7 משתנה מקרי חד מימדי בדיד 20-7

7.1 הגדרה. (משתנה מקרי בדיד) משתנה מקרי (חד מימדי) בדיד X (מ"מ) הוא פונקציה על המרחב מדגם המתאימה ערך מספרי יחיד לכל תוצאה במרחב הסתברות,

$$X : \Omega \to \mathbb{R} \ .$$

הפונקציה משתנה בהתאם לתוצאות הניסוי, בעוד תוצאת הניסוי היא סטוכסטית (אקראית) ולכן השינוי הוא מקרי. בשלב הראשון אנחנו נעסוק רק במשתנים מקריים בדידים וחד-מימדים. זאת אומרת, משתנים מקריים שמקבלים ערכים בדידים (בניגוד למשתנים מקריים רציפים שיכולים לקבל רצף של ערכים), ובנוסף כל משתנה מחזיר מספר בודד $X(\omega) \in \mathbb{R}$ כאשר $\omega \in \Omega$ ו- Ω הינו המרחב המדגם.

7.2 דוגמא. מטילים 2 קוביות. נסמן ב- X את סכום התוצאות שהתקבלו. כמו כן

$$\begin{aligned}
&\omega = \{(1,1)\}, && X(\omega) = 2, \\
&\omega = \{(2,1),(1,2)\}, && X(\omega) = 3, \\
&\omega = \{(1,3),(2,2),(3,1)\}, && X(\omega) = 4, \\
&\vdots \\
&\omega = \{(6,6)\}, && X(\omega) = 12.
\end{aligned}$$

זו דוגמה של מ"מ כאשר X הוא המשתנה המקרי.

7.3 דוגמא. מטילים 3 קוביות הוגנות. נגדיר משתנה מקרי X להיות התוצאה המקסימלית מבין ההטלות. כמו כן

$$\begin{aligned}
&\omega = \{(1,1,1)\}, && X(\omega) = 1, \\
&\omega = \{(2,1,1),(1,2,1),(1,1,2)\}, && X(\omega) = 2, \\
&\vdots \\
&\omega = \{(3,2,1),(2,3,1),(1,2,3),\dots\}, && X(\omega) = 3, \\
&\vdots \\
&\omega = \{(6,6,6)\}, && X(\omega) = 6.
\end{aligned}$$

זו דוגמה של מ"מ כאשר X הוא המשתנה המקרי.

7.4 הגדרה. (תומך של מ"מ בדיד) התומך של משתנה מקרי בדיד X זו קבוצת המספרים ש-X יכול להיות שווה בהסתברות חיובית, קרי

$$\text{Supp}(X) = \{k \in \mathbb{R} : P(X = k) > 0\} \ .$$

7.5 הגדרה. (פונקצית הסתברות / פונקצית התפלגות) הפונקצית ההתפלגות של משתנה מקרי בדיד (מ"מ) X מוגדרת להיות הפונקציה $f_X(k)$ המקבלת המ"מ X ומחזירה את ההסתברות כי למ"מ X יש ערך k:

$$f_X(k) = P(X = k) \ ,$$

עם התכונות

1. $\sum_{k \in X} f_X(k) = 1$.
2. $f_X(k) \geq 0 \quad \forall k$.

זאת אומרת, ההתפלגות של X בנקודה k מציינת את ההסתברות של המאורע

$$\{X = k\} \ .$$

למעשה, פונקצית התפלגות ופונקצית הסתברות הן די דומות מבחינת התכונות, אבל כדאי לא להתבלבל בינהן. פ ונקצית הסתברות מוגדרת על מרחבי מדגם ולכן הקלט של פונקצית ההסתברות יכול להיות מאוד מורכב. מנגד, התפלגויות מקבלות מספרים ומחזירות מספרים.

7.6 דוגמא. נטיל קוביה הוגנת ונניח כי X מחזיר את תוצאת ההטלה. אזי

$$\text{Supp}(X) = \{1,2,3,4,5,6\}$$

וההתפלגות של X היא

$$f_X(k) = \begin{cases} \frac{1}{6} & k \in \text{supp}(X) \\ 0 & \text{אחרת} \end{cases}$$

במילים, $f_X(k) = \frac{1}{6}$ לכל k בתומך, ו- 0 אחרת. חשוב לזכור שהתפלגות היא פונקציה ולכן יש צורך לתאר אותה בכל נקודה. משמע, זה לא מספיק לציין את ערכי הפונקציה רק על התומך, אלא בכל נקודה $k \in \mathbb{R}$.

7.7 דוגמא. אלון הוא סוחר ממולח בשוק ההון. הוא משקיע סכום של $\$1000$ במכשיר פיננסי אשר בכל שנה מכפיל את הכסף בסיכוי $\frac{2}{3}$ או מפסיד חצי מההשקעה בסיכוי $\frac{1}{3}$. השנים הן בלתי תלויות. נגדיר את X להיות שווי ההשקעה של אלון לאחר שנתיים. התומך של X כעת הוא

$$\text{supp}(X) = \{250, 1000, 4000\} \ ,$$

על כן, ההתפלגות של X היא

$$f_X(1000) = \frac{4}{9}, \quad f_X(4000) = \frac{4}{9}, \quad f_X(250) = \frac{1}{9}$$

ו - 0 אחרת.

7.8 דוגמא. מטילים 2 קוביות. נסמן ב- X את סכום התוצאות שהתקבלו. מצאו את ההתפלגות של X.

פיתרון. נתחיל מהתומך. התוצאות האפשרויות בסכום של 2 קוביות הן המספרים הטבעיים מ- 2 עד 12. בכדי לחשב את ההסתברות לכל ערך נשתמש במרחב מדגם סימטרי: גודל מרחב המדגם הוא $6^2 = 36$, ככמות האפשרויות לבחור תוצאה לקוביה הראשונה מתוך 6 ותוצאה לקוביה השניה מתוך 6. נחשב את ההסתברות של כל תוצאה בנפרד:

$$\begin{aligned}
P(X=2) &= \frac{|\{(1,1)\}|}{6^2} &&= \frac{1}{36}, \\
P(X=3) &= \frac{|\{(2,1),(1,2)\}|}{6^2} &&= \frac{2}{36}, \\
P(X=4) &= \frac{|\{(1,3),(2,2),(3,1)\}|}{6^2} &&= \frac{3}{36}, \\
&\vdots \\
P(X=7) &= \frac{|\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}|}{6^2} &&= \frac{6}{36}, \\
P(X=8) &= \frac{|\{(2,6),(3,5),(4,4),(5,3),(6,2)\}|}{6^2} &&= \frac{5}{36}, \\
P(X=9) &= \frac{|\{(3,6),(4,5),(5,4),(6,3)\}|}{6^2} &&= \frac{4}{36}, \\
&\vdots \\
P(X=12) &= \frac{|\{(6,6)\}|}{6^2} &&= \frac{1}{36},
\end{aligned}$$

ובאופן דומה ממשיכים לחשב את האפשרויות לכל הערכים האפשריים בתומך. באופן כללי קל לראות שהנוסחא, כאשר p זו הסכום של ה 2 הקוביות, היא (עיין משוואה (**??**) לעייל)

$$P(X=p) = \frac{1}{36} \times \begin{cases} (p-1) & p < 8 \\ (p-1) - 2(p-7) & 8 \le p < 14 \end{cases}$$

■

7.9 הגדרה. (פונקציית התפלגות מצטברת) פונקצית התפלגות מצטברת מסומנת ע"י $F_X(x)$ של מ"מ בדיד X בעל פונקצית התפלגות $f_X(k)$ מוגדרת להיות

$$F_X(x) = \sum_{\substack{k\in X \\ k\leq x}} f_X(k) \ .$$

במילים אחרות, פונקצית התפלגות מצטברת $F_X(x)$ מוגדרת להיות הפונקציה שתחומה הוא הישר הממשי וטווחה הקטע $[0,1]$ ומקיימת

$$F_X(x) = P(X \leq x) \qquad \forall x \in X \ .$$

7.10 דוגמא. אלון הוא סוחר ממולח בשוק ההון. הוא משקיע סכום של $\$1000$ במכשיר פיננסי אשר בכל שנה מכפיל את הכסף בסיכוי $\frac{2}{3}$ או מפסיד חצי מההשקעה בסיכוי $\frac{1}{3}$. השנים הן בלתי תלויות. נגדיר את X להיות שווי ההשקעה של אלון לאחר שנתיים. מצאו את פונקצית ההתפלגות המצטברת של X וציירו אותה.

פיתרון. כפי שציינו בדוגמה 7.7, התומך של X כעת הוא

$$\text{supp}(X) = \{250, 1000, 4000\} \ .$$

על כן, פונקצית ההתפלגות המצטברת של X היא

$$F_X(k) = P(X \leq k) = \begin{cases} 0 & k < 250, \\ \frac{1}{9} & 250 \leq k < 1000, \\ \frac{5}{9} & 1000 \leq k < 4000, \\ 1 & 4000 \leq k. \end{cases}$$

■

תוחלת היא מדד אמצע, לכן היא מעניקה לנו אינדיקציה מאוד בסיסית בנוגע לערך הממוצע של משתנה מקרי נתון. התוחלת לוקחת את הערכים השונים שמקבל המשתנה המקרי ומשקללת אותם, בהתאם לסבירות של כל ערך, לכדי מספר בודד שמציין את הערך הממוצע של אותו המשתנה. קרי, בדומה למדד המחירים לצרכן אשר משקלל את עלות סל צריכה בהתאם לצריכה, התוחלת היא ממוצע משוקלל של ערכי המשתנה המקרי בהתאם לסבירות של הערכים. בפועל, החישוב של התוחלת הוא מיידי: לוקחים את כל הערכים שמקבל המ"מ X, מכפילים כל ערך בהסתברותו וסוכמים.

7.11 הגדרה. (תוחלת של משתנה נקרי בדיד)

יהי X משתנה מקרי בדיד בעל פונקצית התפלגות $f_X(k)$. התוחלת (expectation) של X היא

$$\mu_X \equiv E[X] = \sum_{k\in X} k f_X(k) \ .$$

7.12 דוגמא. התוחלת של הטלת קוביה היא

$$E[X] = \sum_{k=1}^{6} k\frac{1}{6} = \frac{7}{2} \ .$$

7.13 דוגמא. בן הוא סוחר מתוחכם ולכן הוא מציע לאלון להפקיד אצלו 100 שקלים בתחילת השנה. בסיכוי של $\frac{2}{5}$ הסכום יוכפל, ובסיכוי $\frac{2}{5}$ הסכום יצטמצם בחצי בסוף השנה. מה תהיה תוחלת הכסף של אלון לאחר קבלת ההצעה? נסמן ב- X את שווי כסף של אלון בסוף השנה.

$$E[X] = 200.\frac{2}{5} + 50.\frac{3}{5} = 110.$$

זאת אומרת שהצעתו של בן מגלמת תוחלת רווח של 10%!

7.14 מסקנה. (לינאריות התוחלת)

1. עבור כל קבוע $a \in \mathbb{R}$ התוחלת של a היא

$$E[a] = a\ .$$

2. עבור כל משתנה מקרי X וקבוע $a \in \mathbb{R}$ התוחלת של aX היא

$$E[aX] = aE[X]\ .$$

3. עבור כל צמד משתנים מקריים X ו- Y התוחלת של $X+Y$ היא

$$E[X+Y] = E[X] + E[Y]\ .$$

4. כמובן שהתכונה הנ"ל ניתנת להכללה גם עבור מספר כללי של משתנים , שכן

$$E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]\ .$$

5. עבור המספרים קבועים a_i ומ"מ X_i,

$$E[\sum_{i=1}^{n} a_i X_i] = \sum_{i=1}^{n} a_i E[X_i]\ .$$

7.15 דוגמא. בחנות ירקות קטנה ישנם 3 תפוחים ו-4 אגסים. אלון קונה 2 פירות באקראי. בן המוכר מרוויח על כל תפוח 2 ש"ח ועל כל אגס 4 ש"ח. מצאו את תוחלת הרווח של בן.

פיתרון. נגדיר מ"מ Q עבור הרווח של בן, ומשתנה מקרי X כמספר התפוחים שאלון קנה. על כן,

$$Q = 2X + 4(2 - X) = 8 - 2X$$

$$P(X=1) = \frac{4}{7}, \qquad P(X=2) = \frac{3}{7}.\frac{2}{6} = \frac{1}{7}, \qquad P(X=0) = \frac{4}{7}.\frac{3}{6} = \frac{2}{7}\ .$$

מכאן נובע ש

$$E[X] = 0.\frac{2}{7} + 1.\frac{4}{7} + 2.\frac{1}{7} = \frac{6}{7}\ .$$

נשתמש בתכונת הליניאריות לחישוב התוחלת של

$$E[Q] = 8 - 2E[X] = 8 - \frac{12}{7} = \frac{44}{7}\ .$$

■

דוגמא. בוחרים קוד באורך 10 תווים, כאשר כל תו נבחר באקראי מתוך הספרות 9 – 0 ללא קשר לתווים האחרים שנבחרו. נסמן ב- X מ"מ הסופרים את מספר הרצפים 11 בקוד. חשבו את $E[X]$.

פיתרון. נגדיר מ"מ $X_i,\ i=1,\dots,9$ כאשר X_i מקבל את הערך 1 אם ורק אם במקום ה- i מתחיל רצף של 11. לכן

$$X=\sum_{i=1}^{9} X_i\ .$$

$$\begin{aligned}E[X]=&E\left[\sum_{i=1}^{9}X_i\right]\\=&\sum_{i=1}^{9}E\left[X_i\right]\\=&\sum_{i=1}^{9}\left[1.P(X_i=1)+0.P(X_i=0)\right]\\=&\sum_{i=1}^{9}P(X_i=1)\\=&\sum_{i=1}^{9}\frac{1}{10}.\frac{1}{10}\\=&\frac{9}{100}\ .\end{aligned}$$

■

7.16 הגדרה. (תוחלת פונקציה של משתנה מקרי)

יהי X משתנה מקרי בדיד בעל פונקצית התפלגות $P_X(x)$. התוחלת של המשתנה מקרי $g(X)$ הוא

$$E[g(X)]=\sum_{k\in X} g(x)f_X(k)\ .$$

7.17 דוגמא. ניקח משתנה מקרי X בעל התפלגות

$$P(X=k)=\begin{cases}\frac{1}{3} & k=-1,\\ \frac{1}{6} & k=0,\\ \frac{1}{2} & k=1.\end{cases}$$

חשבו את $E[X]$ ואת $E[X^4]$.

פיתרון. נשתמש בהגדרת התוחלת ונקבל

$$\begin{aligned}E[X]=&(-1).\frac{1}{3}+0.\frac{1}{6}+1.\frac{1}{2} &&=\frac{1}{6},\\ E[X^4]=&(-1)^4.\frac{1}{3}+0^4.\frac{1}{6}+1^4.\frac{1}{2} &&=\frac{5}{6}.\end{aligned}$$

■

7.18 דוגמא. נניח ש X הוא המספר של רכבים אשר עוברים דרך השתיפת רכב בין השעות $16:00$ ו- $17:00$. ל- X יש את ההתפלגות

k	4	5	6	7	8	9
$P(X=k)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

הפונקציה $g(X)=2X-1$ מצייג את הרווח ב \$ עבור X. מצאו את תוחלת הרווח בין השעות $16:00$ ו- $17:00$.

פיתרון.

$$\begin{aligned}
E[g(X)] =&E[2X-1]\\
=&\sum_{k=4}^{9}(2x-1)f_X(k)\\
=&7.\frac{1}{12}+9.\frac{1}{12}+11.\frac{1}{4}+13.\frac{1}{4}+15.\frac{1}{6}+17.\frac{1}{6}\\
=&\$12.67\ .
\end{aligned}$$

■

7.19 הגדרה. (שונות של משתנה מקרי) יהי X משתנה מקרי בדיד בעל פונקציית התפלגות $f_X(k)$ ותוחלת μ. השונות (variance) של X הוא

$$V(X)=E[(X-\mu)^2]=\sum_{k\in X}(k-\mu)^2 f_X(k)\ .$$

הגדרה. (סטיית התקן) סטיית התקן (standard deviation) של משתנה מקרי X מוגדרת להיות שורש השונות ומסומנת ב-

$$\sigma_X=\sqrt{V(X)}\ .$$

7.20 מסקנה. (קיצור דרך לשונות)

$$\begin{aligned}
V(X) =&E[(X-\mu_X)^2]\\
=&E[(X^2-2\mu_X X+\mu_X^2)]\\
=&E[X^2]-2\mu_X E[X]+E[\mu_X^2]\\
=&E[X^2]-2\mu_X\mu_X+\mu_X^2\\
=&E[X^2]-(E[X])^2\ .
\end{aligned}$$

7.21 דוגמא. בר היא בנקאית השקעות מתוחכמת המציעה ללקוחותיה את ההגרלה הבאה: בסיכוי $\frac{2}{5}$ תקבלו רווח של 10. בסיכוי $\frac{1}{5}$ לא תרוויחו ולא תפסידו דבר, ובסיכוי $\frac{3}{5}$ תפסידו 5. חשבו את התוחלת וסטיית התקן של X המייצג את הרווח מהשקעה אצל בר.

פיתרון. תוחלת הרווח היא חיובית שכן

$$\begin{aligned}
E[X] =&10.\frac{2}{5}+0.\frac{1}{5}-5.\frac{2}{5} &&=2,\\
(E[X])^2 =&100.\frac{2}{5}+0^2.\frac{1}{5}+25.\frac{2}{5} &&=50,\\
V(X) =&E[X^2]-(E[X])^2 &&=50-4=46,\\
\Rightarrow\quad \sigma_X =&\sqrt{46}\ .
\end{aligned}$$