סטטיסטיקה ורגרסיה

מצגת הסבר – מקדם מתאם וישר רגרסיה

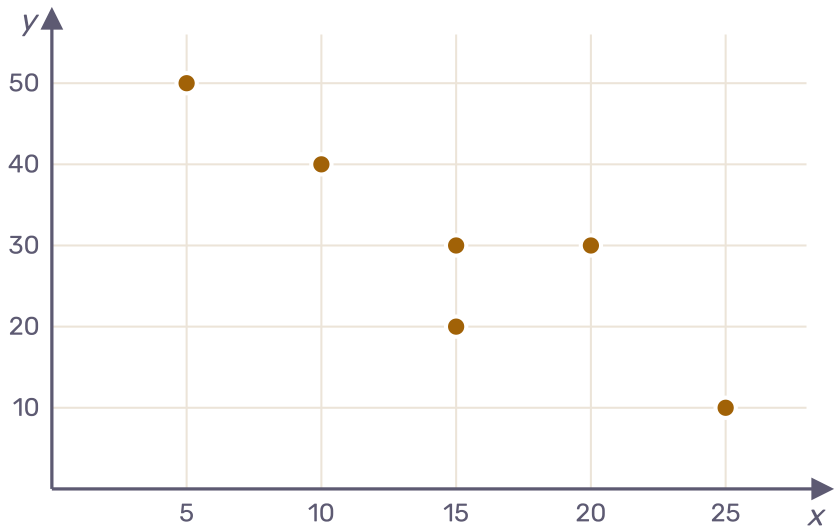
מתמטיקה 4 יח״ל | שאלון 35471 | תכנית חדשה

האתר נבנה ע״י עדי והבינה

מה נלמד?

בבחינה זו בדרך כלל **שאלה 2**. שני משתנים x ו־y שנמדדו על אותם פריטים.

1. ממוצעים $\bar{x},\ \bar{y}$ וסטיות תקן $S_x,\ S_y$
2. **מקדם המתאם** r – כמה חזק הקשר הקווי
3. **ישר הרגרסיה** – וניבוי בעזרתו
4. מה קורה כשמשנים את הנתונים



דיאגרמת פיזור מתוך בגרות קיץ תשפ"ו מועד ב'

מקדם המתאם r – מה הוא אומר?

הסימן = הכיוון

$r > 0$ ← קשר **חיובי**: x גדל – y בדרך כלל גדל.

$r < 0$ ← קשר **שלילי**: x גדל – y בדרך כלל קטן.

הגודל = העוצמה

$$-1 \le r \le 1$$

קרוב ל־1 או ל־-1 ← קשר חזק. קרוב ל־0 ← קשר חלש.

$r = 1$ או $r = -1$ ← **כל** הנקודות על קו ישר אחד.

ישר הרגרסיה עובר תמיד דרך **נקודת הממוצעים** $(\bar{x}, \bar{y})$. ניבוי לפי הישר הוא **הערכה** בלבד (כל עוד $r \neq \pm 1$).

הנוסחאות מהנוסחאון

נוסחאון • סטטיסטיקה

השכיחויות של $x_1, x_2, \ldots, x_n$ הן $f_1, f_2, \ldots, f_n$ בהתאמה.

$$N = f_1 + f_2 + \cdots + f_n$$

ממוצע:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_n f_n}{N}$$

סטיית תקן:

$$S = \sqrt{\frac{(x_1 - \bar{x})^2 f_1 + \cdots + (x_n - \bar{x})^2 f_n}{N}}$$

ציון תקן:

$$z = \frac{x - \bar{x}}{S}$$

שיפוע m של ישר הרגרסיה:

$$m = r \cdot \frac{S_y}{S_x}$$

משוואת ישר הרגרסיה:

$$y - \bar{y} = m(x - \bar{x})$$

מקדם המתאם r:

$$r = \frac{1}{N \cdot S_x \cdot S_y}\left[(x_1 - \bar{x})(y_1 - \bar{y}) + \cdots + (x_N - \bar{x})(y_N - \bar{y})\right]$$

שיטת הפתרון – טבלה אחת לכל החישובים

1. מחשבים $\bar{x}$ ו־$\bar{y}$.
2. טבלה: לכל פריט $x-\bar{x}$, $y-\bar{y}$, הריבועים והמכפלה.
3. $S_x,\ S_y$ – סכום הריבועים חלקי N, ושורש.
4. r – סכום המכפלות חלקי $N\cdot S_x\cdot S_y$.
5. $m=r\cdot\frac{S_y}{S_x}$, ומציבים בישר $y-\bar{y}=m(x-\bar{x})$.
6. ניבוי: מציבים את x הנתון בישר.

דוגמה קטנה

x	y	$x-\bar{x}$	$y-\bar{y}$	מכפלה
1	2	-1	-2	2
2	4	0	0	0
3	6	1	2	2

$\bar{x}=2,\ \bar{y}=4,\ S_x=\sqrt{\frac{2}{3}},\ S_y=\sqrt{\frac{8}{3}}$

$$r=\frac{4}{3\cdot\sqrt{\frac{2}{3}}\cdot\sqrt{\frac{8}{3}}}=\frac{4}{3\cdot\frac{4}{3}}=1$$

כל הנקודות על ישר אחד $(y=2x)\ \leftarrow\ r=1$.

מה קורה כשמשנים את הנתונים?

לזכור (לא בנוסחאון):

השינוי בכל ערכי y	$\bar{y}$	S_y	r	השיפוע m
מוסיפים מספר קבוע c	גדל ב־c	לא משתנה	לא משתנה	לא משתנה
כופלים במספר חיובי k	פי k	פי k	לא משתנה	פי k

למה?

הוספת קבוע מזיזה את כל הדיאגרמה למעלה – המרחקים בין הנקודות לא משתנים.

ומה עם הישר?

הוספת c לכל $y \leftarrow$ הישר כולו עולה ב־c: השיפוע זהה, ומספר החופשי גדל ב־c.

שאלת בגרות פתורה – דיאגרמת פיזור (א, ב)

בגרות קיץ תשפ"ו (2026) – מועד ב' • שאלה 2

x – שטח ירוק לנפש (מ"ר) בשש ערים, y – רמת זיהום האוויר. **א.** איזה היגד נכון?

היגד I: $r = -1$

היגד II: $0 < r < 1$

היגד III: $-1 < r < 0$

הנקודות **יורדות** משמאל לימין ← $r < 0$.

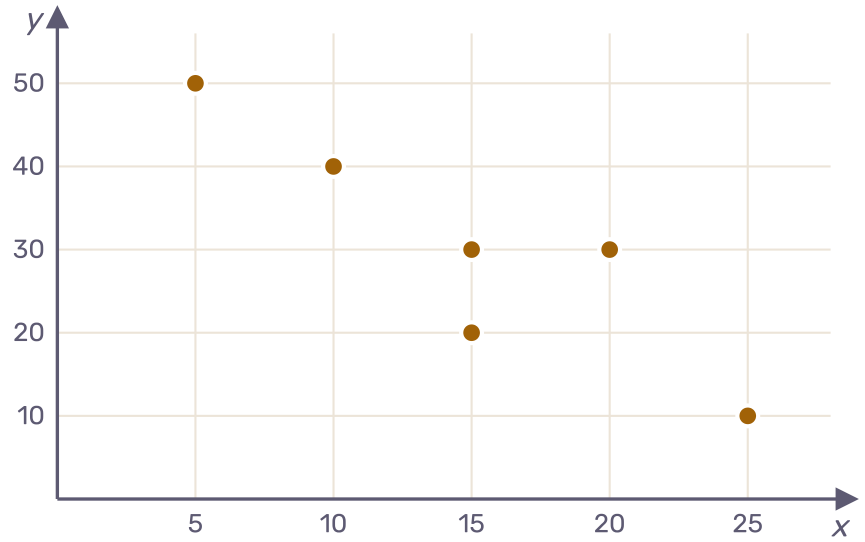
הן לא כולן על ישר אחד (ב־$x = 15$ יש שתי נקודות) ← $r \neq -1$.

היגד III: $-1 < r < 0$

ב. הממוצעים:

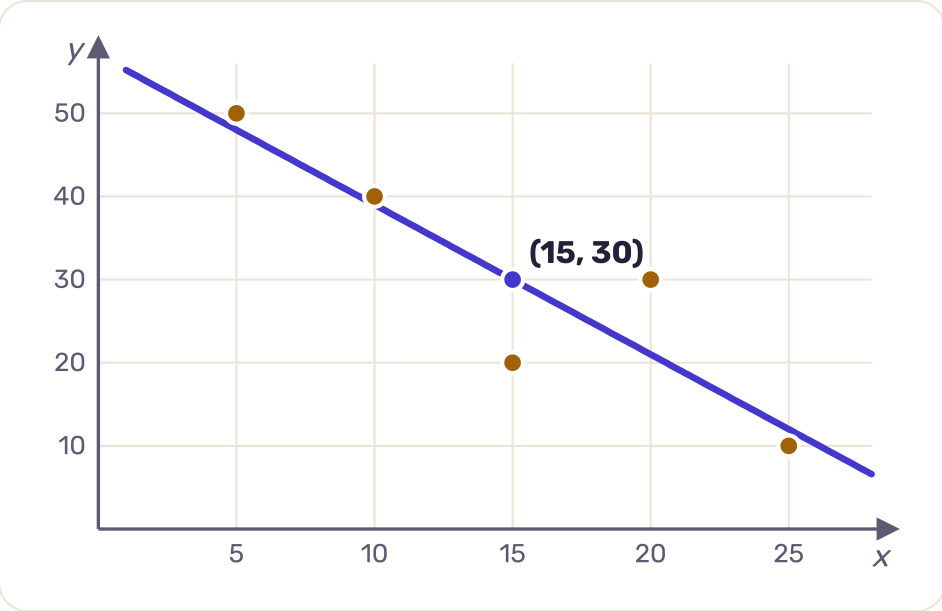
$$\bar{x} = \frac{5 + 10 + 15 + 15 + 20 + 25}{6} = \frac{90}{6} = 15$$

$$\bar{y} = \frac{50 + 40 + 30 + 20 + 30 + 10}{6} = \frac{180}{6} = 30$$



שאלת בגרות פתורה – ישר הרגרסיה (ג, ד)

בגרות קיץ תשפ"ו (2026) – מועד ב' • שאלה 2



ג1 ישר הרגרסיה הוא $y = -1.8x + b$. מצאו את b.

הישר עובר דרך נקודת הממוצעים $(15, 30)$:

$$30 = -1.8 \cdot 15 + b \;\Rightarrow\; b = 57$$

ג2 נבאו את y בעיר שבה $x = 18$:

$$y = -1.8 \cdot 18 + 57 = 24.6$$

ד סטיית התקן של x:

הריבועים:

$$(x - \bar{x})^2 :\; 100,\; 25,\; 0,\; 0,\; 25,\; 100$$

$$S_x = \sqrt{\frac{250}{6}} = \sqrt{\frac{125}{3}} \approx 6.45$$

שאלת בגרות פתורה – r מתוך השיפוע (ה, ו)

בגרות קיץ תשפ"ו (2026) – מועד ב' • שאלה 2

ה נתון $S_y = \sqrt{\frac{500}{3}}$. מצאו את r.

השיפוע הוא $m = r \cdot \frac{S_y}{S_x}$, והשיפוע הנתון -1.8:

$$\frac{S_y}{S_x} = \sqrt{\frac{500}{3} : \frac{125}{3}} = \sqrt{4} = 2$$

$$2r = -1.8 \;\Rightarrow\; r = -0.9$$

ו התברר שכל ערכי y גדולים ב־4. האם השיפוע ישתנה? מהו הישר החדש?

הוספת 4 לכל y: S_y ו־r לא משתנים ← **השיפוע לא משתנה.**

הממוצעים החדשים:

$$\bar{x} = 15, \quad \bar{y} = 30 + 4 = 34$$

$$34 = -1.8 \cdot 15 + b \;\Rightarrow\; b = 61$$

$$y = -1.8x + 61$$

עוד שאלת בגרות – חישוב r בטבלה (א, ב)

בגרות קיץ תשפ"ג (2023) – מועד ב' • שאלה 2

נמדדו 8 דגמים.
x: ארבעה ערכים 9, וארבעה ערכים 11.
y: הערכים 2 ו־18, ושישה ערכים 10.

$$\bar{x} = \frac{4 \cdot 9 + 4 \cdot 11}{8} = 10$$

כל x רחוק 1 מהממוצע:

$$S_x = \sqrt{\frac{8 \cdot 1}{8}} = 1$$

$$\bar{y} = \frac{2 + 6 \cdot 10 + 18}{8} = 10$$

$$S_y = \sqrt{\frac{64 + 0 + 64}{8}} = \sqrt{16} = 4$$

x	y	$x - \bar{x}$	$y - \bar{y}$	מכפלה
9	2	-1	-8	8
9 (3 פעמים)	10	-1	0	0
11 (3 פעמים)	10	1	0	0
11	18	1	8	8

$$r = \frac{8 + 8}{8 \cdot 1 \cdot 4} = \frac{16}{32} = 0.5$$

עוד שאלת בגרות – הישר והשינוי (ג, ד, ה)

בגרות קיץ תשפ"ג (2023) – מועד ב' • שאלה 2

ג ישר הרגרסיה לניבוי y לפי x:

$$y-10=\frac{1}{2}\cdot\frac{4}{1}(x-10)$$
$$y=2x-10$$

ד ניבוי ל־$x=10$:

$$y=2\cdot 10-10=10$$

צפוי: $x=10$ הוא הממוצע, והישר עובר דרך $(10,10)$.

ה כל הזמנים y קוצרו ב־20% (כלומר כפול 0.8). מה קורה ל־r, ל־S_y ולשיפוע?

r – **לא משתנה** (0.8 מצטמצם בנוסחה).

$$S_y=0.8\cdot 4=3.2$$
$$m=\frac{1}{2}\cdot\frac{3.2}{1}=1.6$$

סטיית התקן והשיפוע **קטנים**.

שימו לב – הטעויות הנפוצות

1. מחלקים ב־N – מספר **הפריטים** (הערים, הדגמים), לא במספר העמודות.
2. בסטיית תקן לא לשכוח את **השורש**.
3. r תמיד בין -1 ל־1. יצא מחוץ לתחום? יש טעות.
4. הסימן של r = הסימן של השיפוע.
5. „למצוא b" ← מציבים את $(\bar{x}, \bar{y})$ בישר.
6. ניבוי הוא הערכה. „בדיוק" רק כאשר $r = \pm 1$.

בדיקה עצמית

x	1	2	3	4	5
y	3	5	4	7	6

1. חשבו את $\bar{x},\ \bar{y},\ S_x,\ S_y$.
2. חשבו את r.
3. מצאו את ישר הרגרסיה, ונבאו את y כאשר $x = 6$.
4. כל ערכי y גדלו ב־3. מה יקרה ל־r ולשיפוע?

התשובות (לפתוח אחרי שפתרתם)

1. $\bar{x} = \mathbf{3},\ \bar{y} = \mathbf{5}$

$$S_x = \sqrt{\tfrac{10}{5}} = \sqrt{\mathbf{2}},\quad S_y = \sqrt{\tfrac{10}{5}} = \sqrt{\mathbf{2}}$$

2. סכום המכפלות 8

$$r = \frac{8}{5 \cdot \sqrt{2} \cdot \sqrt{2}} = \mathbf{0.8}$$

3.

$$m = 0.8 \cdot \frac{\sqrt{2}}{\sqrt{2}} = 0.8$$

$$y - 5 = 0.8(x - 3) \ \Rightarrow\ \mathbf{y = 0.8x + 2.6}$$

$$y = 0.8 \cdot 6 + 2.6 = \mathbf{7.4}$$

4. שניהם **לא משתנים** (רק הישר עולה ב־3)