סטטיסטיקה ורגרסיה – כל השאלות עם תשובות סופיות

הכנה לבגרות במתמטיקה – 4 יחידות, כיתה י״ב (שאלון 472) • 8 שאלות

בגרות קיץ תשפ״ג (2023) – מועד ב׳ • שאלה 2

2. מרצה באוניברסיטה רצתה לבדוק אם היעדרות משיעורים בקורס שנתי (המשתנה x) קשורה לינארית לציון במבחן הסופי (המשתנה y).

נתון: ממוצע ההיעדרות מן השיעורים היה $\bar{x} = 10$, הציון הממוצע היה $\bar{y} = 70$, ומקדם המתאם היה שלילי ($r < 0$).

המרצה מצאה את משוואת ישר הרגרסיה לניבוי y מ־x.

א. אחת מן המשוואות 1–4 שלפניכם היא המשוואה שמצאה המרצה. קבעו איזו מהן היא המשוואה, ונמקו את קביעתכם.

1. $y = 10x - 30$
2. $y = x + 60$
3. $y = -2x + 70$
4. $y = -2x + 90$

המרצה חישבה את סטיות התקן בעבור ההיעדרות מן השיעורים ובעבור הציונים במבחן הסופי, וקיבלה: $Sx = 4$, $Sy = 10$.

ב. חשבו את מקדם המתאם r.

ג. מהו מספר ההיעדרויות שישר הרגרסיה מנבא בעבורו ציון 80?

דוד, מרצה אחר, רצה לערוך את אותה בדיקה בנוגע לתלמידיו. הוא מצא את משוואת ישר הרגרסיה לניבוי y מ־x, וגילה כי בעבור כל מספר של היעדרויות – הישר מנבא תמיד את הציון 65.

ד. מצאו את ערכו של כל אחד מן המדדים שבתת־סעיפים (1)–(2) בעבור התלמידים של דוד:

(1) שיפוע ישר הרגרסיה.

(2) הציון הממוצע ($\bar{y}$).

ה. על פי הנתונים שבשאלה, האם אפשר למצוא את הממוצע של היעדרויות התלמידים של דוד? נמקו את תשובתכם.

תשובות סופיות

סעיף א

משוואה 4: $y = -2x + 90$

סעיף ב

$r = -0.8$

סעיף ג

5 היעדרויות

סעיף ד (1)

$m = 0$

סעיף ד (2)

$\bar{y} = 65$

סעיף ה

לא – הישר $y = 65$ אופקי ועובר ב־$(\bar{x}, 65)$ לכל $\bar{x}$.

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

2. חקלאית בעלת מטע עצי שזיף התעניינה בקשר הלינארי שבין קוטר השזיפים (המשתנה x) ובין משקלם (המשתנה y).

היא בחרה באקראי 4 שזיפים, מדדה את הקוטר (במילימטרים) של כל אחד מהם, ומדדה את המשקל (בגרמים) של כל אחד מהם.

נמצא כי $Sy < Sx$ (סטיית התקן של משתנה הקוטר גדולה מסטיית התקן של משתנה המשקל),

וכי מקדם המתאם הוא $r = 0.8$.

לאחר מכן חישבה החקלאית את ישר הרגרסיה לניבוי משקלם של השזיפים לפי קוטרם.

א. **(1)** הביעו באמצעות Sx ו־Sy את שיפוע ישר הרגרסיה.

(2) האם השיפוע של ישר הרגרסיה גדול מ־1 או קטן מ־1? נמקו את תשובתכם.

נתון כי ישר הרגרסיה לניבוי משקל השזיפים לפי קוטרם הוא $y = \frac{3}{4}x + 15$, וכי $\bar{x} = 80$.

לפניכם טבלה, ובה נתונים על המשקל של שלושה מתוך ארבעת השזיפים שנבחרו.

שזיף	המשקל (y)
א	70
ב	70
ג	80
ד	?

ב. **(1)** מצאו את המשקל של שזיף ד.

(2) מצאו את סטיית התקן של משקל השזיפים.

(3) מצאו את סטיית התקן של קוטר השזיפים.

תשובות סופיות

סעיף א (1)

$m = 0.8 \cdot \frac{S_y}{S_x}$

סעיף א (2)

השיפוע קטן מ־1 (כי $m < 0.8$).

סעיף ב (1)

המשקל של שזיף ד: 80 גרם

סעיף ב (2)

$S_y = 5$ גרם

סעיף ב (3)

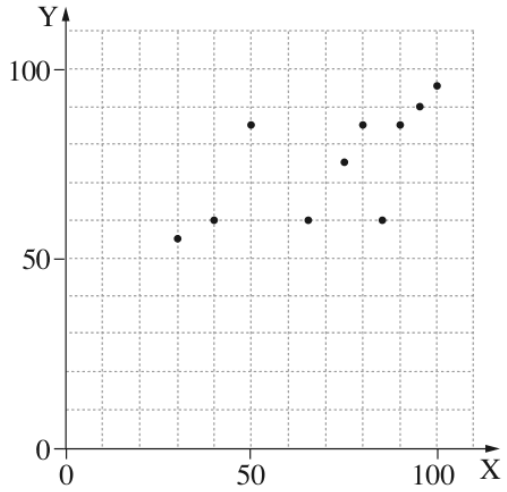
$S_x = 5\frac{1}{3} \approx 5.33$ מילימטר

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

2. תלמידי כיתה י"ב התבקשו לכתוב עבודה ולהגישה. הציון שקיבלו על העבודה שוקלל בציון הסופי של כל תלמיד.

המורה רצתה לבדוק את הקשר בין הציון על העבודה ובין הציון הסופי, ולשם כך סרטטה את דיאגרמת הפיזור של שני הציונים:

X – הציון על העבודה, Y – הציון הסופי. הדיאגרמה שהתקבלה מתוארת בתרשים שלפניכם.



א. האם אפשר להסיק מן הדיאגרמה הנתונה שכל תלמיד שקיבל על העבודה ציון גבוה יותר מתלמיד אחר קיבל **בהכרח** ציון סופי גבוה יותר מן התלמיד האחר? נמקו.

ב. אחד מן המספרים שלפניכם הוא מקדם המתאם המתאים לקשר בין שני המשתנים.

קבעו מיהו מבין המספרים האלה: 0.675 , 0 , 0.999 , -0.8 , 1.6 .

נתונים הממוצעים וסטיות התקן של שני המשתנים: $\overline{X}=71$, $S_X=23$, $\overline{Y}=75$, $S_Y=14$.

ג. מצאו את משוואת ישר הרגרסיה לניבוי הציון הסופי על פי הציון על העבודה.

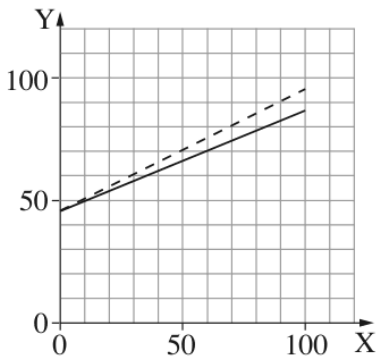
הוחלט להעלות את הציון הסופי של כל תלמיד ב־ 5 נקודות, ובעקבות העלאה זו התקבל ישר רגרסיה חדש.

ד. **(1)** האם השתנתה סטיית התקן S_Y לאחר העלאת הציונים?

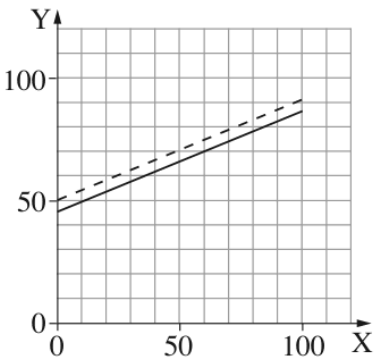
(2) אחד מן הגרפים I–III שבסוף השאלה מייצג את הישר הישן, שלפני העלאת הציון הסופי (מסורטט בקו מלא), ואת הישר החדש, שאחרי ההעלאה (מסורטט בקו מקווקו).

קבעו מיהו הגרף, ונמקו.

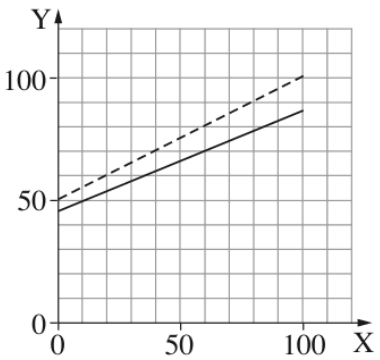
ה. אם קיים בכיתה תלמיד שהציון שלו על העבודה הוא 71 , מה צריך להיות הציון הסופי שלו (לאחר ההעלאה בציון) כדי שהנקודה המייצגת את שני הציונים שלו תהיה על ישר הרגרסיה החדש?



III



II



I

תשובות סופיות

סעיף א
לא
למשל: תלמיד עם 50 על העבודה קיבל 85 בציון הסופי, ותלמיד עם 65 קיבל רק 60.

סעיף ב
$r = 0.675$

סעיף ג
$y = 0.411x + 45.83$ (בעיגול השיפוע ל־0.41: $y = 0.41x + 45.89$)

סעיף ד (1)
לא, S_y נשארת 14.

סעיף ד (2)
גרף II – השיפוע לא משתנה, הישר מוזז 5 יחידות למעלה.

סעיף ה
80

בגרות חורף תשפ"ג (2023) • שאלה 2

2. במחקר שנועד לבדוק את ההשפעה של קורס לגמילה מעישון, השתתפו 9 מעשנים שנכחו בקורס.

בעבור כל אחד מהם נמדדה צריכת הסיגריות היומית לפני הקורס (X), וצריכת הסיגריות היומית לאחר שבוע מתחילת הקורס (Y).

הנתונים שהתקבלו מפורטים בטבלה שלפניכם:

צריכת הסיגריות היומית לפני הקורס – X	5	15	20	22	22	25	28	30	40
צריכת הסיגריות היומית לאחר שבוע מתחילת הקורס – Y	3	10	19	20	22	22	24	30	30

א. חשבו את ממוצע צריכת הסיגריות היומית למשתתף במחקר לפני הקורס ואת ממוצע צריכת הסיגריות היומית למשתתף במחקר לאחר שבוע מתחילת הקורס.

ב. סרטטו דיאגרמת פיזור של Y כתלות ב־ X (כל משבצת במחברת מייצגת 2 סיגריות ליום).

ג. על פי הדיאגרמה שסרטטתם, בחרו במקדם המתאם המתאים לנתונים מבין המספרים שלפניכם, ורשמו אותו:

$r = 1$, $r = 0.949$, $r = 0.212$, $r = -0.633$.

נתונות סטיות התקן $S_X = 9.226$, $S_Y = 8.26$.

ד. מצאו את משוואת ישר הרגרסיה לניבוי Y על פי X .

ה. אלעד מעשן 21 סיגריות ביום. הוא רוצה להשתתף בקורס לגמילה מעישון. על פי ישר הרגרסיה שמצאתם, מהו הניבוי לצריכת הסיגריות היומית שלו לאחר שבוע מתחילת הקורס?

תשובות סופיות

סעיף א

לפני הקורס: 23 סיגריות ביום
אחרי שבוע: 20 סיגריות ביום.

סעיף ב

ראו דיאגרמה.

סעיף ג

$$r = 0.949$$

סעיף ד

$$y = 0.85x + 0.45$$

סעיף ה

כ־18.3 סיגריות ביום (כ־18).

בגרות קיץ תשפ"ב (2022) – מועד ב' • שאלה 2

2. מנהל חברה בדק את הקשר בין ההוצאה החודשית של החברה על פרסום מוצריה, ובין ההכנסות מן המכירות שלה בחודש שלאחר מכן. הוא בדק נתונים מכמה חודשים רצופים.

בטבלה שלפניכם מוצגים הנתונים על ההוצאות וההכנסות באלפי שקלים:

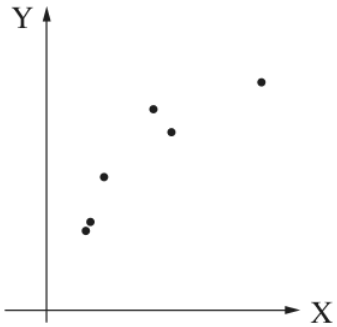
הוצאה על פרסום (X)	9	10	13	24	28	48
הכנסות ממכירות בחודש שלאחר מכן (Y)	180	200	300	450	400	510

המנהל חישב ומצא כי ממוצע ההוצאות לחודש על פרסום הוא $\overline{X} = 22$ וסטיית התקן היא $Sx = 13.6$.

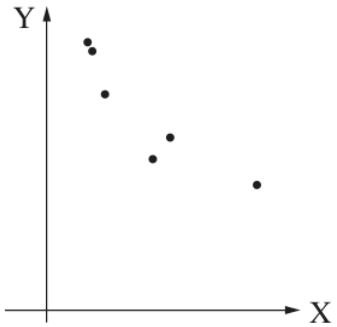
ממוצע ההכנסות ממכירות לחודש הוא $\overline{Y} = 340$ וסטיית התקן היא $Sy = 123.4$.

לפניכם שתי דיאגרמות פיזור (1)–(2). אחת מהן מתארת את הקשר בין שני המשתנים (X ו־ Y).

דיאגרמה 1



דיאגרמה 2



א. מבין שתי הדיאגרמות (1)–(2), מהי הדיאגרמה המתארת את הקשר בין שני המשתנים? נמקו את תשובתכם.

נתונים ארבעה מקדמי מתאם: $r = -0.7$, $r = -0.9$, $r = 1$, $r = 0.9$.

ב. אחד מארבעת מקדמי המתאם מתאים לנתונים. קבעו מיהו ונמקו את קביעתכם.

ג. מהי משוואת ישר הרגרסייה לניבוי ההכנסות ממכירות כתלות בהוצאה על פרסום?

ד. על פי ישר הרגרסייה שמצאתם, מהי ההערכה להכנסות ממכירות (באלפי שקלים) בעבור הוצאה של 19,000 שקלים לחודש על פרסום? נמקו את תשובתכם.

החברה המירה את ההוצאות וההכנסות בשקלים להוצאות והכנסות בדולרים (ולכן כל המספרים בטבלה קטנו בערך פי 3).

ה. מהי ההשפעה של המרת השקלים לדולרים על כל אחד מן הגדלים (1)–(3) שלפניכם (כלומר האם הוא גדל, קטן, או לא השתנה)?

נמקו את תשובותיכם.

(1) $\overline{X}$

(2) Sx

(3) r

תשובות סופיות

סעיף א

דיאגרמה 1 – הקשר חיובי (כש־X גדל, Y גדל בדרך כלל).

סעיף ב

$r = 0.9$

סעיף ג

$y = 8.166x + 160.348$

סעיף ד

כ־315.5 אלפי שקלים (כ־315,500 ש"ח)

סעיף ה (1)

$\bar{X}$ קטן (בערך פי 3).

סעיף ה (2)

S_x קטנה (בערך פי 3).

סעיף ה (3)

r לא משתנה.

בגרות קיץ תשפ"ב (2022) – מועד א' • שאלה 2

2. כדי לבדוק את הקשר בין ציוני בחינת הבגרות במתמטיקה (X) ובין ציוני בחינת הבגרות במדעי המחשב (Y), חושבו בשנה מסוימת הממוצעים, סטיות התקן ומקדם המתאם של הציונים בבחינות האלה. תוצאות החישובים היו:

$\overline{X} = 64$, $\overline{Y} = 72$, $s_x = 14$, $s_y = 9$, $r = 0.77$.

א. מצאו את משוואת ישר הרגרסיה לניבוי הציון במדעי המחשב על פי הציון במתמטיקה.

באותה השנה דני נבחן במתמטיקה ובמדעי המחשב, והציון שקיבל במתמטיקה היה 90 .

ב. על פי ישר הרגרסיה שמצאתם, מהו הציון המשוער של דני בבחינת הבגרות במדעי המחשב?

לאחר שהתברר שהבחינה במתמטיקה באותה השנה הייתה קשה, הוסיפו ל**כל אחד** מן הציונים במתמטיקה 4 נקודות.

ג. מהי ההשפעה של התוספת לציונים על כל אחד מן הגדלים (1)–(3) שלפניכם (כלומר, האם הוא גדל, קטן או לא השתנה)?

נמקו את תשובותיכם.

(1) $\overline{X}$

(2) s_x

(3) r

בשנה שאחרי כן התקבלו בבחינות במתמטיקה ובמדעי המחשב אותם הממוצעים ואותן סטיות התקן כמפורט בתחילת השאלה.

ידוע שמשוואת ישר הרגרסיה של הציונים בשנה זו הייתה: $y = mx + 43.2$.

ד. **(1)** מצאו את הערך של m .

(2) מצאו את מקדם המתאם של הציונים בשנה זו.

תשובות סופיות

סעיף א

$y = 0.495x + 40.32$

סעיף ב

הציון המשוער: $84.87 \approx 85$

סעיף ג (1)

גדל ב־4: $\bar{x} = 68$

סעיף ג (2)

לא השתנתה: $S_x = 14$

סעיף ג (3)

לא השתנה: $r = 0.77$

סעיף ד (1)

$m = 0.45$

סעיף ד (2)

$r = 0.7$

בגרות חורף נבצרים תשפ"ב (2022) • שאלה 2

2. במחקר על הקשר בין טמפרטורת המים X לבין כמות המדוזות Y , נאספו 30 זוגות נתונים והתקבלו התוצאות הבאות: $\bar{X} = 26.5^\circ$, $S_X = 3.5^\circ$, $\bar{Y} = 14$, $S_Y = 2.5$

נערך ניבוי בעזרת ישר רגרסיה לכמות המדוזות לפי טמפרטורת המים.

א. מה יהיה הניבוי לכמות המדוזות כאשר טמפרטורת המים היא הטמפרטורה הממוצעת 26.5° ?

נתון בנוסף כי מקדם המתאם הוא $r = 0.8$.

ב. לפניך שלושה היגדים (1)–(3). קבע אלו מההיגדים נכונים. נמק.

(1) מקדם המתאם לא מאפשר ניבוי טוב כי הוא קטן מ- 1 .

(2) ככל שהטמפרטורה יורדת, ננבא כמות גדולה יותר של מדוזות.

(3) ככל שהטמפרטורה עולה, ננבא כמות גדולה יותר של מדוזות.

ג. מצא את ישר הרגרסיה לניבוי כמות המדוזות לפי הטמפרטורה.

ד. על פי ישר הרגרסיה, מה יהיה הניבוי לכמות המדוזות ביום שבו טמפרטורת המים היא 33° ?

תשובות סופיות

סעיף א

14 מדוזות

סעיף ב

רק היגד (3) נכון.

סעיף ג

$y = \frac{4}{7}x - 1\frac{1}{7} \approx 0.571x - 1.143$

סעיף ד

$17\frac{5}{7} \approx 17.71$, כלומר כ־18 מדוזות.

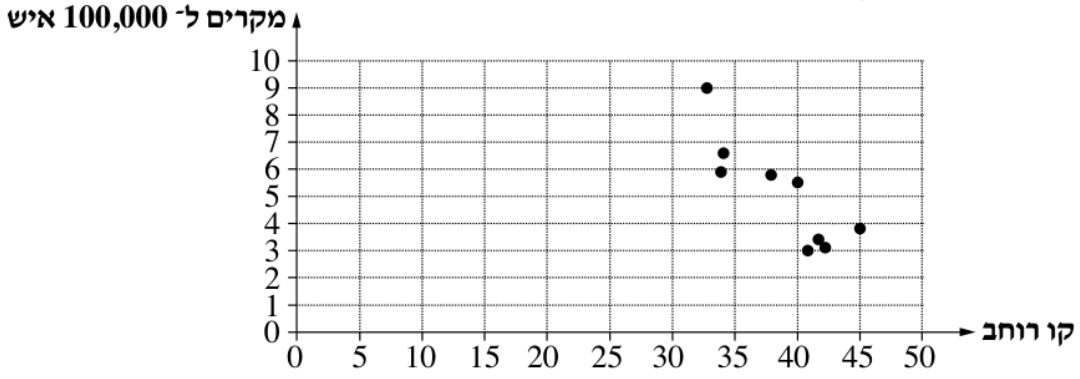
בגרות חורף תשפ"ב (2022) • שאלה 2

2. חוקרים בדקו אם יש קשר בין קו הרוחב שבו נמצא אזור מסוים ובין שיעור מקרי סרטן העור (מלנומה) באוכלוסייה באותו אזור. המחקר נערך בתשעה אזורים שונים בארצות הברית, במשך שלוש שנים. בכל אחד מן האזורים נבדקו 100,000 תושבים. נתוני המחקר מוצגים בטבלה שלפניך.

קו רוחב – x	**שיעור מקרי המלנומה – y** (מקרים ל־ 100,000)	
32.8	9	
33.9	5.9	
34.1	6.6	
37.9	5.8	
40.0	5.5	
40.8	3.0	
41.7	3.4	
42.2	3.1	
45.0	3.8	
$\bar{x} = 38.71$	$\bar{y} = 5.12$	ממוצע
$s_x = 4.04$	s_y	סטיית תקן

א. הראה כי סטיית התקן של שיעור מקרי המלנומה היא $s_y = 1.88$.

לפניך דיאגרמת הפיזור של y כתלות ב־ x .



ב. אחד מבין מקדמי המתאם (1)–(4) שלפניך מייצג את הקשר בין הנתונים. קבע איזה מהם מייצג את הקשר, ונמק את קביעתך.

(1) $r = 1$

(2) $r = -0.857$

(3) $r = 0.651$

(4) $r = -1$

ג. מצא את משוואת קו הרגרסיה לניבוי מקרי המלנומה y על פי קו הרוחב x .

ד. מהו הניבוי לשיעור מקרי המלנומה y בקו רוחב 36 ?

תשובות סופיות

סעיף א

$S_y \approx 1.88$ (מקרים ל־100,000 תושבים)

סעיף ב

$r = -0.857$ (2)

סעיף ג

$y = -0.399x + 20.57$

סעיף ד

בערך 6.2 מקרים ל־100,000 תושבים ($y \approx 6.21$)

הפתרונות נכתבו בעזרת בינה מלאכותית ונבדקו, אך ייתכנו טעויות – ללא אחריות. מומלץ להשוות לפתרון הרשמי.