הסתברות מ' | הרצאה 6 (00094412)

שם: איל

February 21, 2024

נושאי השיעור: השלמות להוכחות משפטים,

נושא ראשון - הוכחת משפטים מהרצאות קודמות

משפט 1. פונקציה של וקטור מקרי

- אם X הוא וקטור מקרי בדיד n מימדי ו-$g:\mathbb{R}^n\to\mathbb{R}^m$ ו-$\underline{Y}=g\left(X\right)$
- אז:

1. $\underline{Y}$וקטור מקרי בדיד
2. $P_{\underline{Y}}\left(y\right)=\sum_{\underline{x}:g\left(\underline{x}\right)=\underline{y}}P_{\underline{X}}\left(x\right)$

הוכחה.

1. הושאר כתרגיל.
2.

- נשים לב כי:

$$P_{\underline{Y}}\left(\underline{y}\right)=P\left(\underline{Y}=\underline{y}\right)$$

$$=P\left(g\left(\underline{X}\right)=\underline{y}\right)$$

$$=P\left(\underline{X}\in\overbrace{\left\{ \underline{x}\in\mathbb{R}^{n}:g\left(\underline{x}\right)=\underline{y}\right\} }^{=A}\right)$$

$$=P\left(\underline{X}\in\left\{ \underline{x}:g\left(\underline{x}\right)=\underline{y}\right\} \right)$$

$$= \sum_{\underline{x}:g(\underline{x})=\underline{y}} P_{\underline{X}}(x)$$

■

משפט 2. יהיו X,Y משתנים מקריים בדידים המוגדרים על אותו מרחב מדגם:

- אזי מתקיים:

1. $P_{Y|X}(y,x)=\begin{cases}\frac{P_{X,Y}(x,y)}{P_{X}(x)} & p_{X}(x)>0\\ 0 & otherwise\end{cases}$

2. נוסחת הכפל:

$$P_{X,Y}(x,y)=P_{X}(x)\cdot P_{Y|X}(y,z)$$

3. נוסחת בייס:

$$P_{X|Y}(x,y)=\begin{cases}P_{Y|X}(y,x)\cdot\frac{P_{X}(x)}{P_{Y}(y)} & p_{Y}(y)>0\\ 0 & otherwise\end{cases}$$

4. נוסחת ההסתברות השלמה:

$$P_{Y}(y)=\sum_{x}P_{X}(x)\cdot P_{Y|X}(y,x)$$

הוכחה.

1. **צ"ל:** $P_{Y|X}(y,x)=\begin{cases}\frac{P_{X,Y}(x,y)}{P_{X}(x)} & p_{X}(x)>0\\ 0 & otherwise\end{cases}$

- לפי הגדרה מתקיים:

$$P_{Y|X}=\begin{cases}P(\{Y=y\}\mid\{X=x\}) & P(\{X=x\})>0\\ 0 & otherwise\end{cases}$$

$$=\begin{cases}\frac{P(\{Y=y\}\cap\{X=x\})}{P(X=x)} & P(\{X=x\})>0\\ 0 & otherwise\end{cases}$$

$$=\begin{cases}\frac{P_{X,Y}(x,y)}{P_{X}(x)} & P_{X}(x)>0\\ 0 & otherwise\end{cases}$$

2. **נוסחת הכפל:**
   - הטענה נובעת מכלל הכפל עבור המאורעות $\{X=x\}$ ו-$\{Y=y\}$.
     - כלומר אם $P_X(x)>0$ אז נקבל:

$$P(\{X=x\}\cap\{Y=y\})=P(\{X=x\})\cdot P(\{Y=y\}\mid\{X=x\})$$

$$=P_X(x)\cdot P_{Y|X}(y,x)$$

     - ואם $P_X(x)=0$ אז יתקיים גם $P_{X,Y}(x,y)=0$ ולכן השוויון עדיין מתקיים.

3. **נוסחת בייס:**
   - אם $P_Y(y)>0$ אז נחלק את טענה 2 במשפט ב-$P_Y(y)$ כדי לקבל את טענה 3 של המשפט.

4. **נוסחת ההסתברות השלמה:**
   - נשים לב כי עבור כל $x\in Supp(X)$, המאורעות $\{X=x\}$ מהווים חלוקה בת מנייה של Ω.
     - זאת מכיוון שהם לא יכולים לקרות בו זמנית וגם האיחוד שלהם יהיה Ω.
       * נשים לב כי יש פה קצת רמאות בטענה הזו והיא לא מדויקת (ההסבר למה זה לא מדויק הושאר כתרגיל).
   - ולכן:

$$P_Y(y)=P(\{Y=y\})=\sum_{x:P_X(x)>0}P(\{Y=y\}\mid\{X=x\})\cdot P\{X=x\}$$

$$=\sum_{x:P_X(x)>0}P_{Y|X}(y,x)\cdot P_X(x)$$

■

משפט 3. תוחלת של פונקציה של משתנה מקרי

- יהיו $X_1,X_2,\ldots,X_n$ משתנים מקריים בדידים על (Ω,P) ותהיה $g:\mathbb{R}^n\to\mathbb{R}$.
- נגדיר וקטור מקרי חדש $Y=g(X_1,\ldots,X_n)$ (גם Y הוא וקטור מקרי)
- התוחלת של Y היא:

$$E[Y]=\sum_{x_1,x_2,\ldots,x_n}g(x_1,x_2,\ldots,x_n)\cdot P_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)$$

– כל זאת בהנחה שהתוחלת של Y מוגדרת, כלומר שהטור מתכנס בהחלט.

הוכחה.

- נבחן את $E\left[Y\right]$ לפי הגדרה:

$$E\left[Y\right]=\sum_{y}y\cdot P_{Y}\left(y\right)$$

- לפי משפט 1 שראינו בהרצאה הזו (פונקציה של וקטור מקרי), מתקיים:

$$P_{Y}\left(y\right)=\sum_{\underline{x}\ :\ g\left(\underline{x}\right)=y}P_{X}\left(x\right)$$

– נציב זאת ב$E\left[Y\right]$ ונקבל:

$$=\sum_{y}y\cdot\left(\sum_{\underline{x}\ :\ g\left(\underline{x}\right)=y}P_{X}\left(x\right)\right)$$

* מכיוון ש-y לא תלוי ב-x, ניתן להכניס אותו לסכום הפנימי ועבור כל $\underline{x}$ ניתן לסמנו $g\left(\underline{x}\right)$:

$$=\sum_{y}\left(\sum_{\underline{x}\ :\ g\left(\underline{x}\right)=y}g\left(\underline{x}\right)\cdot P_{X}\left(x\right)\right)$$

* כאמור, הנחנו שהטור מתכנס בהחלט.
  · ולכן ניתן לקבץ את האיברים ולקבל שהסכימה החיצונית מיותרת.
  · כלומר מתקיים:

$$=\sum_{\underline{x}}g\left(\underline{x}\right)\cdot P_{X}\left(x\right)$$

$\blacksquare$

משפט 4. תכונות בסיסיות של תוחלת (תכונות 1-3 נקראות "לינאריות")

- יהיו X,Y משתנים מקריים בדידים על אותו $\left(\Omega,P\right)$ ויהי $a\in\mathbb{R}$ קבוע.
- אזי מתקיים:

1. אם $P_{X}\left(x\right)=\begin{cases}1 & x=a\\0 & otherwise\end{cases}$ (כלומר $P\left(X=a\right)=1$) אז $E\left[X\right]=a$.

2. אם $Z=a\cdot X$ אז $E\left[Z\right]=E\left[a\cdot X\right]=a\cdot E\left[X\right]$
3. אם $Z=X+Y$ אז $E\left[Z\right]=E\left[X+Y\right]=E\left[X\right]+E\left[Y\right]$
4. אם X,Y בלתי תלויים אז מתקיים $E\left[X\cdot Y\right]=E\left[X\right]\cdot E\left[Y\right]$

מסקנה 5. נובע: $E\left[a\cdot X+b\cdot Y+c\right]=a\cdot E\left[X\right]+b\cdot E\left[Y\right]+c$

הוכחה.

1. אם $P_{X}\left(x\right)=\begin{cases}1 & x=a\\0 & otherwise\end{cases}$ (כלומר $P\left(X=a\right)=1$) אז $E\left[X\right]=a$

- פונקציית ההסתברות היא:

$$P_{X}\left(x\right)=\begin{cases}1 & x=a\\0 & otherwise\end{cases}$$

- ואז התוחלת לפי הגדרה היא:

$$E\left[X\right]=\sum x\cdot P_{X}\left(x\right)=a\cdot\overbrace{P_{X}\left(a\right)}^{=1}=a$$

2. אם $Z=a\cdot X$ אז $E\left[Z\right]=a\cdot E\left[X\right]$

- נגדיר משתנה מקרי חדש $Y=g\left(X\right)$, כאשר $g\left(x\right)=a\cdot x$
- כעת התוחלת לפי הגדרה היא:

$$\begin{aligned}E\left[a\cdot X\right]=E\left[Y\right] & =\sum_{x}g\left(x\right)P_{X}\left(x\right)\\
 & =\sum_{x}a\cdot xP_{X}\left(x\right)\\
 & =a\cdot\overbrace{\sum_{x}x\cdot P_{X}\left(x\right)}^{=E\left[X\right]}\\
 & =a\cdot E\left[X\right]\end{aligned}$$

3. אם $Z=X+Y$ אז $E\left[Z\right]=E\left[X\right]+E\left[Y\right]$

- נסמן $g\left(x,y\right)=x+y$

- בדומה לסעיף הקודם, מתקיים:

$$E\left[X+Y\right]=\sum_{x,y}\overbrace{(x+y)}^{=g(x,y)}P_{X,Y}\left(x,y\right)$$

– נפתח את הטור ונקבל:

$$=\sum_{x,y}xP_{X,Y}\left(x,y\right)+\sum_{x,y}yP_{X,Y}\left(x,y\right)$$

– מכיוון שהטור מתכנס בהחלט, מתקיים:

$$=\sum_{x}\left(\sum_{y}x\cdot P_{X,Y}\left(x,y\right)\right)+\sum_{y}\left(\sum_{x}y\cdot P_{X,Y}\left(x,y\right)\right)$$

* ניתן להוציא את x,y מהסכומים הפנימיים ולקבל:

$$=\sum_{x}x\cdot\overbrace{\left(\sum_{y}P_{X,Y}\left(x,y\right)\right)}^{=P_X(x)}+\sum_{y}y\cdot\left(\sum_{x}P_{X,Y}\left(x,y\right)\right)$$

· ומכיוון ש$\sum_y P_{X,Y}\left(x,y\right)$ זו ההסתברות השולית לפי X, מתקיים $\sum_y P_{X,Y}\left(x,y\right)=P_X\left(x\right)$

· ובאותו אופן, $\sum_x P_{X,Y}\left(x,y\right)=P_Y\left(y\right)$ כי זו ההסתברות השולית לפי Y.

* כעת מתקיים:

$$=E\left[X\right]+E\left[Y\right]$$

4.

- נסמן $g\left(x,y\right)=x\cdot y$
- מתקיים:

$$E\left[X\cdot Y\right]=\sum_{x,y}\left(x\cdot y\right)P_{X,Y}\left(x,y\right)$$

- X,Y בלתי תלויים ולכן:

$$P_{X,Y}\left(x,y\right)=P_X\left(x\right)\cdot P_Y\left(y\right)$$

- נציב זאת ונקבל:

$$E\left[X\cdot Y\right]=\sum_{x,y}\left(x\cdot y\right)P_{X}\left(x\right)\cdot P_{Y}\left(y\right)$$

- הנחנו שהטור מתכנס בהחלט ולכן ניתן לפצל את הסכימה:

$$E\left[X\cdot Y\right]=\sum_{x}\sum_{y}\left(x\cdot y\right)P_{X}\left(x\right)\cdot P_{Y}\left(y\right)$$

$$=\sum_{x}x\cdot P_{X}\left(x\right)\cdot\sum_{y}y\cdot P_{Y}\left(y\right)$$

$$=E\left[X\right]\cdot E\left[Y\right]$$

$\blacksquare$

משפט 6. מציאת פונקציית הסתברות שולית מפונקציית הסתברות משותפת:

- אם $\underline{X}$ וקטור מקרי n מימדי $\underline{X}=X_{1},X_{2},\ldots,X_{n}$.
- אז לכל k מתקיים:

$$P_{X_{k}}=\sum_{x_{1},\ldots,x_{k-1},x_{k+1},\ldots,x_{n}}P_{\underline{X}}\left(x_{1},\ldots,x_{k-1},x_{k+1},\ldots,x_{n}\right)$$

הוכחה.

- נגדיר פונקציה $g:\mathbb{R}^{n}\to\mathbb{R}$ באופן הבא:

$$g\left(\underline{x}\right)=x_{k}$$

- אזי לפי משפט 1 מתקיים:

$$P_{X_{k}}\left(y\right)=P_{Y}\left(y\right)=\sum_{\underline{x}\ :\ g\left(\underline{x}\right)=y}P_{\underline{X}}\left(\underline{x}\right)$$

  - נציב את הגדרת g, שהיא $g\left(\underline{x}\right)=x_{k}$:

$$\sum_{\underline{x}\ :\ x_{k}=y}P_{\underline{X}}\left(\underline{x}\right)$$

* ונשים לב שהסכימה הזו היא:

$$= \sum_{x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n} P_{\underline{X}}\left(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n\right)$$

■

דוגמה 7. תוחלת של נמלה.
נמלה הולכת על ציר המספרים השלמים.
בכל צעד, בסיכוי $p \in (0,1)$ היא הולכת צעד אחד $(+1)$ ימינה, אחרת היא הולכת צעד אחד שמאלה (-1).
כל צעד בלתי תלוי בצעדים האחרים.
צ"ל: מצאו תוחלת מיקום הנמלה אחרי n צעדים.
פיתרון:

- נסמן ב-X את מיקום הנמלה.
- דרך א':
  - אזי:

$$E\left[X\right] = \sum_x x \cdot P_X\left(x\right)$$

  - נבחן את $P_X\left(x\right)$:

$$P_X\left(x\right) = \begin{cases} ? & -n \le x \le n,\ x \in \mathbb{Z},\ n\ and\ x\ are\ both\ odd\ or\ both\ even \\ 0 & otherwise \end{cases}$$

- דרך ב':
  - נגדיר את Y = מספר צעדי ימין מתוך n הצעדים הראשונים.
    * מתקיים כי Y מתפלג בינומית עם הפרמטרים (n,p) והתוחלת שלו היא $n \cdot p$.
  - הקשר בין Y ל-X הוא:

$$X = \overbrace{(+1) \cdot Y}^{Right\ \ steps} + \overbrace{(-1) \cdot (n - Y)}^{Left\ \ steps}$$

$$X = 2 \cdot Y - n$$

* כלומר:

$$E[X]=E[2\cdot Y-n]$$

$$=2\cdot E[Y]-n$$

$$=2\cdot(n\cdot p)-n$$

- דרך ג':
  - נגדיר Z = מספר צעדי שמאלה.
    * אזי Z מתפלג בינומית אם $n,1-p$.
    * מתקיים:

$$X=Y-Z$$

· כלומר אלה מאורעות תלויים, כי אם יודעים את Y אז יודעים את Z.
· ועדיין, לפי לינאריות התוחלת מתקיים:

$$E[X]=E[Y-Z]=E[Y]-E[Z]$$

$$=n\cdot p-n\cdot(1-p)$$

$$=n\cdot(2p-1)$$

הערה:
אם $p>\frac{1}{2}$ אז יש סיכוי שהיא תמשיך עד אינסוף.
אם $p<\frac{1}{2}$ אז יש סיכוי
אם $p=\frac{1}{2}$ אז היא בטוח תחזור לאפס מתישהו.

נמלה בדו-מימד כן תחזור לאפס. בממוצע לוקח לה אינסוף צעדים לחזור.
בשלושה מימדים ומעלה, יש סיכוי שהנמלה לא תחזור.

דוגמה 8.

- בכד יש G כדורים ירוקים ו-$N-G$ כדורים כחולים.
- שולפים כדור אחר כדור, n כדורים, ללא החזרה.
- הניחו סיכוי שווה לכל כדור.
- מהי תוחלת מספר הירוקים במדגם?

פיתרון:

- דרך א':
  - נסמן ב-X את מספר הכדורים הירוקים שנשלפו.
    - X מתפלג היפר-גאומטרי.
    - כלומר:

$$P_X(x)=\begin{cases}\dfrac{\binom{G}{x}\binom{N-G}{N-x}}{\binom{N}{n}} & 0\le x\le G,\ \ 0\le n-x\le N-G\\ 0 & otherwise\end{cases}$$

- דרך ב':
  - נגדיר מאורעות:
    - עבור $1\le k\le n$ נגדיר A_k = הכדור שהוצא הוא ירוק.
      - מתקיים:

$$P(A_k)=\frac{G}{N}$$

  - נגדיר משתנים מקריים:
    - עבור $1\le k\le n$ נגדיר I_k:

$$I_k=\begin{cases}1 & A_k\ \ happened\\ 0 & A_k\ \ didn't\ \ happen\end{cases}$$

      - למשל, אם $n=5$ אז:

k	1	2	3	4	5
$Ball$	$green$	$blue$	$blue$	$green$	$green$
I_k	1	0	0	1	1
A_k	קרה	לא קרה	לא קרה	קרה	קרה

    - נשים לב כי $\sum_k I_k=X$

- אזי מתקיים:

$$E\left[X\right]=E\left[\sum_{k=1}^{n}I_{k}\right]$$

  – ומלינאריות התוחלת מתקיים:

$$\sum_{k=1}^{n}E\left[I_{k}\right]$$

- נשים לב כי לפי הגדרת התוחלת מתקיים

$$E\left[I_{k}\right]=\sum_{i}i\cdot P\left(I_{k}=i\right)$$

$$=0\cdot P\left(I_{k}=0\right)+1\cdot P\left(I_{k}=1\right)$$

$$=P\left(I_{k}=1\right)$$

$$=P\left(A_{k}\right)$$

$$=\frac{G}{N}$$

- ולכן:

$$E\left[X\right]=\sum_{k=1}^{n}\frac{G}{N}=n\cdot\frac{G}{N}$$

הכללה של העיקרון מהדוגמאות - שיטת האינדיקטורים:

הגדרה 9. אינדיקטור.

- אם A מאורע, אז למשתנה המקרי I המוגדר על ידי:

$$I=\begin{cases}1 & A\ happened\\0 & A\ didn't\ happen\end{cases}$$

- קוראים "המציין" או ה"אינדיקטור" של A.
- סימון: I_A

<u>הערה:</u>

- מתקיים:

$$E\left(I_A\right)=E\left(A\right)$$

- ומתקיים:
  - אם $A_1,A_2,\ldots,A_n$ הם מאורעות ו-X הוא משתנה מקרי שסופר כמה מתוכם קרו, אז ניתן לרשום את X על ידי:

$$X=\sum_{k=1}^{n}I_{A_k}$$

  - בפרט מתקיים:

$$E\left[X\right]=E\left[\sum_{k=1}^{n}I_{A_k}\right]=\sum_{k=1}^{n}E\left[I_{A_k}\right]$$

$$=\sum_{k=1}^{n}P\left(A_k\right)$$

- הטענה הזו נכונה גם עבור $\left(A_k\right)_{k=1}^{\infty}$, כלומר עבור מספר בן מנייה של מאורעות.

<u>הערה שנייה:</u>

- בשיטת האינדיקטורים אין צורך לדרוש ש-A_k יהיו בלתי תלויים.
  - ובדוגמה השנייה באמת התקיים ש-$A_1,A_2,\ldots,A_n$ הם מאורעות תלויים.

<u>הערה שלישית:</u>

- נניח כי בדוגמה השנייה היינו מחזירים כל כדור שנשלף.
- אז X היה מפולג בינומית עם הפרמטרים $n,p=\frac{G}{N}$
  - ומתקיים:

$$E\left[X\right]=n\cdot p=n\cdot\frac{G}{N}$$

- נשים לב שהפיתרון בשיטת האינדיקטורים נשאר תקף גם במקרה הזה כאשר $A_1, A_2, \ldots, A_n$ בלתי תלויים.

משפט 10. נוסחת הזנב.

- אם X הוא משתנה מקרי המקבל רק ערכים טבעיים, אז:

$$E\left[X\right]=\sum_{x=0}^{\infty}P\left(X>x\right)=\sum_{x=1}^{\infty}P\left(X\geq x\right)$$

הוכחה.

- נגדיר $A_k=\{X\geq k\}$ עבור $k=1,2,\ldots,$
- אזי מתקיים צָ סופר כמה מתוך המאורעות A_1, A_2 קרו
- באופן מפורש מתקיים:

$$X=\sum_{k=1}^{\infty}I_{A_k}=\sum_{k=1}^{m}I_{\{X\geq k\}}$$

  – ולכן התוחלת היא:

$$E\left[X\right]=E\left(\sum_{k=1}^{\infty}I_{\{X\geq k\}}\right)$$

$$=\sum_{k=1}^{\infty}P\left(X\geq k\right)$$

* נסמן $k'=k+1$

$$=\sum_{k'=0}^{\infty}P\left(X\geq k'+1\right)$$

* מכיוון ש-X שלם, מתקיים:

$$=\sum_{k'=0}^{\infty}P\left(X>k'\right)$$

■

משפט 11. תכונות נוספות של תוחלת.
יהיו X,Y משתנים מקריים בדידים על אותו Ω.

1. אם $P\left(X\geq0\right)=1$ אז $E\left[X\right]\geq0$
2. אם $P\left(X\geq0\right)=1$ וגם $E\left[X\right]=0$ אז $P\left(X=0\right)=1$
3. אם $P\left(X\geq Y\right)=1$ אז $E\left[X\right]\geq E\left[Y\right]$

הוכחה.

1. **צ"ל:** אם $P\left(X\geq0\right)=1$ אז $E\left[X\right]\geq0$
   - נניח כי $P\left(X\geq0\right)=1$
     - זאת אומרת ש-$Supp\left(X\right)\subseteq[0,\infty)$
   - ולכן לכל $x<0$ מתקיים $P_X\left(x\right)=0$
   - כלומר

$$E\left[X\right]=\sum_{x}x\cdot P_X\left(x\right)=\sum_{x\geq0}x\cdot P_x\left(x\right)\geq0$$

2. **צ"ל:** אם $P\left(X\geq0\right)=1$ וגם $E\left[X\right]=0$ אז $P\left(X=0\right)=1$
   - לפי החלק הראשון של המשפט, מתקיים:

$$E\left[X\right]=\sum_{x\geq0}x\cdot P_X\left(x\right)=0$$

     - ולכן לכל x מתקיים:

$$x\cdot P_X\left(x\right)=0$$

   - אם $x\neq0$ אז $P_X\left(x\right)=0$
     - כלומר $P_X\left(x\right)=1$
3. **צ"ל:** אם $P\left(X\geq Y\right)=1$ אז $E\left[X\right]\geq E\left[Y\right]$
   - נניח כי $P\left(X\geq Y\right)=1$.
   - נגדיר $Z=X-Y$
     - אזי מתקיים:

$$P\left(Z\geq0\right)=P\left(X-Y\geq0\right)=1$$

- ולכן לפי החלק הראשון של המשפט מתקיים:

$$E\left[Z\right]\geq0$$

- ומלינאריות של התוחלת, מתקיים:

$$E\left[Z\right]=E\left[X\right]-E\left[Y\right]$$

■

נושא שלישי - שונות:

דוגמה 12. רולטה.

- ברולטה יש 37 מספרים (מ-0 עד 36):
  - $2,4,6,8,\ldots,36$ - אדום
  - $1,3,5,7,\ldots,35$ - שחור.
  - 0 - ירוק.
- נבחר מספר באקראי. יש 2 הימורים אפשריים:
  1. על צבע (שחור או אדום)
     - אם זוכים מקבלים בחזרה פי 2
  2. על מספר
     (א) אם זוכים מקבלים בתמורה פי 36.9
- יש לנו שתי אפשרויות משחק:
  1. להמר 1 ש"ח על צבע שחורך
     - אם יצא שחור - הרווח הוא $(-1)+2=+1$
     - אחרת, הרווח הוא $-1=-1$
  2. להמר 10 ש"ח על מספר 17.
     - אם יצא 17 אז הרווח הוא $-10+369=+359$
     - אחרת, הרווח הוא $=-10$
- במה נבחר?

פיתרון:

- נגדיר משתנים מקריים:

- X = רווח בסגנון משחק 1
- Y = רווח בסגנון משחק 2

- מתקיים:

$$P_X(x)=\begin{cases}\frac{18}{37} & x=1\\ \frac{19}{37} & x\neq 1\\ 0 & otherwise\end{cases}$$

$$P_Y(y)=\begin{cases}\frac{1}{37} & y=359\\ \frac{36}{37} & y=-10\\ 0 & otherwise\end{cases}$$

- נחשב תוחלת עבור X:

$$E[X]=\sum_x x\cdot P_X(x)=1\cdot\frac{18}{37}+(-1)\cdot\frac{19}{37}=-\frac{1}{37}$$

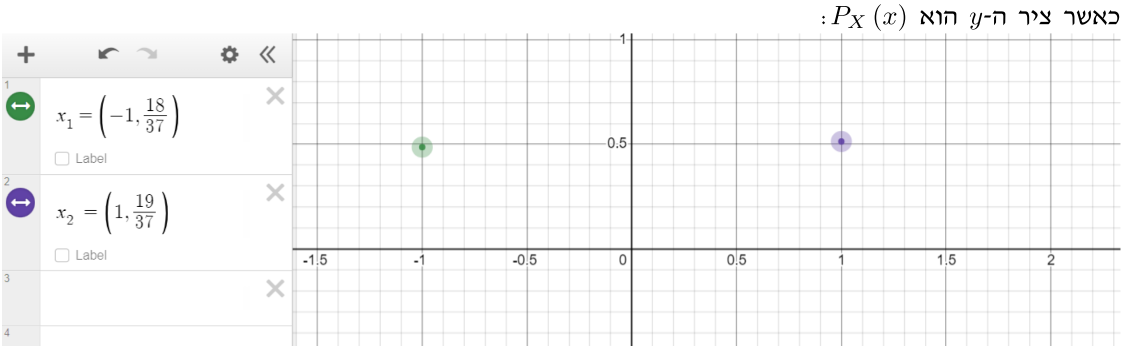
- נחשב תוחלת עבור Y:

$$E[Y]=\sum_y y\cdot P_Y(y)=359\cdot\frac{1}{37}+(-10)\cdot\frac{36}{37}=-\frac{1}{37}$$

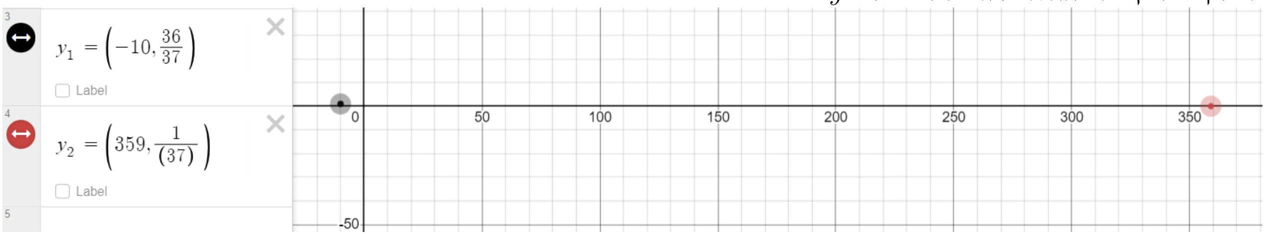
- כלומר:

$$E[X]=E[Y]$$

- מכיוון שממוצע הרוח למשחק הוא התוחלת (מהגדרת התוחלת), נקבל שממוצע הרווחים בשני סגנונות המשחק זהה.
- נתבונן בפונקציית ההסתברות של X לפי x:
  - כאשר ציר ה-y הוא $P_X(x)$:



- נתבונן בפונקציית ההסתברות של Y לפי y:



- אמנם בשני הסגנונות הרווח הממוצע למשחק הוא $-\frac{1}{37}$, בסגנון הראשון ה"סטיות" סביב הממוצע נוטות להיות קטנות יותר מאשר בסגנון השני.
- נרצה לחפש הגדרה לגודל מתמטי שיכמת את גדלי הסטיות (גודל הפיזור) הללו.

- לכן נגדיר:

הגדרה 13. שונות, סטיית תקן.

- בהינתן משתנה מקרי X בדיד עם תוחלת מוגדרת, ה**שונות** שלו היא:

$$\overbrace{Var}^{Variance}(X) = E\left[(X - E(X))^2\right]$$

  – בהנחה שהתוחלת מוגדרת.

- **סטיית התקן** של X היא:

$$\overbrace{SD}^{Standard\ \ deviation} = \sqrt{Var(X)}$$

- <u>משמעות</u>:
  – אם X מייצג גודל אקראי בניסוי, אז השונות של X היא ממוצע הסטיות של הגודל הזה מהממוצע שלו מועלות בריבוע על פי $n \to \infty$ חזרות על הניסוי.
  – סטיית התקן היא שורש של הגול הזה.

דוגמה 14. שונות של רולטה.

- עבור הדוגמה עם הרולטה, נחשב את השונות בשני הסגנונות:
  – עבור הסגנון הראשון:

$$Var(X) = E\left[(X - E[X])^2\right]$$

$$= E\left[\left(X - \left(-\frac{1}{37}\right)\right)^2\right]$$

* זו תוחלת של פונקציה של X, ולכן:

$$= \sum_x \left(\left(x - \left(-\frac{1}{37}\right)\right)^2 \cdot P_X(x)\right)$$

$$= \left(1 + \frac{1}{37}\right)^2 \frac{18}{37} + \left(-1 + \frac{1}{37}\right) \cdot \frac{19}{38} \approx 1.002$$

– עבור הסגנון השני:

$$Var(Y) = E\left[(Y - E[Y])^2\right]$$

$$= E\left[\left(Y + \frac{1}{37}\right)^2\right]$$

$$= \sum_y \left(\left(y + \frac{1}{37}\right)^2 \cdot P_Y(y)\right)$$

$$= \left(-10 + \frac{1}{37}\right)^2 \cdot \frac{36}{37} + \left(359 + \frac{1}{37}\right)^2 \cdot \frac{1}{37}$$

$$\approx 3700$$

• ואם נחשב סטיית תקן, נקבל:

$$SD(X) = \sqrt{Var(x)} \approx 1$$

$$SD(Y) = \sqrt{Var(Y)} \approx 65$$

– באופן לא פורמלי, הסטיות מהממוצע נוטות להיות גדולות יותר בסגנון השני מאשר בסגנון הראשון.

משפט 15. תכונות של שונות.

- יהיו X,Y משתנים מקריים בדידים עם תוחלות ושונות מוגדרות.
- יהיו $a,b\in\mathbb{R}$ קבועים.
- מתקיים:

1. תמיד $Var\left(X\right)\geq0$
2. אם $Var\left(X\right)=0$ אז $P\left(X=E\left[X\right]\right)=1$
3. $Var\left(a\cdot X+b\right)=a^{2}\cdot Var\left(X\right)$
4. אם X,Y בלתי תלויים אז $Var\left(X\pm Y\right)=Var\left(X\right)+Var\left(Y\right)$
5. $Var\left(X\right)=E\left[X^{2}\right]-\left(E\left[X\right]\right)^{2}$

דוגמה 16.

- מצאו שונות של משתנה מקרי גאומטריי עם פרמטר $p\in\left(0,1\right)$.

פיתרון:

- יהא $X\sim Geo\left(p\right)$.
  - ראינו שהתוחלת של משתנה מקרי גאומטרי היא $\frac{1}{p}$, ולכן $E\left[X\right]=\frac{1}{p}$.
- מחפשים את $Var\left(X\right)$.
- דרך א' - לפי הגדרה:

$$Var\left(X\right)=E\left[\left(X-E\left[X\right]\right)^{2}\right]$$

$$=E\left[\left(X-\frac{1}{p}\right)^{2}\right]$$

  - זו תוחלת של פונקציה של משתנה מקרי, ולכן:

$$=\sum_{x}\left(x-\frac{1}{p}\right)^{2}\cdot P_{X}\left(x\right)$$

$$=\sum_{x}\left(x-\frac{1}{p}\right)^{2}\cdot q^{x-1}\cdot p$$

- דרך ב' - שימוש במשפט:
  - לפי סעיף 5 במשפט האחרון, $Var\left(X\right)=E\left[X^{2}\right]-\left(E\left[X\right]\right)^{2}$.

* נשים לב כי $E\left[X^2\right]$ זו תוחלת של פונקציה של משתנה מקרי:

$$E\left[X^2\right] = \sum_x x^2 \cdot P_X(x)$$

$$= \sum_{x=1}^{\infty} x^2 \cdot q^{x-1} \cdot p$$

· נרשום את הטור הזה בצורה אחרת:

$$= p \cdot q \cdot \sum_{x=1}^{\infty} x(x-1) q^{x-2} + \overbrace{\sum_{x=1}^{\infty} x \cdot q^{x-1} \cdot p}^{=E[X]}$$

· נשים לב כי הטור השמאלי ניתן לכתיבה כנגזרת:

$$p \cdot q \cdot \sum_{x=1}^{\infty} x(x-1) q^{x-2} = p \cdot q \cdot \frac{d^2}{dq^2}\left(\overbrace{\sum_{x=0}^{\infty} q^x}^{=\frac{1}{1-q}}\right)$$

· כלומר:

$$E\left[X^2\right] = p \cdot q \cdot \frac{2}{(1-q)^3} + \frac{1}{p}$$

$$= p \cdot q \cdot \frac{2}{\left(\underbrace{1-q}_{=p}\right)^3} + \frac{1}{p}$$

$$= \frac{2q}{p^2} + \frac{1}{p}$$

$$E\left[X^2\right] = \frac{2q+p}{p^2}$$

* ולכן מתקיים:

$$Var(X) = E\left[X^2\right] - (E[X])^2$$

$$= \frac{2q+p}{p^2} - \frac{1}{p^2}$$

- ולכן:

$$= \frac{2 \cdot q + p - 1}{p^2}$$

$$= \frac{2q - \left(\overbrace{1-p}^{=q}\right)}{p^2}$$

$$= \frac{q}{p^2}$$

- לדוגמה, אם $p = \frac{1}{10}$, ממוצע הטלות מטבע עד לקבלת "ראש" הוא $\frac{1}{p} = 10$
  - וממוצע ההפרש בין מספר ההטלות עד לקבלת "ראש" מועלות בריבוע הוא $90 = \frac{\frac{9}{10}}{\left(\frac{1}{10}\right)^2}$
    - השורש של ממוצע ההפרש הזה הוא בערך 9.5