זוג משתנים אקראיים בדידים

מטרה לאפיין ניסוי אקראי בעל שתי תוצאות, והקשר בין התוצאות האלה.

הרצאה הקודמת:
$p[X=x_j]$
$F_X(x)=p(X\leqslant x)$

PDF (הגדרה 2.1):

$$p_{XY}[x_j,y_k]=\Pr[X=x_j,Y=y_k]$$

CDF (הגדרה 2.2):

$$F_{XY}(x,y)=p(X\leqslant x,Y\leqslant y)$$

תזכורת
$p(X=0)=1-p$
$p(X=1)=p$

דוגמה: זוג תוצאות של ניסוי Bernoulli

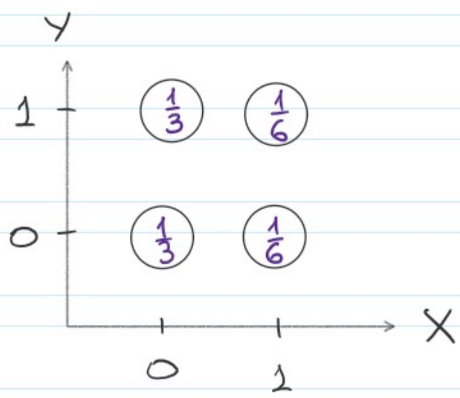
$X\sim Ber(p_1)$ $\quad p_1=\frac{1}{3}$

$Y\sim Ber(p_2)$ $\quad p_2=\frac{1}{2}$

ניתוח דוגמה:

* תוצאות ניסוי אפשריות
* הסתברויות לתוצאות

PDF

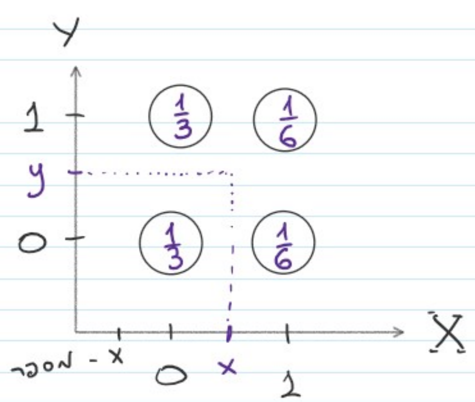


$$p[X=0,Y=0]=(1-p_1)(1-p_2)$$
$$=\left(1-\frac{1}{3}\right)\left(1-\frac{1}{2}\right)=$$

$$p[X=0,Y=1]=(1-p_1)\,p_2$$
$$p[X=1,Y=0]=p_1\cdot(1-p_2)=\frac{1}{6}$$

CDF $\quad F_{XY}(x,y)=p(X\leqslant \underset{\smile}{x},Y\leqslant \underset{\smile}{y})$

$x,y\in\mathbb{R}$



* $x<0$

$$p(X\leqslant x,Y\leqslant y)=0$$

$y<0$

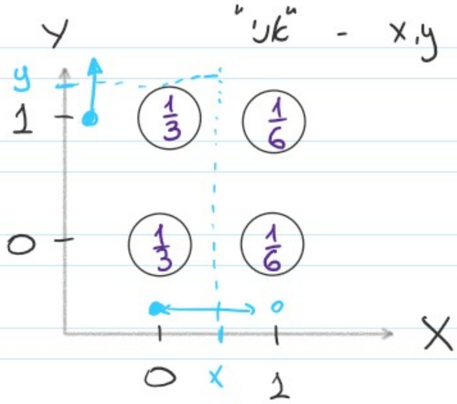
$$F_{XY}(x,y)=0 \qquad \begin{matrix}x<0\\ \text{או}\\ y<0\end{matrix}$$

* $0\leqslant x<1$
$0\leqslant y<1$

מספר - ה   0 x 1

X,Y - ניסוי

x,y - "אני"



* $0\le x<1$
$0\le y<1$

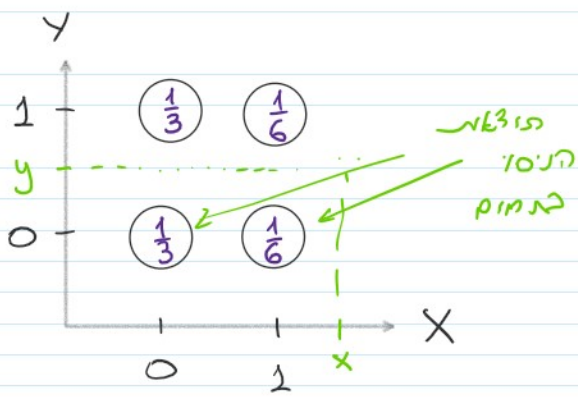
$$F_{XY}(x,y)=\underbrace{P(X=0,Y=0)}_{\text{נקודה היחידה בתחום}}=1/3$$

$$F_{XY}(x,y)=p(X\leqslant x,Y\leqslant y)$$

* $0\le x<1$
$1\le y$
ישנם 2 תוצאות ניסוי בתחום

$$F_{XY}(x,y)=\frac{1}{3}+\frac{1}{3}=\frac{2}{3}$$

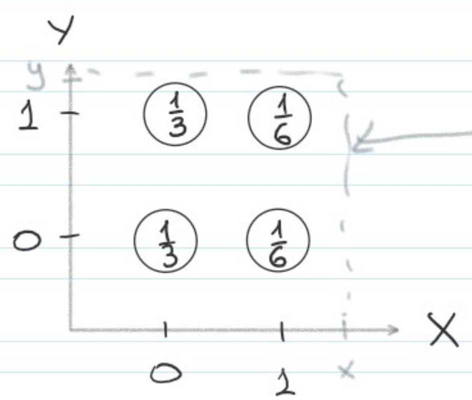
* $1\le x$
$0\le y<1$



$$F_{XY}(x,y)=\frac{1}{6}+\frac{1}{3}=\frac{1}{2}$$

* $1\le x$
$1\le y$

כל תוצאות ניסוי אפשריות בתחום



$$F_{XY}(x,y)=1$$

תכונות PDF (תכונה 2.1): תחום ערכים ו"סכום" ערכים

$$0\leqslant p_{XY}[x_j,y_k]\leqslant 1\ \forall i,k$$

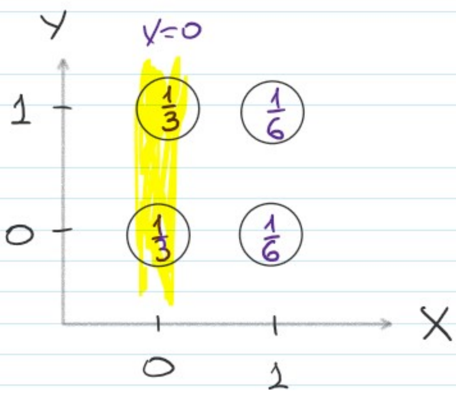
$$\sum_{j,k}p_{XY}[x_j,y_k]=1$$

2.1.1 התפלגות שולית

מטרה לקבל/לאפיין התפלגות של כל אחד מהמשתנים בנפרד, מתוך התפלגות משותפת.

$$F_X(x) = F_{XY}(x,\infty) \qquad p_X[x_k] = \sum_j p_{XY}[x_k,y_j]$$
$$F_Y(y) = F_{XY}(\infty,y) \qquad p_Y[y_j] = \sum_k p_{XY}[x_k,y_j]$$

שאלה: $p[X=x_k \text{ , כלשהו } Y]$



דוגמה: חשב $p[X=0]$

$$p[X=0] = p[X=0,Y=0] + p[X=0,Y=1]$$
$$= \sum_k p[X=0,Y=k]$$
$$p[Y=0] = \tfrac{1}{3}+\tfrac{1}{6} = \tfrac{1}{2} = p[Y=1]$$
$$p[X=1] = \tfrac{1}{6}+\tfrac{1}{6} = \tfrac{1}{3}$$

2.1.2 אי תלות סטטיסטית

מטרה לאפיין מצב של תוצאות ניסוי בלתי תלויות.

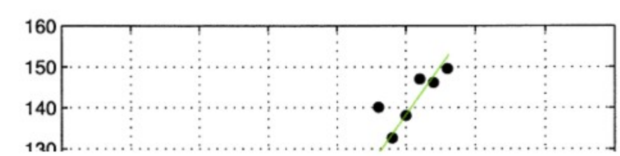
משתנים בלתי תלויים סטטיסטית (הגדרה 2.4): משתנים נקראים בלתי תלויים סטטיסטית אם ורק אם מתקיים $p[X=x_k, Y=y_j] = p[X=x_k]\cdot p[Y=y_j]$

$$p_{XY}[x_k,y_j] = p_X[x_k]p_Y[y_j] \tag{א2.5}$$
$$F_{XY}(x,y) = F_X(x)F_Y(y) \tag{ב2.5}$$

מספיק אחד מהן

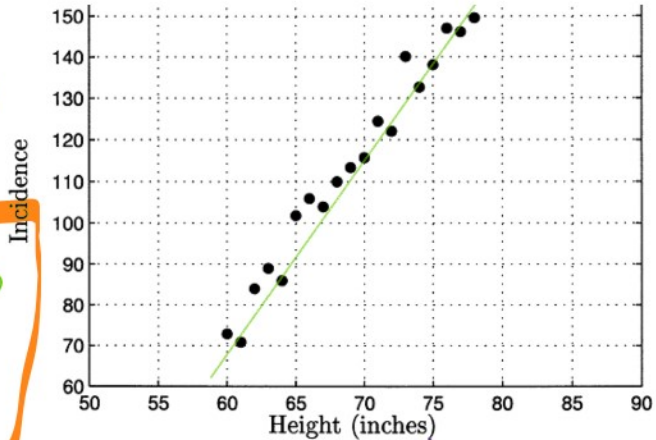
הערה 2.1 ! התפלגות משותפת ⇐ התפלגות שולית.
התפלגות שולית ⇐ התפלגות משותפת רק במקרה של משתנים בלתי תלויים.

דוגמה לתלות מדומה



מקרים
ל- 1000

הגרף כולל
נשים בתוך
סטטיסטיקה!



גובה

Figure 7.10: Incidence of prostate cancer per 1000 individuals older than age 55 versus height.

*נשים
לרוב יותר
נמוכות
מגברים

תוחלת

תכונות:

$$E[g(X)] = \underbrace{\sum}_{\text{סכום על כל הערכים אפשריים}} \underbrace{g(x_k)}_{\text{ערך}} \underbrace{p_X(x_k)}_{\text{סיכוי לערך זה}}$$

$$E[g(X,Y)] = \sum_k \sum_j g(x_k, y_j)\, p_{XY}[x_k, y_j]$$

* ליניאריות של תוחלת

$$E[aX + bY] = aE[X] + bE[Y]$$

Kay עמוד 187

הוכחה עבור $a=b=1$ $\quad g(X,Y)=x+y$

$$
\begin{aligned}
E_{X,Y}[X+Y] &= \sum_i \sum_j (x_i + y_j) p_{X,Y}[x_i, y_j] \\
&= \sum_i \sum_j x_i p_{X,Y}[x_i, y_j] + \sum_i \sum_j y_j p_{X,Y}[x_i, y_j] \\
&= \sum_i x_i \underbrace{\sum_j p_{X,Y}[x_i, y_j]}_{p_X[x_i]} + \sum_j y_j \underbrace{\sum_i p_{X,Y}[x_i, y_j]}_{p_Y[y_j]} \quad \text{(from (7.6))} \\
&= E_X[X] + E_Y[Y] \quad \text{(definition of expected value).}
\end{aligned}
$$

פירוק ל- 2 סכומים

התפלגות שולית - $p_X[x_i]$

תכונות של משתנים בלי תלויים (בלבד) (תכונה 2.3):

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

$$E[g_1(X)g_2(Y)] = E[g_1(X)]E[g_2(Y)]$$

$$\mathrm{Var}[aX+bY] = a^2\,\mathrm{Var}[X] + b^2\,\mathrm{Var}[Y]$$

הרצאה קודמת

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

קשר בין המשתנים

חסר $Var[X+Y]$

$$Z = X+Y \qquad \#1 \quad Var[Z] = E\Big[\big(X+Y-\overbrace{E[X+Y]}^{E[X]+E[Y]}\big)^2\Big]$$

$$\#2 \quad = E\Big[\big(\underbrace{X-E[X]}_{a} + \underbrace{Y-E[Y]}_{b}\big)^2\Big]$$

$$(a+b)^2 = a^2 + 2ab + b^2$$

$$\#3 \quad = \underbrace{E\Big[\big(X-E[X]\big)^2\Big]}_{Var[X]} + 2\,\underbrace{E\Big[\big(X-E[X]\big)\big(Y-E[Y]\big)\Big]}_{Cov[X,Y]} + \underbrace{E\Big[\big(Y-E[Y]\big)^2\Big]}_{Var[Y]}$$

$$\#4 \quad Var[X] + 2\,Cov[X,Y] + Var[Y]$$

Covariance = שונות משותפת

שונות משותפת (covariance) (הגדרה 2.6):

$$\mathrm{Cov}[X,Y] = E\Big[\big(X-E[X]\big)\big(Y-E[Y]\big)\Big] \quad \#1$$

$$= E[XY] - E[X]E[Y] \quad \#2$$

$$\#1 = E\big[XY - XE[Y] - YE[X] + E[X]E[Y]\big]$$

$$= E[XY]$$

$$- E\big[X\,E[Y]\big] \rightarrow E[Y] = b \quad \text{קבוע} \Rightarrow E[X \cdot b] = b\,E[X] = E[Y] \cdot E[X]$$

$$- E\big[Y\,E[X]\big] \rightarrow \cancel{E[X] \cdot E[Y]}$$

$$+ E\big[1 \cdot E[X]E[Y]\big] \rightarrow \cancel{E[X] \cdot E[Y]}$$

$$= \#2$$

תכונות של שונות משותפת (תכונה 2.4):

$$\mathrm{Cov}[X,X] = Var[X]$$

$$\mathrm{Cov}[X,Y] = \mathrm{Cov}[Y,X]$$

$$\mathrm{Cov}[X,a] = 0$$

$$\mathrm{Cov}[aX,bY] = ab\,\mathrm{Cov}[X,Y]$$

$$\mathrm{Cov}[X,Y] = \mathrm{Cov}[X+a,Y+b]$$

$$\mathrm{Var}[X \pm Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] \pm 2\,\mathrm{Cov}[X,Y]$$

חיזוי לינארי

מטרה להשתמש בקשר לינארי בין המשתנים כדי ליצור חיזוי מאחד לשני.

מטרה להשתמש בקשר לינארי בין המשתנים כדי ליצור חיזוי מאחד לשני.

חיזוי לינארי (הגדרה 2.8): חיזוי מהצורה

$$\hat{Y} = ax + b,$$

כאשר x הוא מקרה פרטי של X עבורו נעשה חיזוי.

$$mse = \underbrace{E}_{\text{תוחלת/ממוצע}}\Big[\big(\underbrace{Y}_{\text{ערך "אמיתי"}} - \underbrace{\hat{Y}}_{\text{חיזוי}}\big)^{\overset{\text{ריבוע}}{2}}\Big]$$

mean square error (MSE)

שגיאה ריבועית ממוצעת

mse הוא פונקציה של $\overbrace{a,b}^{\text{המשתנים של mse}}$

$$\begin{aligned}
mse(a,b) &= E[(Y - \overbrace{(aX + b)}^{\text{חיזוי}})^2] \\
&= E[Y^2 - 2aXY + a^2X^2 - 2bY + 2abX + b^2] \\
&= E[Y^2] - 2aE[XY] + a^2E[X^2] - 2bE[Y] + 2abE[X] + b^2
\end{aligned}$$

$$\frac{\partial mse(a,b)}{\partial a} = E\left[2(Y - aX - b)(-X)\right]$$

(2.17) $$= -2E[XY] + 2aE[X^2] + 2bE[X] = 0$$

$$\frac{\partial mse(a,b)}{\partial b} = E\left[2(Y - aX - b)(-1)\right]$$

(2.18) $$= -2E[Y] + 2aE[X] + 2b = 0$$

גזירה לצורך חיפוש מינימום של הפונקציה

ניתן לרשום בצורה מטריציאונית

(2.19) $$\begin{bmatrix} E[X^2] & E[x] \\ E[X] & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} E[XY] \\ E[Y] \end{bmatrix}.$$

סיכום של תוצאת הגזירה

הפתרון המתקבל הוא

(2.20א) $$a_{opt} = \frac{E[XY] - E[X]E[Y]}{E[X^2] - E^2[X]} = \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]}$$

(2.20ב) $$b_{opt} = E[Y] - \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]}E[X]$$

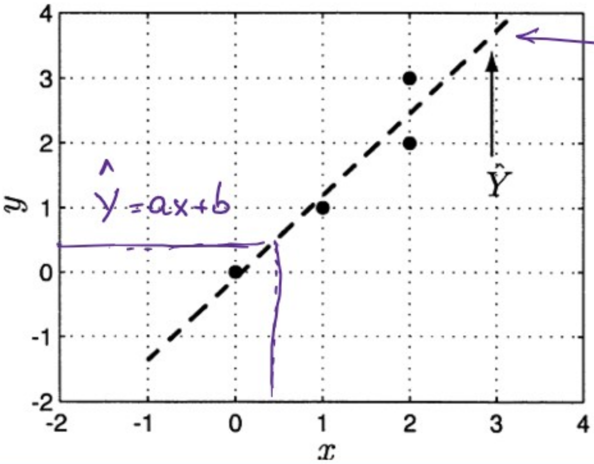
פתרון של 2 משוואות עם 2 נעלמים

לסיכום,

(2.21) $$\boxed{\begin{aligned} \hat{Y} &= a_{opt}x + b_{opt} \\ &= E[Y] + \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]}\left(x - E[X]\right) \end{aligned}}$$

הצבה בתוך נוסחה של קו לינארי

דוגמה מספרית:



מטרה: לאפיין קו אופטימלי

במובן שגיאה ריבועית

ממוצעת מינימלית (MMSE)

minimum

$X,Y = \{\underbrace{(0,0),(1,1),(2,2),(2,3)}_{\text{4 אפשרויות}}\}$

עם הסתברות זהה

PDF משותף

* התפלגות שולית:

X,Y: סה"כ 4 אפשרויות עם הסתברות זהה

$$P_X[k] = \begin{cases} \frac{1}{4} & k=0 \\ 1/4 & k=1 \\ \frac{1}{2} & k=2 \end{cases} \qquad P_Y[j] = \frac{1}{4} \quad j=0,1,2,3$$

$$\hat{Y} = a_{opt}x + b_{opt} = E[Y] + \frac{\text{Cov}[X,Y]}{\text{Var}[X]}(x - E[X])$$

$E[X]$

$E[Y]$

$Var[X] = E[X^2] - E^2[X]$

$Cov[X,Y] = E[XY] - E[X]E[Y]$ *

$$E[X] = 0\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 2\cdot\frac{1}{2} = 1\frac{1}{4} = 5/4 \qquad E[Y] = 0\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 2\cdot\frac{1}{4} + 3\cdot\frac{1}{4} = \frac{6}{4} = \frac{3}{2}$$

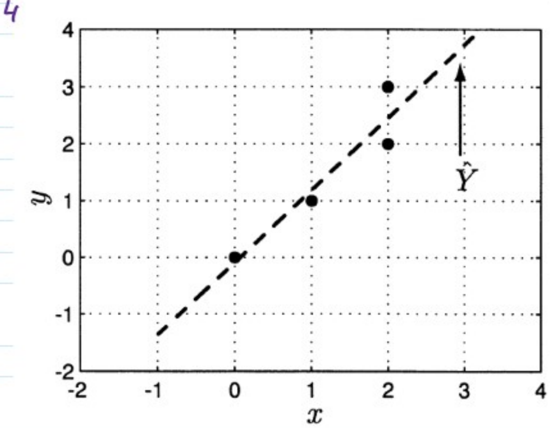
$$E[X^2] = 0^2\cdot\frac{1}{4} + 1^2\cdot\frac{1}{4} + 2^2\frac{1}{2} = \frac{9}{4} \Rightarrow Var[X] = \frac{9}{4} - \left(\frac{5}{4}\right)^2 = \frac{11}{16}$$

$$E[XY] = 0\cdot0\cdot\frac{1}{4} + 1\cdot1\cdot\frac{1}{4} + 2\cdot2\cdot\frac{1}{4} + 2\cdot3\cdot\frac{1}{4} = \frac{11}{4}$$

$$Cov[XY] = \frac{11}{4} - \frac{5}{4}\cdot\frac{3}{2} = \frac{7}{8}$$

סיכום:

$$\hat{Y} = \frac{3}{2} + \frac{7/8}{11/16}\left(x - \frac{5}{4}\right) = \frac{14}{11}x - \frac{1}{11}$$



שאלה (מתבקשת): מהי מידת קשר הליניארי בין המשתנים?

שאלה (מתבקשת): מהי מידת קשר הלינארי בין המשתנים?

הקדמה: נתון מ"מ X

* $E[X-E[X]]=0$
* $Var[bX]=b^2 Var[X]$

$b=\sqrt{Var[X]}$

משתנה מנורמל עם תוחלת 0 ושונות 1 $\Leftarrow$ $\tilde{X}=\frac{X-E[X]}{\sqrt{Var[X]}}$

בהינתן $\tilde{X}$, $\tilde{Y}$ (גם הוא מנורמל), מתקיים קשר לינארי:

$$\tilde{Y}=\underbrace{\frac{Cov[X,Y]}{Var[X]\,Var[Y]}}\tilde{X}$$

ρ_{xy} = מקדם קורלציה

$Cov[X,Y]=0$

$\Updownarrow$

* $\rho_{xy}=0$ ← אין קשר לינארי
* $\rho_{xy}=\pm1$ ← קשר לינארי
* שגיאת חיזוי לינארי הופכת לפי $(1-\rho^2_{xy})$ $|\rho_{xy}|$ יותר קרוב ל-1 $\Leftarrow$ השגיאה קטנה

$|\rho_{xy}|\le 1$

* $\rho_{xy}=Cov[X,Y]=0$ $\Leftarrow$ $\hat{Y}=E[Y]$

קשרים בין המשתנים (שמע)

אורטוגונליים (הגדרה 2.11): עבור X,Y אורטוגונליים מתקיים

$$E[XY]=0 \tag{2.27}$$

חוסרי קורלציה (הגדרה 2.12): עבור X,Y חוסרי קורלציה מתקיים

$$\mathrm{Cov}[X,Y]=\rho_{XY}=0 \tag{2.28}$$

בלתי תלויים לינארית

קשר בין אי תלות לחוסר קורלציה (הגדרה 2.13): כאשר X,Y הם בלתי תלויים, הם גם חוסרי קורלציה.

הערה 2.2 ! בלתי תלויים $\Leftarrow$ חסרי קורלציה - החץ הוא רק בכיוון אחד (בפרט למקרה פרטי של התפלגות גאוסית, בהמשך).

דוגמה למצב של תלות משמעותי,

תלות משמעותי,
אבל ללא תלות
לינארי.