר שהסיכויים שווים.

2. הגבנוניות של מ"מ X היא המומנט המרכזי הרביעי שלו. היא מודדת את טבע הסטיות מהתוחלת: גבנוניות נמוכה פירושה סטיות גדולות אבל נדירות, וגבנוניות גבוהה פירושה סטיות קטנות אך נפוצות. הגבנוניות אי-שלילית.

פונקציה יוצרת מומנטים: עבור מ"מ X הפונקציה יוצרת המומנטים שלו מוגדרת בתור $M_X\left(t\right)=\mathbf{E}\left(e^{tX}\right)$.

הנגזרת מסדר n ב 0 של הפונקציה שווה למומנט ה n של המ"מ, במובן שבו היא קיימת אמ"מ הוא קיים והם שווים במקרה זה.

אם הפונקציה מוגדרת בסביבה של 0, היא גזירה אינסוף פעמים ב 0. במקרה זה הפונקציה יוצרת המומנטים שווה לטור הטיילור שלה, כלומר:

$$M_X\left(t\right)=\sum_{k=0}^{\infty}\frac{\mathbf{E}\left(X^k\right)}{k!}t^k$$

יהי מ"מ X,Y אזי מתקיים כי פונקצית הצפיפות של הסכום שלהם היא **הקונבולוציה** של הפונקציות צפיפות (הקונבולוציה חילופית):

$$f_{X+Y}(t)=f_X*f_Y(t)\coloneqq\int_{-\infty}^{\infty}f_X(x)f_Y(t-x)dx$$

משפטים חשובים וקירובים

התכנסויות של סדרות מ"מ: עבור סדרת מ"מ X_n:

1. הסדרה מתכנסת בהסתברות ל c, כלומר $X_n\xrightarrow{P}c$, אם ורק אם: $\forall\epsilon>0:P\left(\left|X_n-a\right|<\epsilon\right)\to1$.

2. הסדרה מתכנסת כמעט בוודאות ל c, כלומר $X_n\xrightarrow{\text{a.s.}}c$, אם ורק אם: $P\left(\lim_{n\to\infty}X_n=c\right)=1$. התכנסות כמעט בוודאות גוררת התכנסות בהסתברות.

$$\lim_{n\to\infty}X_n=a=\bigcap_{m\in\mathbb{N}}\bigcup_{N\in\mathbb{N}}\bigcap_{n>N\in\mathbb{N}}\left|X_n-a\right|<\frac{1}{m}$$

3. הסדרה מתכנסת בהתפלגות למ"מ X, כלומר $X_n\xrightarrow{D}X$, אם ורק אם: $\forall a:P\left(X_n\le a\right)\to P\left(X\le a\right)$.

חוקי המספרים הגדולים: עבור סדרת מ"מ X_n בלתי מתואמים (בזוגות כמובן) בעלי תוחלת μ ושונות σ^2, אם נגדיר $\overline{X}_n=\sum_{i=1}^{n}X_i/n$ אזי:

1. החוק החלש של המספרים הגדולים: $\overline{X}_n$ מתכנסת בהסתברות ל μ.

2. החוק החזק של המספרים הגדולים: אותה טענה רק עם התכנסות כמעט בוודאות.

משפט הגבול המרכזי: עבור סדרת מ"מ X_n שווי התפלגות, בלתי תלויים משותף, בעלי תוחלת μ ושונות σ^2, אם נגדיר $\overline{X}_n=\sum_{i=1}^{n}X_i/n$ אזי:

$$\left(\frac{\overline{X}_n-\mu}{\sigma}\right)\sqrt{n}\xrightarrow{D}N\left(0,1\right)$$

קירובים נורמליים של מ"מ: ממשפט הגבול המרכזי נובע שעבור מ"מ $X_1,\ldots,X_n$ שהם שווי התפלגות, בלתי תלויים במשותף, בעלי תוחלת μ ושונות σ^2, מתקיים:

$$\sum_{i=1}^{n}X_i\approx N\left(\mu n,\sigma^2n\right)\qquad\qquad\frac{\sum_{i=1}^{n}X_i}{n}\approx N\left(\mu,\frac{\sigma^2}{n}\right)$$

תיקון רציפות: כאשר מקרבים מ"מ בדיד X ע"י מ"מ רציף $\tilde{X}$, ההסתברות $P\left(X=a\right)$ מוחשבת ע"י המ"מ הרציף ע"י ההסתברות:

$$P\left(\left|\tilde{X}-a\right|\le\frac{1}{2}\right)$$

שרשראות מרקוב

שרשרת מרקוב: שרשרת מרקוב היא סדרת מ"מ $\left\{X_n\right\}_{n=0}^{\infty}$ כך שכל משתנה בה בעל תומך סופי, יכול להיות תלוי רק במשתנים שבאו לפניו, ובהינתן מ"מ שבא לפניו, הוא ב"ת בכל מ"מ שבא לפני המ"מ שערכו ידוע.

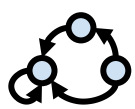
מטריצות מרקוב ווקטורי התפלגות: וקטור התפלגות הוא וקטור ממשי שסכום רכיביו הוא 1 והוא אי-שלילי. מטריצת מרקוב היא מטריצה $A\in\mathbb{R}^{n\times n}$ בה כל עמודה היא וקטור התפלגות.

מטריצת מרקוב מתארת שרשרת מרקוב במובן שע"י קביעת X_0 לפי וקטור התפלגות ב $\mathbb{R}^n$, והגדרת $X_n\sim C_{X_{n-1}}\left(A\right)$, מקבלים שרשרת מרקוב.

בשרשרת זו מתקיים שאם $X_n\sim u$, אז:

$$X_{n+k}\sim A^ku\qquad\qquad P\left(X_{n+k}=a\mid X_n=b\right)=\left[A^k\right]_{a,b}$$

תיאור מטריצות מרקוב בצורה ויזואלית: אפשר לייצג מטריצות מרקוב בצורה ויזואלית ע"י גרפים המתארים את ההסתברות לעבור מכל מצב אחד לאחר. למשל:



$$\begin{bmatrix}0&0&0.2\\1&1&0.8\\0&0&0\end{bmatrix}$$

ההתפלגות הסטציונרית: עבור מטריצת מרקוב $A\in\mathbb{R}^{n\times n}$, ההתפלגות הסטציונרית שלה מוגדרת בתור הוקטור העצמי של 1 עבור המטריצה כאשר הוא מנורמל להיות וקטור התפלגות (אם קיים ואם יחיד). ההתפלגות מתארת את ההתנהגות של מ"מ בשרשרת בשלב רחוק בעתיד.

שאלות בלי תלות בזמן: בשביל לפתור שאלות בתהליכי מרקוב לגבי מאורעות או מ"מ שלא מתייחסים לזמן ספציפי בתהליך, בדרך כלל ננסה לפתור את השאלות האלה מכל מצב בתהליך (ע"י הסתברות, תוחלת, או שונות שלמה) לפתור מערכת משוואות לינארית, ובכך לקבל את התשובה.

הרחבת זיכרון: שאלות מסוימות בתהליכי מרקוב יתייחסו למידע מכמה משתנים בתהליך שאין גישה אליו ישירות ממנו. לשם כך נבנה תהליך מרקוב חדש שהמצבים בו מתארים את המצב הכי עדכני בתהליך מרקוב המקורי שרלוונטי לשאלה שעונים עליה. מכאן נפתור כרגיל בתהליך החדש.

טורי טיילור שימושיים: מתקיים לכל ממשי x:

$$e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}$$

$$\cos\left(x\right)=\sum_{n=0}^{\infty}\frac{\left(-1\right)^n}{\left(2n\right)!}x^{2n}\qquad\qquad\sin\left(x\right)=\sum_{n=0}^{\infty}\frac{\left(-1\right)^n}{\left(2n+1\right)!}x^{2n+1}$$