# תרגול כיתה 8 – מבוא להסתברות וסטטיסטיקה
## מ"מ דו מימדי, התפלגות משותפת, שונות משותפת שונות מותנה ומקדם המתאם

**נוסחאות**

(1) $E(\sum x) = \sum E(x)$

(2a) $V(X+Y) = V(X) + V(Y)$   $[X, Y \text{ independent}]$

(2b) $V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$   $[X, Y \text{ dependent}]$

(3) $Cov(X,Y) = E(XY) - E(X)E(Y)$

(4) $\rho(X,Y) = \dfrac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$

(5) $\mathrm{E}(X \mid Y=y) = \sum_x x\, \mathrm{P}(X=x \mid Y=y) = \sum_x x \dfrac{\mathrm{P}(X=x, Y=y)}{\mathrm{P}(Y=y)}$

**שאלה 1**

בכד חמישה כדורים הממוספרים מ-1 עד 5: כדורים מס' 1 ו-2 אדומים, כדורים מס' 3 ו-4 ירוקים וכדור מס' 5 לבן. מוציאים מהכד שני כדורים, באופן מקרי וללא החזרה. יהיו:

X – מספר הכדורים הירוקים שיצאו.

Y – מספר הכדורים שיצאו ועליהם מספרים זוגיים.

א. מצא את פונקצית ההתפלגות המשותפת של X ו- Y ואת ההתפלגויות השוליות.

ב. מצא את ההסתברות ש- (Y = X) ואת ההסתברות ש- (Y < X)

ג. חשב את השונות המשותפת $Cov(X,Y)$. האם X ו-Y מתואמים? הערך את חוזק המתאם.

ד. חשב את מקדם המתאם בין X ו- Y.

ה. ידוע כי יצא לפחות כדור אחד אדום. מה ההסתברות שיצא בדיוק כדור אחד שעליו מספר זוגי?

ו. חשב $P(X \mid Y=1)$ ואת $E(X \mid Y=1)$ .

*פתרון:*

א. טבלת ההתפלגות המשותפת של X ו-Y (השוליות בשולי הטבלה, עמודה ושורה קיצוניות)

| Y \ X | 0 | 1 | 2 | P(y=k) |
|---|---|---|---|---|
| 0 | 1/10 | 2/10 | 0 | 3/10 |
| 1 | 2/10 | 3/10 | 1/10 | 6/10 |
| 2 | 0 | 1/10 | 0 | 1/10 |
| P(x=k) | 3/10 | 6/10 | 1/10 | 1 |

דרך החישוב בתאי הטבלה דלעיל:

תא (X=0,Y=0): ההסתברות להוציא כדור ראשון לא ירוק ולא זוגי (2/5) וכדור שני לא ירוק ולא זוגי (1/4).

$$P\left(X=0,Y=0\right)=\frac{2}{5}\cdot\frac{1}{4}=\frac{1}{10}$$

תא (X=1,Y=0):

[ (כדור ראשון לא זוגי וירוק וכדור שני לא זוגי ולא ירוק) = (1/5)(2/4)

+ (כדור ראשון לא זוגי ולא ירוק וכדור שני לא זוגי וירוק= (1/4)(2/5) ]

$$P\left(X=1,Y=0\right)=\frac{1}{5}\cdot\frac{2}{4}+\frac{2}{5}\cdot\frac{1}{4}=\frac{2}{10}$$

תא (X=2,Y=0): [אי אפשר לשלוף 2 כדורים ירוקים וגם שלא יהיו בידנו כדורים עם מס' זוגי]

$$P\left(X=2,Y=0\right)=0$$

תא (X=1,Y=1):

[(ירוק+לא ירוק ולא זוגי) + (ירוק ולא זוגי+לא ירוק וזוגי) + (לא ירוק וזוגי+ירוק ולא זוגי)
+ (לא ירוק ולא זוגי+ירוק וזוגי)]

$$P\left(X=1,Y=1\right)=\frac{1}{5}\cdot\frac{2}{4}+\frac{1}{5}\cdot\frac{1}{4}+\frac{1}{5}\cdot\frac{1}{4}+\frac{2}{5}\cdot\frac{1}{4}=\frac{6}{20}=\frac{3}{10}$$

וכך הלאה.

<= ההתפלגויות השוליות [P(X=k), P(Y=k)] מופיעות בשולי הטבלה (עמודה ושורה קיצוניות)

ב. נעזר בטבלה שבנינו למצוא את ההסתברויות:

$$\begin{aligned}P\left(X=Y\right)&=P\left(X=0,Y=0\right)+P\left(X=1,Y=1\right)+P\left(X=2,Y=2\right)\\&=1/10+3/10+0=4/10\\P\left(X>Y\right)&=P\left(X=1,Y=0\right)+P\left(X=2,Y=0\right)+P\left(X=2,Y=1\right)\\&=2/10+0+1/10=3/10\end{aligned}$$

ג. לחישוב השונות המשותפת (קווריאנס) נחשב תחילה את התוחלות $E\left[X\right],E\left[Y\right]$ לפי ההגדרה:

$$E\left[X\right]=\sum_{k=0}^{2}kP\left(X=k\right)=0+1\cdot 6/10+2\cdot 1/10=4/5$$

בגלל הסימטריות בשאלה ניתן מיד לומר שזו גם התוחלת של Y.

השונות המשותפת: לפי ההגדרה $Cov\left(X,Y\right)=E\left[XY\right]-E\left[X\right]\cdot E\left[X\right]$

נחשב את המרכיב $E\left[XY\right]$ :

$$E[XY]=\sum_{i,j=0}^{2,2} xyP(X=i,Y=j)=1\cdot 1\cdot 3/10+1\cdot 2\cdot 1/10+2\cdot 1\cdot 1/10=7/10$$

נציב הכל בנוסחה:

$$Cov(X,Y)=E[XY]-E[X]\cdot E[X]=7/10-(4/5)^2=0.06$$

לא קיבלנו $Cov(X,Y)=0$, לכן X ו-Y מתואמים.

> השונות המשותפת תלויה ביחידות, לכן איננו יכולים לדעת עד כמה המתאם אכן קטן או גדול רק לפיה.

ד. מטעמי סימטריה $\sigma(X)=\sigma(Y)$. נחשב תחילה: $\sigma(X)=\sqrt{V(X)}$.

$$E\left[X^2\right]=\sum_{k=0}^{2} y^2P(Y=k)=0+1^2\cdot 6/10+2^2\cdot 1/10=1$$

$$\sigma(X)=\sqrt{V(X)}=\sqrt{E\left[X^2\right]-E^2[X]}=\sqrt{1-0.8^2}=0.6$$

ונציב בנוסחה למקדם המתאם:

$\rho(X,Y)=\dfrac{Cov(X,Y)}{\sigma(X)\sigma(Y)}=\dfrac{0.06}{0.6\cdot 0.6}=1/6$ (רואים שהמתאם לא גדול אבל גם לא קרוב ל- 0).

ה. כעת נסמן ע"י X את מספר הכדורים האדומים, נשים לב שמטעמי סימטריה (מרחב ההסתברות סימטרי ויש 2 כדורים אדומים ו-2 ירוקים) התפלגותם זהה להתפלגות הירוקים (לכן אפשר להשתמש בטבלה).

$$\begin{aligned}&P(Y=1\mid X\geq 1)\\&=\frac{P(Y=1\cap X=1)+P(Y=1\cap X=2)}{\left[P(X=1)+P(X=2)\right]}\\&=\frac{0.3+0.1}{0.6+0.1}=0.571\end{aligned}$$

ו. ההסתברות המותנה:

$$P(X\mid Y=1)=\frac{P(X,Y=1)}{P(Y=1)}=\frac{P(X,Y=1)}{0.6}=\begin{cases}0.2/0.6=1/3 & ,x=0\\0.3/0.6=1/2 & ,x=1\\0.1/0.6=1/6 & ,x=2\\0 & ,else\end{cases}$$

התוחלת המותנה:

$$E(X\mid Y=1)=\sum_{x=0}^{2} x\,\mathrm{P}(X=x\mid Y=1)=0+1\cdot\frac{1}{2}+2\cdot\frac{1}{6}=\frac{5}{6}$$

<u>שאלה 2</u>

שתי קוביות משחק נזרקות. נגדיר מ"מ: X – התוצאה הגבוהה ביותר ו-Y – מספר הקוביות עם תוצאה זוגית.

א. בנה את טבלת ההתפלגות של $(X,Y)$.

ב. חשב את ההסתברויות הבאות: $P(X \geq 4, Y=1)$, $P(X \geq 4, Y=2)$, $P(X+Y \leq 6)$, $P(2Y < X)$

ג. חשב את ההתפלגויות השוליות של X ו-Y.

ד. האם X ו-Y בלתי תלויים? נמק.

ה. מצא את התפלגות של $Z = X+Y$.

ו. מהי התוחלת והשונות של Z.

*<u>פתרון</u>*:

X- התוצאה הגבוהה ביותר, Y- מספר הקוביות עם תוצאה זוגית.

א). טבלת ההתפלגות המשותפת:

| X / Y | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 0 | 1/36 | 0 | 3/36 | 0 | 5/36 | 0 |
| 1 | 0 | 2/36 | 2/36 | 4/36 | 4/36 | 6/36 |
| 2 | 0 | 1/36 | 0 | 3/36 | 0 | 5/36 |

ב). חישוב ההסתברויות:

$$P(X \geq 4, Y=1) = \frac{1}{36}(4+4+6) = \frac{14}{36} = \frac{7}{18}$$

$$P(X \geq 4, Y=2) = \frac{1}{36}(3+0+5) = \frac{8}{36} = \frac{2}{9}$$

$$P(X+Y \leq 6) = 1 - P(X+Y > 6) = 1 - \frac{1}{36}(6+0+5) = \frac{25}{36}$$

$$P(2Y < X) = \frac{1}{36}(\underbrace{1+0+3+0+5+0}_{y=0} + \underbrace{2+4+4+6}_{y=1} + \underbrace{0+5}_{y=2}) = \frac{30}{36}$$

ג) ההתפלגויות השוליות הן:

| i | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| P(X=i) | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 |

| i | 0 | 1 | 2 |
|---|---|---|---|
| P(Y=j) | 9/36 | 18/36 | 9/36 |

ד) X,Y אינם בלתי תלויים מכיוון שלמשל:

$$P(X=3,Y=0)=\frac{3}{36}=0.08333\neq P(X=3)P(Y=0)=\frac{5}{36}\cdot\frac{9}{36}=0.0347$$

ה) עבור Z=X+Y נחשב מהטבלה המשותפת לעיל:

| s | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| P(Z=s) | 1/36 | 0 | 5/36 | 3/36 | 9/36 | 7/36 | 6/36 | 5/36 |

נמחק תאים בהם ההסתברות 0, ונקבל את טבלת התפלגות Z הסופית:

| s | 1 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| P(Z=s) | 1/36 | 5/36 | 3/36 | 9/36 | 7/36 | 6/36 | 5/36 |

ו) התוחלת והשונות של Z:

התוחלת: $E(Z)=\sum_{s=1}^{8} sP(Z=s)$

$$E(Z)=\frac{1}{36}\left(1\cdot1+3\cdot5+4\cdot3+5\cdot9+6\cdot7+7\cdot6+8\cdot5\right)=5.472$$

השונות: $V(Z)=E\left(Z^2\right)-E^2\left(Z\right)$

$$E(Z^2)=\frac{1}{36}\left(1^2\cdot1+3^2\cdot5+4^2\cdot3+5^2\cdot9+6^2\cdot7+7^2\cdot6+8^2\cdot5\right)=32.917$$

$$V(Z)=E\left(Z^2\right)-E^2\left(Z\right)=32.917-5.472^2=2.974$$

במקרה זה חישבנו מפורשות את הטבלה של Z, לכן זו הדרך הקצרה ביותר.

=> דרך אחרת, העושה שימוש ב-X ו-Y ותכונות השונות והתוחלת, ללא שימוש בטבלה של Z:

התוחלת –

$$E(Z)=E(X+Y)=E(X)+E(Y)=4\frac{17}{36}+1=5.472$$

השונות –

$$V(Z)=V(X+Y)=V(X)+V(Y)+2Cov(X,Y)$$

את ה-$Cov$ נחשב ע"י:

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$
$$E(XY)=\sum_{j=0}^{2}\sum_{i=1}^{6}jiP(X=i,Y=j)$$

התוצאה המספרית פחות חשובה, חשוב לדעת שזו הדרך למקרה הצורך (כמו במבחן...).

## תוחלת ושונות מותנה, תוחלת ושונות נשנית

### נוסחאות

תוחלת מותנה:

$$\mathrm{E}(X\mid Y=y)=\sum_x x\,\mathrm{P}(X=x\mid Y=y)=\sum_x x\,\frac{\mathrm{P}(X=x,Y=y)}{\mathrm{P}(Y=y)}$$

תוחלת נשנית:

$$\mathrm{E}\left(X\right)=\mathrm{E}\left(\mathrm{E}\left(X\mid Y\right)\right)=\sum_y \mathrm{E}(X\mid Y=y)\cdot\mathrm{P}(Y=y)$$

שונות נשנית:

$$\mathrm{Var}\left(Y\right)=\mathrm{E}\left(\mathrm{Var}\left(Y\mid X\right)\right)+\mathrm{Var}\left(\mathrm{E}\left(Y\mid X\right)\right)$$

<u>שאלה 3</u>

יואב וחנן משחקים בזריקת כדורים לסל לסירוגין. יואב קולע בסיכוי 0.7 וחנן בסיכוי 0.4. חנן מתחיל במשחק. המשחק מסתיים כאשר חנן קולע ראשון.

א. מהי תוחלת מספר הקליעות של יואב בזמן זה?

ב. מהי השונות?

*<u>פתרון:</u>*

א). נסמן ב- $X$ את הקליעה הראשונה של חנן. $X\sim G\left(0.4\right)$.

נסמן ב- $Y$ את מספר הקליעות של יואב עד הקליעה הראשונה של חנן. ניתן לראות שמס' הקליעות של יואב מתפלג בצורה מותנית באופן: $Y\mid X\sim Bin\left(X-1,0.7\right)$. לכן לכל $k\in\mathbb{N}$ מתקיים-

$$Y\mid X=k\sim Bin(k-1,0.7).$$

לפיכך נוכל השתמש בתוחלת של התפלגות בינומית, $E(Y\mid X=k)=np=(k-1)\cdot 0.7$.

ומכאן מתקיים גם ש- $E(Y\mid X)=0.7(X-1)$.

כעת נעזר בכלל התוחלת הנשנית:

$$E(Y)=E\left(E\left(Y\mid X\right)\right)=E(0.7X-0.7)=0.7E(X)-0.7$$
$$=\frac{0.7}{0.4}-0.7=\boxed{1.05}$$

ב). ראינו שמס' הקליעות של יואב מתפלג בצורה מותנית באופן: $Y\mid X\sim Bin\left(X-1,0.7\right)$.

לכן גם- $Y\mid X=k\sim Bin(k-1,0.7)$ לכל $k\in\mathbb{N}$.

נוסחת השונות של התפלגות בינומית תיתן, $V(Y\mid X=k)=np\left(1-p\right)=0.7\cdot 0.3(k-1)$

לפיכך מתקיים גם ש- $V(Y\mid X)=0.7\cdot 0.3(X-1)=0.21(X-1)$

כעת נפעיל את כלל השונות הנשנית:

$$V(Y)=V\left[E(Y\mid X)\right]+E\left[V(Y\mid X)\right]$$
$$=V(0.7X-0.7)+E(0.21X-0.21)$$
$$=0.7^2\cdot\frac{0.6}{0.4^2}+\frac{0.21}{0.4}-0.21=\boxed{2.1525}$$

<u>שאלה 4</u>

במשחק מזל מטילים קובייה הוגנת פעם אחת. תהי $N$ תוצאת ההטלה. כעת מטילים מטבע $N$ פעמים. מצידו האחד של המטבע חקוק "1" ומצידו האחר "0". הסתברויות הטלת המטבע $C$ הן:

$$P(C=0)=1-p,\ \ P(C=1)=p,\ \ p\in(0,1).$$

הרווח בכל הטלת המטבע היא תוצאת ההטלה. מהי תוחלת סה"כ הרווח במשחק ומהי שונותו?

*<u>פתרון:</u>*

נגדיר מ"מ $N\sim U[1,6]$ – התפלגות הטלת הקובייה.

ניתן לראות שהרווח מסדרת הטלות המטבע הוא התפלגות המותנה (בהינתן $N$), נגדיר אותו באופן הבא-

$$X\mid N\sim Bin(N,p)$$

לחישוב התוחלת נעזר בנוסחת התוחלת הנשנית (דרך התניה)-

$$E(X)=E\big(E(X\mid N)\big)=\underbrace{E(N\cdot p)}_{X|N\sim Bin(N,p)}=p\underbrace{E(N)}_{N\sim U[1,6]}=p\cdot\underbrace{3.5}_{\frac{(1+6)}{2}}$$

לחישוב השונות נעזר בכלל השונות הנשנית-

$$\begin{aligned}V(X)&=E\big(V(X\mid N)\big)+V\big(E(X\mid N)\big)\\&=E\big(N\cdot p(1-p)\big)+V(N\cdot p)\\&=p(1-p)\cdot E(N)+p^2\cdot V(N)\\&=p(1-p)\cdot\frac{7}{2}+p^2\cdot\frac{6^2-1}{12}=\boxed{\frac{7}{2}p-\frac{7}{12}p^2}\end{aligned}$$

<u>שאלה 5</u>

מטילים קובייה הוגנת $n$ פעמים. נגדיר את המ"מ הבאים: $X$ – מס' הפעמים שהתקבל 6,
$Y$ – מס' הפעמים שהתקבל 5. חשב את $Cov(X,Y)$.

*<u>פתרון:</u>*

חישוב ישיר הוא ארוך. ננקוט בדרך אחרת. מהנתונים ידוע כי

$$X\sim Bin(n,1/6),\ \ Y\sim Bin(n,1/6)$$

נרצה להשתמש בנוסחת הקווריאנס (הנכונה תמיד ולא משנה מה הקשר בין המשתנים),

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

מהחיבוריות של ההתפלגות הבינומית נקבל- $X+Y\sim Bin(n,2/6)$.

לכן נוכל לבודד את המבוקש בשאלה-

$$\tfrac{1}{2}\left[V\left(X+Y\right)-V\left(X\right)-V\left(Y\right)\right]=Cov\left(X,Y\right)$$

השונות של מ"מ בינומי נתונה ע"י $Z\sim Bin\left(n,p\right)\Rightarrow V\left(Z\right)=np\left(1-p\right)$

לכן- $Cov\left(X,Y\right)=\frac{1}{2}\left[n\cdot\frac{2}{6}\cdot\frac{4}{6}-2\cdot n\cdot\frac{1}{6}\cdot\frac{5}{6}\right]=\boxed{\frac{-n}{36}}$

<u>שאלה 6</u>

מטילים מטבע הוגן $n$ פעמים $\Omega=\left\{H,T\right\}^n$. נגדיר: $X$ – מס' הפעמים שהתקבל "H" מתוך $n$,

$Y$ – מס' הפעמים שהתקבל "T" מתוך $n$. חשב את המתאם $\rho\left(X,Y\right)$.

*<u>פתרון:</u>*

נשתמש בהגדרת מקדם המתאם $\rho\left(X,Y\right)=\frac{Cov\left(X,Y\right)}{\sqrt{V\left(X\right)V\left(Y\right)}}$

נחשב את $Cov\left(X,Y\right)$ במונה ע"י שימוש בתכונות הקווריאנס ונעזר בכך ש- $X+Y=n$ ,

$$Cov\left(X,Y\right)=Cov\left(X,N-X\right)=Cov\left(X,N\right)+Cov\left(X,-X\right)$$

המעבר האחרון לעיל נובע מתכונת הקווריאנס:

$$Cov(aX,cW+dV)=acCov(X,W)+ad\,Cov(X,V)$$

נפתח את אגף ימין ונשתמש בכך שתוחלת של קבוע היא הקבוע,

$$=\left[E\left(X\cdot n\right)-E\left(X\right)E\left(n\right)\right]-V\left(X\right)$$
$$=nE\left(X\right)-nE\left(X\right)-V\left(X\right)=-V\left(X\right)$$

כעת, השונות של $Y$ היא- $V\left(Y\right)=V\left(n-X\right)=\left(-1\right)^2V\left(X\right)=V\left(X\right)$

נציב בנוסחת המתאם ונקבל, $\rho\left(X,Y\right)=\frac{-V\left(X\right)}{\sqrt{V\left(X\right)V\left(X\right)}}=\boxed{-1}$