פרק 4: משתנים מקריים, פונקציות התפלגות ותוחלת

1 הקדמה

בפרק 1 הגדרנו מרחב הסתברות, שסומן על-ידי השלשה (Ω, A, P). התחלנו עם ניסוי מקרי: קראנו לאוסף כל התוצאות האפשריות של הניסוי מרחב מדגם וסימנו אותו באות Ω. A ציין אוסף של תת-קבוצות, הנקראות מאורעות, של מרחב המדגם. ולבסוף פונקצית ההסתברות P היא פונקציה שהתחום שלה הוא A והטווח שלה הקטע $[0,1]$. המטרה שלנו הייתה ועדיין לייחס הסתברויות למאורעות. נשתמש במונח 'משתנה מקרי' שיוגדר עתה לתיאור מאורעות, ובמונח 'פונקצית התפלגות מצטברת' כדי לתת הסתברות למאורעות מסוימים המוגדרים במונחים של משתנים מקריים. בעצם שתי הקונספציות של משתנה מקרי ופונקצית התפלגות מצטברת יעזרו לנו במציאת הסתברות של מאורעות שהיא המטרה שלנו.

2 הגדרות

הגדרה **1** משתנה מקרי

נתון מרחב הסתברות (Ω, A, P). משתנה מקרי, המסומן על-ידי X או $X(\cdot)$ הוא פונקציה שהתחום שלה הוא מרחב המדגם Ω והטווח שלה הוא הישר הממשי.

הפונקציה X חייבת להיות כזאת שהקבוצה $A_r = \{\omega; X(\omega) \le r\}$ שייכת לשדה A לכל מספר ממשי r.

הגדרה אקוויולנטית

נתון מרחב הסתברות (Ω, A, P). פונקציה $X = X(\omega)$, המתאימה לכל נקודה ω במרחב המדגם Ω מספר ממשי $X(\omega)$, נקראת בשם משתנה מקרי חד-ממדי.

<u>הערות:</u>

1. ההגדרה האקוויולנטית היא החלק החשוב של הגדרת משתנה מקרי. העובדה שאנו דורשים גם שקבוצת כל ה- ω-ות עבורן $\{X(\omega) \le r\}$ תהיה מאורע (כלומר אלמנט של A לכל מספר ממשי r) היא לא ממש אילוץ למטרות שלנו מאחר והכוונה שלנו להשתמש במשתנה מקרי רק לצורך תיאור מאורעות. לעיתים רחוקות נתייחס למשתנה המקרי בעצמו, בעצם העניין שלנו הוא במאורעות המוגדרים במונחים של משתנה מקרי.
2. נשים לב שבהגדרת משתנה מקרי לא משתמשים בפונקצית ההסתברות P של מרחב ההסתברות (Ω, A, P).
3. המונח "משתנה מקרי" הוא מונח לא מוצלח שזכה להשתרש, מכיוון שמדובר בעצם על פונקציה ולא על משתנה.
4. בהגדרה שלנו סימנו משתנה מקרי על-ידי X או $X(\cdot)$, למרות ש $X(\cdot)$ הוא סימון שלם יותר, המדגיש שמשתנה מקרי הוא פונקציה, נשתמש בדרך כלל בסימון המקוצר X .

5. משתנים מקריים יסומנו באותיות לטיניות גדולות ערך של המשתנה המקרי יסומן באותיות לטיניות קטנות וכו'.

<u>דוגמא **1**:</u>

נניח הניסוי הוא זריקת מטבע. יהי X משתנה מקרי הסופר את מספר העצים. $\Omega = \{$פלי,עץ$\}$, ומתקיים $X($עץ$) = 1$, ו- $X($פלי$) = 0$. כלומר, המשתנה המקרי X התאים מספר ממשי לכל תוצאה של הניסוי. מכיוון שאנו קוראים ל-X משתנה מקרי עלינו להראות מתמטית שהוא מקיים את ההגדרה כלומר, עלינו להראות ש $\{\omega; X(\omega) \le r\}$ שייכת ל $\boldsymbol{A}$ לכל מספר ממשי r.

ב A ארבע תת-קבוצות: $A = \{\Phi, \{$עץ$\}, \{$פלי$\}, \Omega\}$. עכשיו:

עבור $r < 0$ $\{\omega; X(\omega) \le r\} = \Phi$

עבור $0 \le r < 1$ $\{\omega; X(\omega) \le r\} = \{$פלי$\}$

עבור $r \ge 1$ $\{\omega; X(\omega) \le r\} = \Omega = \{$פלי,עץ$\}$

כלומר לכל r הקבוצה $\{\omega; X(\omega) \le r\}$ שייכת ל $\boldsymbol{A}$ ולכן $X(\cdot)$ הוא משתנה מקרי.

<u>דוגמא **2**:</u>

נניח הניסוי הוא הטלת זוג קוביות. $\Omega = \{(i, j)\,; i, j = 1, \ldots, 6\}$. כלומר ב-Ω 36 נקודות מדגם. ניתן להגדיר מספר משתנים מקריים, לדוגמא - אם נגדיר:

X = סכום התוצאות על הקוביות. אזי אם $\omega = (i, j)$, $X(\omega) = i + j$.

Y = הערך המוחלט של ההפרש בין התוצאות. אזי אם $\omega = (i, j)$, $Y(\omega) = |i - j|$.

ניתן להראות ש X ו- Y הם משתנים מקריים. המ"מ X מקבל את הערכים 2, 3 12 והמ"מ Y מקבל את הערכים 0, 1 5 .

בשתי הדוגמאות שלמעלה תארנו את המשתנה המקרי במונחים של הניסוי המקרי ולא על-ידי תיאור של הצורה הפונקציונלית שלהם. בדרך כלל כך נתאר משתנים מקריים.

הגדרה **2** פונקצית התפלגות מצטברת

פונקצית התפלגות מצטברת של משתנה מקרי X , מסומנת על-ידי $F_X(\cdot)$,מוגדרת להיות פונקציה שהתחום שלה הוא הישר הממשי והטווח שלה הוא הקטע [0 , 1] והמקיימת $F_X(x) = P(X \le x) = P[\{\omega : X(\omega) \le x\}]$ לכל מספר ממשי x.

פונקצית ההתפלגות המצטברת מוגדרת באופן יחיד לכל משתנה מקרי X. אם הפונקציה ידועה ניתן להשתמש בה לחישוב הסתברויות של מאורעות המוגדרים במונחים של המשתנה

המקרי המתאים לה. (נשים לב שבהגדרה הזו אנו משתמשים בדרישה ש $\{\omega;\ X(\omega)\le r\}$
שייכת ל $\boldsymbol{A}$ לכל מספר ממשי r שהופיעה בהגדרת המשתנה המקרי X).

<u>דוגמא **3**:</u>

בניסוי זריקת מטבע. נניח כי המטבע הוגנת. יהי X משתנה מקרי הסופר את מספר העצים.
אזי,

$$F_X(x)=\begin{cases}0 & x<0\\ \frac{1}{2} & 0\le x<1\\ 1 & 1\le x\ .\end{cases}$$

<u>דוגמא **4**:</u>

בניסוי הטלת זוג הקוביות. יהי Y הערך המוחלט של ההפרש בין התוצאות. פונקצית ההתפלגות המצטברת של Y נתונה על-ידי: (נתאר גם את הפונקציה בגרף)

$$F_Y(y)=\begin{cases}0 & y<0\\ \frac{6}{36} & 0\le y<1\\ \frac{16}{36} & 1\le y<2\\ \frac{24}{36} & 2\le y<3\\ \frac{30}{36} & 3\le y<4\\ \frac{34}{36} & 4\le y<5\\ 1 & 5\le y\end{cases}$$

תכונות פונקצית ההתפלגות המצטברת :

1. $F_X(\infty)=\lim_{x\to\infty}F_X(x)=1$, $F_X(-\infty)=\lim_{x\to-\infty}F_X(x)=0$,
2. $F_X(\cdot)$ היא פונקציה מונוטונית לא יורדת, כלומר $F_X(a)\le F_X(b)$ עבור $a<b$.
3. $F_X(\cdot)$ היא פונקציה רציפה מימין, כלומר $\lim_{0<h\to0}F_X(x+h)=F_X(x)$.

נוכיח רק את תכונה מס' 2. נשים לב שהמאורע

$$\{\omega;\ X(\omega)\le b\}=\{X\le b\}=\{X\le a\}\cup\{a<X\le b\}$$

וכן ש $\{X\le a\}\cap\{a<X\le b\}=\phi$. ולכן

$$F_X(b)=P(X\le b)=P(X\le a)+P(a<X\le b)\ge P(X\le a)=F_X(a)$$

תכונה 3. ש $F_X(\cdot)$ רציפה מימין, נובעת מההגדרה שלנו את $F_X(x)$ שהיא $P(X \le x)$.

הגדרה **3** פונקצית התפלגות מצטברת

כל פונקציה $F(\cdot)$ שהתחום שלה הוא הישר הממשי והטווח שלה הוא הקטע $[0,1]$ והמקיימת את 3 התכונות שלמעלה, מוגדרת להיות פונקצית התפלגות מצטברת.

הגדרה זו מאפשרת לנו להשתמש במונח "פונקצית התפלגות מצטברת" מבלי להזכיר משתנה מקרי.

3. פונקציות הסתברות וצפיפות

הגדרנו משתנה מקרי ופונקצית התפלגות של משתנה מקרי. פונקצית ההתפלגות המצטברת מתארת את התפלגות הערכים של המשתנה המקרי. עבור שני סוגים של משתנים מקריים ניתן לתאר באופן קל יותר את התפלגות הערכים שלהם באמצעות פונקצית ההסתברות או פונקצית צפיפות ההסתברות. שני הסוגים האלה הם משתנה מקרי בדיד ומשתנה מקרי רציף שאותם נגדיר להלן.

3.1 משתנים מקריים בדידים

הגדרה **4** משתנה מקרי בדיד

משתנה מקרי X נקרא משתנה מקרי בדיד אם הטווח של X הוא בר-מניה. אם X הוא משתנה מקרי בדיד אז פונקצית ההתפלגות המצטברת שלו מוגדרת בדידה.

כשאנו אומרים שהטווח של X הוא בר-מניה הכוונה שקיימת סדרה x_1, x_2, x_3 סופית או אינסופית של מספרים ממשיים שהמ"מ X יכול לקבל. אם X משתנה מקרי בדיד עם ערכים בדידים x_n x_1, x_2, x_3 אזי $\Omega = \bigcup_n \{\omega; X(\omega) = x_n\} = \bigcup_n \{X = x_n\}$ ומכיוון שעבור כל $i \ne j$ מתקיים $\{X = x_i\} \cap \{X = x_j\} = \phi$ נובע ש - $1 = P(\Omega) = \sum_n P(X = x_n)$ לפי האקסיומה של ההסתברות.

הגדרה **5** פונקצית הסתברות של משתנה מקרי בדיד

אם X משתנה מקרי בדיד המקבל ערכים בדידים x_n x_1, x_2, x_3 , אזי הפונקציה המסומנת על-ידי $P_X(x)$ והמוגדרת על-ידי

$$P_X(x) = \begin{cases} P(X = x_j) & x = x_j \quad j = 1, 2, \ldots, n, \ldots \\ 0 & x \ne x_j \end{cases}$$

היא פונקצית ההסתברות של X .

$P_X(\cdot)$ היא פונקציה שהתחום שלה הוא הישר הממשי והטווח שלה הוא הקטע [0,1] .
הערכים של המשתנה המקרי נקראים לעיתים נקודות מסה, והפונקציה $P_X(x_j)$ מציינת את המסה המחולקת לכל אחת מנקודות המסה.

משפט **1**

יהי X משתנה מקרי בדיד. ניתן לקבל את $F_X(\cdot)$ מתוך $P_X(\cdot)$ ולהפך.

הוכחה

יהיו x_1, x_2 נקודות מסה של המ"מ X. נניח שהפונקציה $P_X(x_j)$ נתונה אזי $F_X(x) = \sum_{\{j; x_j \le x\}} P_X(x_j)$.

ולהפך, נניח ש $F_X(\cdot)$ נתונה אזי $P_X(x_j) = F_X(x_j) - \lim_{0<h\to 0} F_X(x_j - h)$ כלומר ניתן לקבל את $P_X(x_j)$ לכל נקודת מסה x_j , אבל $P_X(x) = 0$ לכל $x \neq x_j$, $j = 1, 2, \ldots$ כך ש $P_X(x)$ מוגדרת לכל x ממשי.

<u>דוגמא **5**</u>

כדי להדגים את משמעותו של משפט 1, נניח ניסוי של זריקת קוביה. יהי X התוצאה על הקוביה:

$$F_X(x) = \begin{cases} 0 & x < 1 \\ 1/6 & 1 \le x < 2 \\ 2/6 & 2 \le x < 3 \\ 3/6 & 3 \le x < 4 \\ 4/6 & 4 \le x < 5 \\ 5/6 & 5 \le x < 6 \\ 1 & 6 \le x \end{cases} \qquad P_X(x) = \begin{cases} \frac{1}{6} & x = 1, 2, \ldots, 6 \\ 0 & \text{אחרת} \end{cases}$$

לפי משפט 1, עבור פונקצית הסתברות $P_X(\cdot)$ נתונה ניתן לקבל את $F_X(x)$ לכל x .

לדוגמא, עבור $x = 2.5$:

$$F_X(2.5) = \sum_{\{j; x_j \le 2.5\}} P_X(x_j) = P_X(1) + P_X(2) = \frac{2}{6}$$

ואם $F_X(\cdot)$ נתונה, ניתן לקבל את $P_X(x)$ לכל x. לדוגמא עבור $x = 3$,

$$P_X(3) = F_X(3) - \lim_{0<h\to 0} F_X(3-h) = \frac{3}{6} - \frac{2}{6} = \frac{1}{6}$$

הצורה הגרפית של פונקצית ההתפלגות המצטברת של משתנה מקרי בדיד היא פונקצית מדרגות. לפונקצית ההתפלגות המצטברת של משתנה מקרי בדיד יש ׳מדרגה׳ בכל נקודת מסה. כלומר, בנקודת המסה x_j , ל- $F_X(\cdot)$ יש מדרגה בגודל $P_X(x_j)$, ו- $F_X(\cdot)$ שטוחה בין נקודות מסה.

הגדרה **6** פונקצית הסתברות בדידה

כל פונקציה $P(\cdot)$ שהתחום שלה הוא הישר הממשי והטווח שלה הוא הקטע $[0,1]$ מוגדרת להיות פונקצית הסתברות של משתנה מקרי בדיד אם עבור קבוצה בת מניה של ערכים

$x_1, x_2, x_3 \qquad x_n$

1. $P(x_j) > 0$ עבור $j = 1, 2, \ldots.$
2. $P(x) = 0$ עבור $x \neq x_j$, $j = 1, 2, \ldots.$,
3. $\sum P(x_j) = 1$, כאשר הסכום על כל הנקודות $x_1, x_2, x_3 \qquad x_n$.

הגדרה זו מאפשרת לנו להשתמש במונח ״פונקצית הסתברות״ מבלי להתייחס למשתנה מקרי מסוים.

<u>דוגמא **6**</u>

נניח הניסוי הוא הטלת זוג קוביות. נגדיר:

X = סכום התוצאות על הקוביות.

Y = הערך המוחלט של ההפרש בין התוצאות.

נקודות המסה של X הן $2, 3 \qquad 12$. פונקצית ההסתברות של X נתונה בטבלה הבאה:

x	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
$P_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

פונקצית ההסתברות של Y נתונה בטבלה הבאה:

y	**0**	**1**	**2**	**3**	**4**	**5**
$P_Y(y)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

(נתאר את פונקצית ההסתברות וההתפלגות המצטברת בגרף).