הרצאה 3 - בוקר

יום ראשון 01 נובמבר 2020 13:04

דוגמה:

שחקני ליגת על של כ״סל

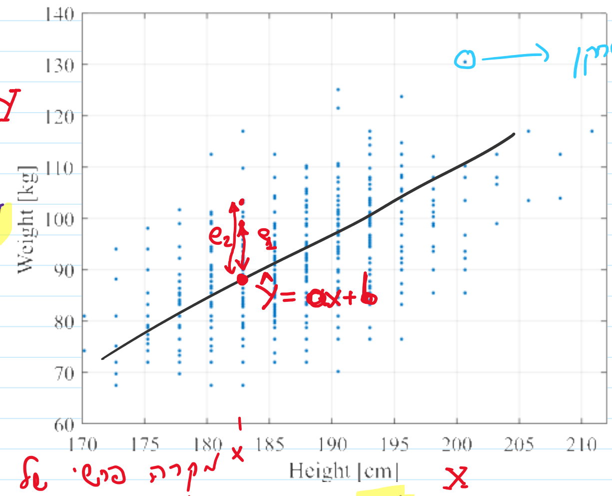
סה״כ
1034
שחקנים

תובנה חשובה:

ניתן לשים לב, כי יש "קשר" מסוים בין גובה למשקל

$\Downarrow$

לדוגמה, ניתן לחזות את המשקל ע״י ידיעת הגובה



מטרה: קשר לינארי בין זוג משתנים אקראיים

הצגת הבעיה:

א חיזוי לינארי

$$\underbrace{\hat{Y}}_{\text{חיזוי למ״א } Y} = \overset{?}{a}\underbrace{x}_{\text{מקרה פרטי של מ״א } X} + \overset{?}{b}$$

חיזוי למ״א Y
(חיזוי משקל)

מקרה פרטי של מ״א X
(בדוגמה לעיל גובה מסוים)

ב איכות/דיוק החיזוי

שגיאה ריבועית ממוצעת — mean - square error

$$mse = E\left[(\underbrace{Y}_{\text{ערך אמיתי}} - \underbrace{\hat{Y}}_{\text{חיזוי}})^2\right]$$

* a,b אופטימליים במובן שגיאה ריבועית ממוצעת מינימלית

מהם ערכי a,b אשר גורמים למינימום שגיאה?

$$mse_{(a,b)} = E\left[(Y - aX - b)^2\right]$$

$a,b \Leftarrow$ מינימום mse

* חישוב מינימום ע״י גזירה

הערכים a, b עבורם השגיאה היא מינימלית ניתן לקבל ע"י גזירה והשוואה של הנגזרת לאפס, כלהלן.

$$\begin{aligned} mse(a,b) &= \boxed{E[(Y - aX - b)^2]} \\ &= E[Y^2 - 2aXY + a^2X^2 - 2bY + 2abX + b^2] \\ \tfrac{\partial}{\partial a} &= \underbrace{E[Y^2]}_{=0} - 2aE[XY] + a^2E[X^2] - 2bE[Y] + 2abE[X] + b^2 \end{aligned} \quad (2.17)$$

תזכורת: תוחלת היא פעולה לינארית

את המינימום ניתן לקבל ע"י הגזירה, תוך שימוש בכלל השרשרת, $\left[f\left(g(x)\right)\right]' = f'\left(g(x)\right)g'(x)$,

ביטוי שבסוגר

$$\begin{aligned} \frac{\partial mse(a,b)}{\partial a} &= E\left[2(Y - aX - b)(-X)\right] \\ &= -2E[XY] + 2aE[X^2] + 2bE[X] = 0 \end{aligned} \quad (2.18)$$

$$\begin{aligned} \frac{\partial mse(a,b)}{\partial b} &= E\left[2(Y - aX - b)(-1)\right] \\ &= -2E[Y] + 2aE[X] + 2b = 0 \end{aligned} \quad (2.19)$$

פתרון: $\Rightarrow a_{opt}, b_{opt} \Rightarrow$ mse מינימלי

ניתן לרשום בצורה מטריציאונית

$$\begin{bmatrix} E[X^2] & E[x] \\ E[X] & 1 \end{bmatrix} \begin{bmatrix} a_{opt} \\ b_{opt} \end{bmatrix} = \begin{bmatrix} E[XY] \\ E[Y] \end{bmatrix}. \quad (2.20)$$

הפתרון המתקבל הוא

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

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

תשובה סופית:

$$\begin{aligned} \hat{Y} &= a_{opt}x + b_{opt} \\ &= E[Y] + \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]}\left(x - E[X]\right) \end{aligned} \quad (2.22)$$

קוד Matlab לחיזוי לינארי:

```
Ex = mean(X);
Ey = mean(Y);
Cxy = mean(X.*Y) - Ex*Ey;
Vx = var(X);
Yh = Ey+Cxy/Vx*(X-Ex);
e = Y - Yh; % error
mse = mean(e.^2);
```

לפי הגדרה $\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y]$

קו מגמה לינארי

חיזוי לינארי

רגרסיה לינארית

שגיאה ממוצעת (לא ריבועית) mean(e)

שגיאה ריבועית ממוצעת מתקבל

שגיאת חיזוי - שגיאה ממוצעת (תכונה 2.7):

$$E\left[Y - (a_{opt}X + b_{opt})\right] = 0 \quad (2.24)$$

שגיאת חיזוי - שגיאה ריבועית ממוצעת (תכונה 2.6): שגיאת החיזוי נתונה ע"י

$$mse_{min} = E\left[\left(Y - (a_{opt}X + b_{opt})\right)^2\right] = \mathrm{Var}[Y](1 - \rho_{XY}^2)$$

$$\rho_{XY} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}}$$

מקדם קורלציה (שונות משותפת מנורמלת)

* מדד לקשר לינארי בין שני משתנים אקראיים:

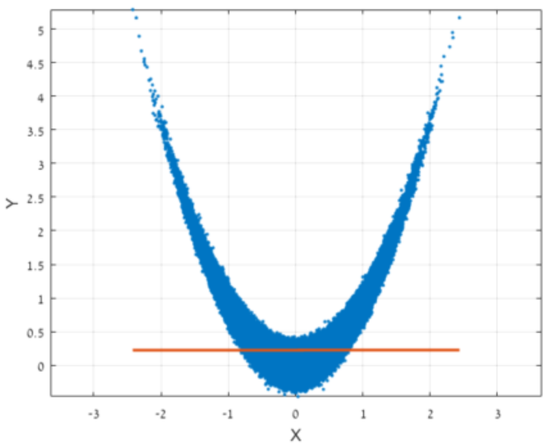
* מדד לקשר לינארי בין 2 משתנים אקראיים:

$\rho_{xy} = \pm 1$ קשר לינארי ← אין שגיאת חיזוי

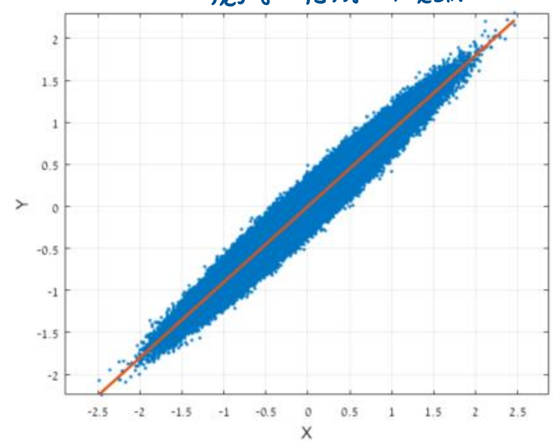
$\rho_{xy} = 0$ אין קשר לינארי ← X לא "עוזר" לדעת מהו Y

באופן כללי $|\rho| \le 1$

לא מתאים לחיזוי לינארי



מתאים לחיזוי לינארי



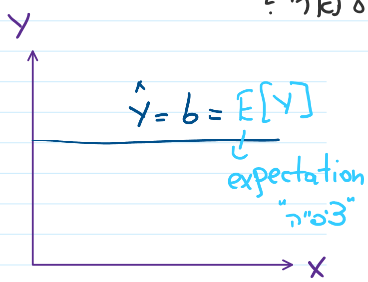
ניתוח מצב של $Cov[X,Y]=0 \Leftarrow \rho_{xy}=0$

* משמעות: חסרי תלות לינארי ⇐ $mse = Var[Y]$

* מה זה Var (שונות), E (תוחלת) בהקשר של חיזוי לינארי?

חיזוי ע"י קבוע

$a=0$



* חיזוי ע"י קבוע בלבד

$$mse(b) = E[(Y-b)^2].$$

שגיאה ריבועית ממוצעת ←

$$\begin{aligned} mse(b) &= E[(Y-b)^2] \\ &= E[Y^2 - 2bY + b^2] \\ &= E[Y^2] - 2bE[Y] + b^2 \end{aligned}$$

גזירה לצורך חיפוש מינימום ←

$$\frac{\partial mse(b)}{\partial b} = -2E[Y] + 2b = 0$$

אופטימלי - מקטין mse למינימום

ערך

$$b_{opt} = E[Y]$$

שגיאה מינימלית

$$mse_{min} = E[(Y - \underbrace{E[Y]}_{b})^2] = Var[Y].$$

$Cov[X,Y] = \rho_{xy} = 0$

* במקרה של חוסר קשר לינארי בין X,Y ← הערך של שיפוע המקבל יהיה $a = 0$

הערך של $b = E[Y]$ היא תוחלת של Y - חיזוי אופטימלי

$mse_{min} = Var[Y]$ שגיאת חיזוי במקרה זה היא שונות

דוגמה מספרית
עמוד 194 בספר
של Kay

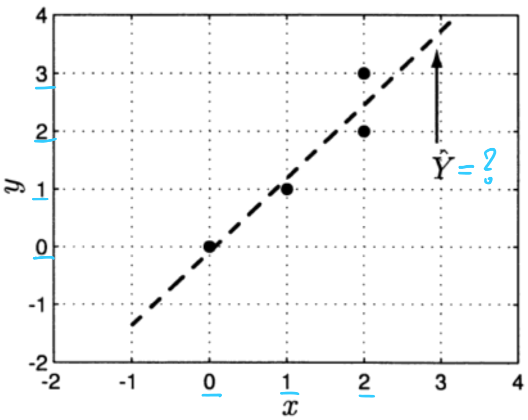
יכול לקבל אחד מתוך 4 ערכים אפשריים בהסתברות זהה

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$$

מטרה: חיזוי לינארי

שלבים:

* התפלגויות שוליות של X,Y
* $E[X], E[Y]$
  $Var[X], Cov[X,Y]$
* סיכום: משוואה של קו מגמה.



פתרון: * התפלגות שולית ← התפלגות של כל אחד בנפרד

$$p_Y[j] = \begin{cases} \frac{1}{4} & j=0 \\ \frac{1}{4} & j=1 \\ \frac{1}{4} & j=2 \\ \frac{1}{4} & j=3 \end{cases} \qquad p_X[i] = \begin{cases} \frac{1}{4} & k=0 \\ \frac{1}{4} & k=1 \\ \frac{1}{2} & k=2 \end{cases} \rightarrow \underbrace{p_{XY}[2,2] + p_{XY}[2,3]}_{\text{סכום עבור כל ערכי Y כאשר X=2}} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

4 ערכים עם הסתברות זהה

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

ערך הסתברות

$$\mathrm{Var}[X] = E[X^2] - E^2[X] = \frac{9}{4} - \left(\frac{5}{4}\right)^2 = \frac{11}{16}$$

$$\left(\frac{5}{4}\right)^2$$

$$E[X^2] = \frac{1}{4}\cdot 0^2 + \frac{1}{4}\cdot 1^2 + \frac{1}{2}\cdot 2^2 = \frac{9}{4}$$

הסתברות $(\text{ערך})^2$

$$\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y] = \frac{7}{8}$$

$$E[XY] = 0\cdot 0\cdot\frac{1}{4} + 1\cdot 1\cdot\frac{1}{4} + 2\cdot 2\cdot\frac{1}{4} + 2\cdot 3\cdot\frac{1}{4} = \frac{11}{4}$$

ערך מ-X
ערך מ-Y
הסתברות XY

$(k=2)(j=3)\underbrace{p_{XY}[2,3]}_{P[X=2,\,Y=3]}$

תזכורת: תוחלת

$$E[X] = \sum_i \overbrace{x_i}^{\text{ערך}} \overbrace{p_X[x_i]}^{\text{הסתברות של ערך}}$$

$$E[g(X)] = \sum_i \overbrace{g(x_i)}^{X^2,\ (\text{ערך})^2} p_X[x_i]$$

$$E[XY] = \sum_i \sum_k x_i y_k p_{XY}[x_i, y_k]$$

$$\boxed{\hat{Y} = E[Y] + \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]}(x - E[X]) = \frac{14}{11}x - \frac{1}{11}}$$

חישוב שגיאה: ניתן לחשב ב-2 דרכים:

1. הצבה של ערכי X וחישוב

$$E[X\cdot Y]=\sum_i\sum_k x_i y_k p_{\cancel{k}\,XY}[x_i,y_k]$$

$$E[g(X,Y)]=\sum_i\sum_k g(x_i,y_k)p_{\cancel{k}\,XY}[x_i,y_k]$$

ג. הצבה של ערכי X וחישוב

$(\text{הפרש})^2$ ה- ממוצע

$$mse=\frac{1}{4}\sum_{j=0}^{3}\left(y-\hat{y}\right)^2$$

$$=\underbrace{\frac{1}{4}}_{\text{ממוצע}}\Big[\Big(0-\Big(\frac{14}{11}\cdot 0-\frac{1}{11}\Big)\Big)^2+$$

$$+\Big(1-\Big(\frac{14}{11}\cdot 1-\frac{1}{11}\Big)\Big)^2+$$

$$+\Big(2-\Big(\frac{14}{11}\cdot 2-\frac{1}{11}\Big)\Big)^2+$$

$$+\underbrace{\Big(3-\Big(\frac{14}{11}\cdot 2-\frac{1}{11}\Big)\Big)^2}_{\text{שגיאה בריבוע}}\Big]=$$

ב. חישוב הביטוי

$$\mathrm{Var}[Y](1-\rho_{XY}^2)$$

משתנה אקראי רציף

מטרה: לאפיין ניסוי בעל תוצאה רציפה, כאן מתח.

אפיון: CDF . זהה למ"א בדיד

$$F_X(x)=P(X\le x)$$

PDF : נגזרת של CDF :

$$f_X(x)=\frac{\partial F_X(x)}{\partial x}$$

$$F_X(x)=\int_{-\infty}^{x}f_X(p)\,dp$$

תכונות :

* $\cancel{1\geq}\ f_X(x)\geqslant 0$
* $\int_{-\infty}^{\infty}f_X(x)dx=1$ סה"כ הסתברות

* $P(X=a)=0 \quad \forall a\in\mathbb{R}$

לכל a ממש.

הערה חשובה:

* $\underbrace{F_X(x)}_{\text{גם בדיד גם רציף}}\ ,\ \underbrace{P_X[x_k]}_{\text{מ"א בדיד}}$ - משמעות הסתברות

$f_X(x)$: צפיפות הסתברות - יכולה להיות לא יחידה
ומשמעו אינה הסתברות

* $p(X=a)$ ישנם אינסוף ערכים אפשריים רציפים של X ⇐ הסתברות
לערך מסוים בודד = 0

* תוחלת

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \qquad E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

* שונות: מכוסס על תוחלת, זהה למ"ש בד'ג'. $\mathrm{Var}[X] = E[X^2] - E^2[X]$

התפלגות אקספוננציאלית (מעריכית)

מטרה לאפיין תופעות עבורם התסברות משתנה באופן אקספוננציאלי (עם הזמן).

הגדרה (PDF 3.3): עבור $X \sim Exp(\lambda)$ מתקיים

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & x \geqslant 0 \\ 0 & \text{אחרת} \end{cases} \qquad (3.5)$$

λ - פרמטר של התפלגות

* **דוגמה 3.1:** חשב תוחלת של X, כאשר $X \sim Exp(\lambda)$.

פתרון:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \xrightarrow{\text{הצבה}} E[X] = \lambda \int_0^{\infty} x \exp(-\lambda x)dx = \frac{1}{\lambda}$$

$$\boxed{\int x \exp(ax)dx = \exp(ax)\left[\frac{x}{a} - \frac{1}{a^2}\right]}$$

* $\mathrm{Var}[X] = \frac{1}{\lambda^2}$

* $$F_X(x) = \int_{-\infty}^{x} f_X(p)\, dp \xrightarrow{\text{הצבה}} F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

$$= \begin{cases} 1 - \exp(-\lambda x) & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

דוגמה מספרית:

דוגמה 3.4: מודל של הופעת תקלות במערכת מסויימת מתוארת ע"י התפלגות $Exp(\lambda)$, כאשר $\lambda = \frac{1}{1000}\left[\frac{1}{\text{שעה}}\right]$. שאלות:

(א) מהו סיכוי לתקלה אחרי 1000 שעות?

(ב) מהו סיכוי לתקלה אחרי 100 שעות ולפני 1000 שעות?

(ג) מהו משך הזמן עבורו הסיכוי לתקלה הוא $\frac{1}{2}$?

ערך "חציון"

X - זמן תקלה [שעה]

סיכוי לפני 1000 שעה

$$\boxed{P(\underbrace{X < 1000}_{\text{הגדרה של CDF}}) = F_X(1000) = 1 - \exp\left(-\frac{1000}{1000}\right) \cong 0.63}$$

סיכוי אחרי 1000 שעה

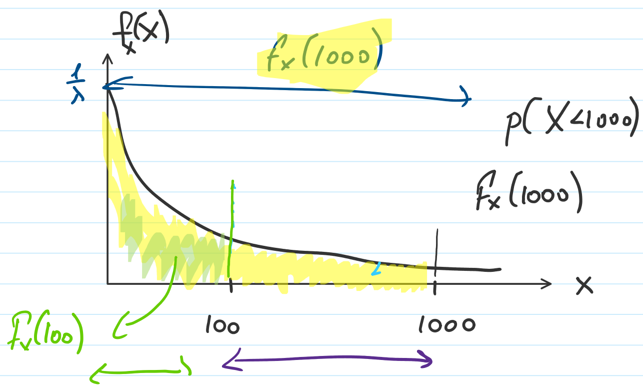
$$P(X > 1000) = 1 - P(X < 1000) = 1 - F_X(1000)$$

הגדרה של PDF $\quad F_X(x) = P(X \leq x)$

* מאחר והסתברות $P(X=x) = 0$

$$\Rightarrow P(X \leq x) = P(X < x)$$

$\Rightarrow \quad P(X \le x) = P(X < x)$



$P(X \le 1000) = 1 - P(X < 1000)$

$= 1 - F_X(1000)$

$\cong 0.36$

ג

$P(100 < X < 1000)$

$= F(1000) - F(100)$

ד

$$\Pr(X < x) = \Pr(X \leqslant x) = \frac{1}{2}$$

$$\Rightarrow F_X(x) = 1 - F_X(x) = \frac{1}{2}$$

$$\Rightarrow x = \frac{\ln\left(\frac{1}{2}\right)}{-0.001} \approxeq 693[\text{שעה}]$$