4. תוחלות ומומנטים

תוחלת היא עקרון חשוב בבעיות הקשורות למשתנים מקריים או התפלגויות. הפסקאות הבאות ידונו בהגדרות ותוצאות הנוגעות לתוחלות.

4.1 תוחלת

הגדרה **10** תוחלת

יהי X משתנה מקרי. התוחלת של X המסומנת ב $E(X)$ או μ_X מוגדרת על-ידי:

1. $E(X)=\sum_j x_j P_X(x_j)$ אם X משתנה מקרי בדיד עם נקודות מסה x_j x_1, x_2

2. $E(X)=\int_{-\infty}^{\infty} x f_X(x)dx$ אם X משתנה מקרי רציף עם פונקצית צפיפות $f_X(x)$

3. $E(X)=\int_{-\infty}^{\infty}[1-F_X(x)]dx-\int_{-\infty}^{0}F_X(x)dx$ עבור משתנה מקרי כלשהו.

בהגדרה 1: $E(X)$ מוגדרת כסכום של טור בתנאי שהטור מתכנס בהחלט, אחרת אנחנו אומרים שהתוחלת לא קיימת.

בהגדרה 2: $E(X)$ מוגדרת כאינטגרל בתנאי שהאינטגרל קיים, אחרת אנחנו אומרים שהתוחלת לא קיימת.

בהגדרה 3: אנחנו דורשים ששני האינטגרלים יהיו סופיים כדי שלמשתנה תהיה תוחלת.

נשים לב מה משמעות ההגדרה: ב- $\sum_j x_j P_X(x_j)$ הסכום הוא על ערכי המשתנה המקרי X מוכפלים בהסתברויות לקבל את הערכים האלו. כך $E(X)$ הוא הממוצע המשוקלל של הערכים שהמ"מ המקרי X מקבל, כשהמשקלים הם ההסתברויות לקבל את הערכים האלו. ערכים שיש להם הסתברות גבוהה יקבלו משקל גדול יותר. אותו דבר נכון גם להגדרה 2. שם הערך x מוכפל בהסתברות , בערך, שהמ"מ X יקבל את הערך x, כלומר $f_X(x)dx$ ואז אינטגרל על כל הערכים.

הערות

1) בהגדרת התוחלת של משתנה מקרי, השתמשנו רק בפונקצית הסתברות או הצפיפות או ההתפלגות המצטברת. כלומר הגדרנו את התוחלת של הפונקציות האלה מבלי להתייחס למשתנים המקריים. לתוחלת שהוגדרה נקרא לכן 'התוחלת של ההתפלגות המצטברת' או 'התוחלת של הצפיפות'. אנו יכולים לכן לדבר על התוחלת של ההתפלגות כמו גם על התוחלת של המשתנה המקרי.

2) $E(X)$ היא מרכז הכובד של ההתפלגות, נקודת שיווי המשקל של ההתפלגות. לכן התוחלת של X היא מדד למיקום מרכזי של המשתנה המקרי X. מדדים אחרים למיקום או מיקום מרכזי נכיר בהמשך.

3) הגדרה 3 נכונה לכל המשתנים המקריים בעוד שהגדרת 1 מתאימה רק למשתנים מקריים בדידים ואילו הגדרה 2 רק למשתנים מקריים רציפים. יכולנו להסתפק לכן בהגדרה 3. הסיבה להוספת הגדרות 1 ו- 2 היא שהן יותר אינטואיטיביות וקל יותר ברוב המקרים לחשב באמצעותן את התוחלת. ניתן להראות, למרות שלא נעשה זאת, שהגדרה 1 נובעת מ- 3 במקרה הבדיד והגדרה 2 במקרה הרציף. נשתמש בהגדרה 3 במקרים שקל יותר לחשב בעזרתה את התוחלת.

<u>דוגמא **10**</u>

ניח הניסוי הוא הטלת זוג קוביות. נחשב את התוחלת של הממי"ם:

X = סכום התוצאות על הקוביות. $E(X) = 7$

Y = הערך המוחלט של ההפרש בין התוצאות. $E(Y) = \frac{70}{36}$

<u>דוגמא **11**</u>

יהי X משתנה מקרי רציף בעל פונקצית הצפיפות הבאה: $(\lambda > 0)$

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

אזי $E(X) = \frac{1}{\lambda}$. (נראה לפי הגדרה 2 ו- 3).

<u>דוגמא **12**</u>

יהי X משתנה מקרי בעל פונקצית הצפיפות הבאה:

$$f(x) = \begin{cases} \frac{1}{x^2} & x > 1 \\ 0 & \text{אחרת} \end{cases}$$

אזי $E(X) = \int_1^{\infty} x \cdot \frac{1}{x^2} dx = \int_1^{\infty} \frac{1}{x} dx = \lim_{b\to\infty} (\ln b) = \infty$, במקרה כזה נאמר ש $E(X)$ לא קיימת.

נוכל גם לומר כי התוחלת של X שווה לאינסוף מאחר שברור כי האינטגרל שמגדיר את התוחלת הוא אינסופי.

4.2 תוחלת של פונקציה של משתנה מקרי

יהי X משתנה מקרי ותהי $g(\cdot)$ פונקציה שהתחום והטווח שלה הוא הישר הממשי. התוחלת של $g(X)$ המסומנת ב $E[g(X)]$ מוגדרת על-ידי:

1. $E[g(X)] = \sum_j g(x_j)P_X(x_j)$ אם X משתנה מקרי בדיד עם נקודות מסה x_1, x_2 (בתנאי שהטור מתכנס בהחלט).

2. $E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$ אם X משתנה מקרי רציף עם פונקצית צפיפות $f_X(x)$ (בתנאי $\int_{-\infty}^{\infty} |g(x)|f_X(x)dx < \infty$).

התוחלת של פונקציה של משתנה מקרי מוגדרת רק במונחים של פונקצית ההסתברות או הצפיפות או של המ"מ X, ניתן לכן להגדיר אותה מבלי להתייחס למשתנה מקרי.

הערה: נשים לב שכאשר $g(x) = x$, אזי $E[g(X)] = E(X)$ התוחלת של X. נעבור למקרה פרטי חשוב נוסף של תוחלת של פונקציה של משתנה מקרי , השונות.

4.3 שונות

התוחלת של משתנה מקרי X מדדה מיקום מרכזי של ההתפלגות. השונות של משתנה מקרי X היא מדד לפיזור של ההתפלגות.

הגדרה **11** שונות

יהי X משתנה מקרי. נסמן את התוחלת של X על-ידי $\mu_X = E(X)$. "השונות של X" המסומנת ב- $V(X)$ או σ_X^2 מוגדרת על-ידי: $V(X) = E[(X-\mu_X)^2]$.

1. $V(X) = \sum_j (x_j - \mu_X)^2 P_X(x_j)$ אם X משתנה מקרי בדיד עם נקודות מסה x_j x_1, x_2 .

2. $V(X) = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)dx$ אם X משתנה מקרי רציף עם פונקצית צפיפות $f_X(x)$.

3. $V(X) = \int_{-\infty}^{\infty} 2x[1-F_X(x)+F_X(-x)]dx - \mu_X^2$ עבור משתנה מקרי כלשהו.

השונות מוגדרת רק אם הטור ב- 1 מתכנס או אם האינטגרלים ב- 2, 3 קיימים. שוב השונות של משתנה מקרי מוגדרת רק במונחים של פונקצית ההסתברות או הצפיפות או ההתפלגות המצטברת של המ"מ X וניתן לכן להגדיר אותה מבלי להתייחס למשתנה מקרי.

השונות היא מדד לפיזור. השונות היא בעצם הממוצע המשוקלל של המרחקים בריבוע מהתוחלת. אם הערכים שהמשתנה המקרי מקבל נוטים להיות רחוקים מהתוחלת השונות גדולה יותר. ברור מהגדרה 1 ו- 2 ונכון גם לגבי הגדרה 3 שהשונות היא לא שלילית.

הגדרה **12** סטית תקן

יהי X משתנה מקרי. "סטית התקן של X" מסומנת ב- σ_X ומוגדרת על-ידי $\sigma_X = +\sqrt{V(X)}$. (השורש החיובי של השונות).

סטית התקן של X כמו השונות היא מדד לפיזור של ערכי המשתנה המקרי. בשימושים רבים היא עדיפה על השונות מכיוון שהיא תימדד באותן יחידות מדידה כמו המשתנה המקרי.

<u>דוגמא 13</u>

נניח הניסוי הוא הטלת זוג קוביות. נחשב את השונות של המשתנה המקרי

X = סכום התוצאות על הקוביות.

$$V(X) = \sum_{x=2}^{12}(x-7)^2 P_X(x) = (2-7)^2 \cdot \frac{1}{36} + (3-7)^2 \cdot \frac{2}{36} + \cdot \quad \cdot \quad \cdot + (12-7)^2 \cdot \frac{1}{36} = \frac{210}{36}$$

<u>דוגמא 14</u>

יהי X משתנה מקרי רציף בעל פונקצית הצפיפות הבאה: $(\lambda > 0)$

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

השונות של X :

$$V(X) = \int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)dx = \int_{0}^{\infty}(x-\frac{1}{\lambda})^2 \lambda e^{-\lambda x}dx = \frac{1}{\lambda^2}$$

משפט **3** : תכונות התוחלת

1. $E[c] = c$ עבור c קבוע.
2. $E[cg(X)] = cE[g(X)]$ עבור c קבוע.
3. $E[c_1 g_1(X) + c_2 g_2(X)] = c_1 E[g_1(X)] + c_2 E[g_2(X)]$
4. $E[g_1(X)] \leq E[g_2(X)]$ אם $g_1(x) \leq g_2(x)$ לכל x ממשי. (הוכחה בכיתה).

משפט **4** :

אם X משתנה מקרי אזי $V(X) = E[(X-\mu_X)^2] = E[X^2] - \mu_X{}^2 = E[X^2] - (E[X])^2$.

בתנאי ש $E[X^2]$ קיימת.

משפט **5** : אי שוויון מרקוב

יהי X משתנה מקרי ותהי $g(\cdot)$ פונקציה לא שלילית שהטווח שלה הישר הממשי. אזי

לכל $k > 0$ $$P[g(X)] \ge k] \le \frac{E[g(X)]}{k}$$

בפרט: אם $g(x) = x$ נובע $P[X] \ge k] \le \frac{E[(X)]}{k}$. (הוכחה בכיתה)

אי שווין צ׳ביצ׳ב (תוצאה של אי-שוויון מרקוב)

יהי X משתנה מקרי בעל שונות סופית אזי : $$P[\left|X-\mu_X\right| \ge \varepsilon] \le \frac{V(X)}{\varepsilon^2}$$

אי שווין צ׳ביצ׳ב נותן חסם עליון להסתברות שהמשתנה המקרי X יקבל ערכים במרחק של יותר מ- ε מתוחלתו.

בפרט, יש לנו חסם עליון להסתברות למצוא את X במרחק של יותר מ- r סטיות תקן מתוחלתו:

לכל $r > 0$ $$P[\left|X-\mu_X\right| \ge r\sigma_X] = P[(X-\mu_X)^2 \ge r^2\sigma_X^2] \le \frac{1}{r^2}$$

הוכחה: ניקח $g(x) = (x-\mu_X)^2$ ו- $k = r^2\sigma^2$ ולפי משפט 5 הנ״ל מתקיים.

<u>הערה:</u> אם X משתנה מקרי בעל שונות סופית אזי $P[\left|X-\mu_X\right| < r\sigma_X] \ge 1-\frac{1}{r^2}$ שזה בדיוק שכתוב של המשוואה הקודמת. משתמשים באי שוויון צ׳ביצ׳ב בדרכים שונות. אנו נשתמש בו יותר מאוחר על מנת להוכיח את החוק של המספרים הגדולים. נשים לב שהמשוואה האחרונה אומרת: $P[\mu_X - r\sigma_X < X < \mu_X - r\sigma_X] \ge 1-\frac{1}{r^2}$ כלומר, ההסתברות ש-X ייפול במרחק של r יחידות של סטית תקן מתוחלתו גדולה/שווה ל- $1-\frac{1}{r^2}$. לדוגמא עבור r=2, $P[\mu_X - 2\sigma_X < X < \mu_X - 2\sigma_X] \ge \frac{3}{4}$ או לכל משתנה מקרי X בעל שונות סופית, לפחות שלושה רבעים של המסה של X נופלת בתוך שתי סטיות תקן מהתוחלת.

בדרך כלל, כדי לחשב הסתברות של מאורע המוגדר במונחים של משתנה מקרי X אנו נדרשים לדעת את ההתפלגות של X. אי-שוויון צ׳ביצ׳ב נותן לנו חסם, שאינו תלוי בהתפלגות של X, להסתברות של מאורעות מסוימים המתוארים במונחים של משתנה מקרי והתוחלת והשונות שלו.

4.4 מומנטים ופונקציה יוצרת מומנטים

הגדרה 13 מומנטים

יהי X משתנה מקרי. המומנט מסדר r של X מסומן ב- μ_r , ומוגדר על-ידי: $\mu_r = E[X^r]$

אם התוחלת קיימת. נשים לב ש $\mu_1 = E[X] = \mu_X$.

הגדרה 14 מומנטים מרכזיים

יהי X משתנה מקרי. המומנט המרכזי מסדר r של X סביב a מוגדר על-ידי $E[(X-a)^r]$.

ואם $a = \mu_X$,נקבל את המומנט המרכזי מסדר r של X סביב התוחלת, המסומן ב- μ_r' ,

ומוגדר על-ידי: $\mu_r' = E[(X-\mu_X)^r]$.

נשים לב ש $\mu_1' = E[(X-\mu_X)] = 0$ וכן ש $\mu_2' = E[(X-\mu_X)^2] = V(X)$, השונות של X. כמו כן אם פונקצית הצפיפות (ההסתברות) של X סימטרית סביב μ_X אזי כל המומנטים האי-זוגיים סביב μ_X שווים 0.

הגדרה 15 מומנט פקטוריאלי

יהי X משתנה מקרי. המומנט הפקטוריאלי מסדר r של X מוגדר כ :

(עבור r מספר שלם חיובי) $E[X(X-1)\cdot \cdot \cdot (X-r+1)]$.

עבור משתנים מקריים מסוימים (בדרך כלל בדידים) קל יותר לחשב מומנטים פקטוריאליים מאשר מומנטים. בכל אופן ניתן לקבל את המומנטים מהמומנטים הפקטוריאליים ולהפך. מכיוון שהמומנטים של ההתפלגות חשובים, היה שימושי אם יכולנו למצוא פונקציה שתייצג את כל המומנטים. פונקציה כזו נקראת פונקציה יוצרת מומנטים.

הגדרה **16** פונקציה יוצרת מומנטים

יהי X משתנה מקרי עם צפיפות (הסתברות) $f_X(\cdot)$ ($P_X(\cdot)$) . התוחלת של e^{tX} מוגדרת להיות "פונקציה יוצרת מומנטים" של X אם התוחלת קיימת לכל ערך t באינטרוול מסוים $h > 0\ ;\ \ -h < t < h$. פונקציה יוצרת מומנטים מסומנת ב- $M_X(t)$ ושווה ל-

עבור X רציף : $M_X(t) = E[e^{tX}] = \int e^{tx} f_X(x)dx$

עבור X בדיד : $M_X(t) = E[e^{tX}] = \sum_x e^{tX} P_X(x)$

נשים לב שפונקציה יוצרת מומנטים מוגדרת במונחים של פונקצית צפיפות (או הסתברות), ומכיוון שפונקצית צפיפות הוגדרה מבלי להתייחס למשתנים מקריים (ראה הגדרות 6 ו- 9) ניתן לדון בפונקציה יוצרת מומנטים מבלי להתייחס למשתנים מקריים.

אם קיימת פונקציה יוצרת מומנטים, אזי $M_X(t)$ גזירה ורציפה בסביבה מסוימת של t. ואם נגזור r פעמים את $M_X(t)$ לפי t נקבל:

$$\frac{d^r}{dt^r} M_X(t) = \int_{-\infty}^{\infty} x^r e^{tx} f_X(x)dx$$

ועבור $t \to 0$ נמצא כי $\frac{d^r}{dt^r} M_X(0) = E[X^r] = \mu_r$. כשצד שמאל של המשוואה פירושו הנגזרת מסדר r של $M_X(t)$ כאשר $t \to 0$. כך, ניתן לקבל את המומנטים של ההתפלגות על ידי גזירה של הפונקציה יוצרת המומנטים ומכאן שמה.

אם בהגדרה 16 נפתח את הטור e^{tX} נקבל את הפיתוח לטור של $M_X(t)$ במונחים של המומנטים של $P_X(\cdot)$:

$$M_X(t) = E[e^{tX}] = E[1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdot\cdot\cdot\ ]$$

$$= 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \frac{t^3}{3!}E[X^3] + \cdot\cdot\cdot$$

$$= \sum_{j=0}^{\infty} \frac{t^j}{j!} E[X^j]$$

ושוב אנו רואים כי אפשר לקבל את $E[X^r] = \mu_r$ מתוך $M_X(t)$. μ_r הוא המקדם של $\frac{t^r}{r!}$