

רגרסיה ומתאםמשוואות קו הרגרסיה

סימונים:

$$S_{XX} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2$$

$$S_{YY} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n \cdot \bar{y}^2$$

$$S_{XY} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y}$$

כאשר: S_{XY} – השונות המשותפת, $\sqrt{S_{YY}}, \sqrt{S_{XX}}$ – סטיות התקן של X ו-Y בהתאמה.

בסימונים דלעיל – נוסחאות קו הרגרסיה:

קו הרגרסיה של Y כתלות ב-X: $\hat{Y} = b_{Y|X} X + a_{Y|X}$

$$b_{Y|X} = \frac{S_{XY}}{S_{XX}} = r \frac{\sqrt{S_{YY}}}{\sqrt{S_{XX}}}; \quad a_{Y|X} = \bar{Y} - b_{Y|X} \bar{X}$$

נוסחה חלופית (נוחה לעבודה):

$$b_{Y|X} = \frac{\sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2}$$

קו הרגרסיה של X כתלות ב-Y: $\hat{X} = b_{X|Y} X + a_{X|Y}$

$$b_{X|Y} = \frac{S_{XY}}{S_{YY}} = r \frac{\sqrt{S_{XX}}}{\sqrt{S_{YY}}}; \quad a_{X|Y} = \bar{X} - b_{X|Y} \bar{Y}$$

נוסחת מקדם המתאם הליניארי (של פירסון) (r):

$$r = \frac{S_{XY}}{\sqrt{S_{XX}} \sqrt{S_{YY}}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sqrt{(\sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2)(\sum_{i=1}^n y_i^2 - n \cdot \bar{y}^2)}}$$

חוזק המתאם: מתאם חלש: $0 \leq |r| < 0.4$; מתאם בינוני: $0.4 \leq |r| < 0.7$; מתאם גבוה: $0.7 \leq |r| \leq 1$

שאלה 1

נתונים ציונים של 5 תלמידים במתמטיקה (x) ובפיזיקה (y) בסולם ציונים [0-10]:

x	y
6	7
7	7
8	6
9	8
10	7

- א. חשב את מקדם המתאם בין הציונים במתמטיקה לציונים בפיזיקה.
 ב. מצא את משוואת הרגרסיה של הציון בפיזיקה על סמך הציון במתמטיקה.
 ג. מצא את משוואת הרגרסיה של הציון במתמטיקה על סמך הציון בפיזיקה.
 ד. תלמיד קיבל ציון 8.5 במתמטיקה, איזה ציון תנבא לו בפיזיקה?.

פתרון:

(א). נתון: $n = 5$

חישובי עזר:

הממוצעים: $\bar{X} = 8, \bar{Y} = 7$

$$S_{XX} = (6^2 + 7^2 + 8^2 + 9^2 + 10^2) - 5 \cdot 8^2 = 10$$

$$S_{YY} = (7^2 + 7^2 + 6^2 + 8^2 + 7^2) - 5 \cdot 7^2 = 2$$

$$S_{XY} = (6 \cdot 7 + 7 \cdot 7 + 8 \cdot 6 + 9 \cdot 8 + 10 \cdot 7) - 5 \cdot 7 \cdot 8 = 1$$

נציב בנוסחת מקדם המתאם:

$$r = \frac{S_{XY}}{\sqrt{S_{XX}} \sqrt{S_{YY}}} = \frac{1}{\sqrt{2} \cdot \sqrt{10}} = +0.224$$

כלומר, קיים מתאם לינארי חיובי (חלש) בין הציונים במתמטיקה לציונים בפיזיקה.

(ב). קו הרגרסיה המבוקש $\hat{Y} = aX + b$

$$b = \frac{S_{XY}}{S_{XX}} = \frac{1}{10} = 0.1$$

$$a = \bar{Y} - b\bar{X} = 7 - 0.1 \cdot 8 = 6.2$$

והקו המבוקש: $\hat{Y} = 0.1X + 6.2$

(ג). קו הרגרסיה המבוקש $\hat{X} = aY + b$

$$b = \frac{S_{XY}}{S_{YY}} = \frac{1}{2} = 0.5$$

$$a = \bar{X} - b\bar{Y} = 8 - 0.5 \cdot 7 = 4.5$$

והקו המבוקש: $\hat{X} = 0.5Y + 4.5$

$$\hat{Y}(8.5) = 0.1X + 6.2 = 0.1 \cdot 8.5 + 6.2 = 7.05 \quad (\text{ד})$$

שאלה 2נתון קו הרגרסיה $\hat{Y} = -2X + 5$.

- א. אם ידוע שהמוצע של X הוא 1.5, מהו הממוצע של Y ?
- ב. אם ידוע שסטיית התקן של Y גדולה פי 3 מסטיית התקן של X , מהו מקדם המתאם בין X ל- Y ?

פתרון:

א) מקו הרגרסיה נובע: $a = 5, b = -2$. נציב בנוסחה $a_{Y|X} = \bar{Y} - b_{Y|X} \bar{X}$

$$5 = \bar{Y} - (-2) \cdot 1.5 = \bar{Y} + 3 \quad \Rightarrow \quad \boxed{\bar{Y} = 2}$$

ב) נתון $\sqrt{S_{YY}} = 3\sqrt{S_{XX}}$, מכאן-

$$-2 = r \cdot \frac{\sqrt{S_{YY}}}{\sqrt{S_{XX}}} = r \cdot \frac{3\sqrt{S_{XX}}}{\sqrt{S_{XX}}} = 3r \quad \Rightarrow \quad \boxed{r = -2/3}$$

שאלה 3

חושב מקדם המתאם (r) עבור נתוני רגרסיה, שבדקו התאמה בין מבחן במתמטיקה למבחן בפיזיקה, ונמצא ש-
 $r = 0.82$. לאחר מכן ניתן פקטור של 10 נק' לכ"א מהציונים במתמטיקה והוחלט להוריד 5% מכ"א מהציונים בפיזיקה. איך ישפיע השינוי על מקדם המתאם?

פתרון:

השפעת טרנספורמציות לינאריות ($x' = a + bx; y' = c + dx$) על מקדם המתאם:

אם $b \cdot d > 0 \Rightarrow$ לא יהיה שינוי במקדם המתאם: $r' = r$.

אם $b \cdot d < 0 \Rightarrow$ השינוי במקדם המתאם יהיה: $r' = -r$.

הטרנספורמציות: $x' = x + 10, y' = 0.95 \cdot y$

מכיוון ש- $b \cdot d = 1 \cdot 0.95 > 0$, אזי לא יחול שינוי במקדם המתאם והוא יישאר כמו שהוא: $r = 0.82$.