סטטיסטיקה בניהול



תוכן העניינים

1. סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה 1
2. סטטיסטיקה תיאורית- הצגה של נתונים 6
3. סטטיסטיקה תיאורית- סכימה 14
4. סטטיסטיקה תיאורית - מדדי מיקום מרכזי 18
5. סטטיסטיקה תיאורית - מדדי פיזור - הטווח, השונות וסטיית התקן 25
6. סטטיסטיקה תיאורית - מדדי פיזור - טווח בין רבעוני 28
7. סטטיסטיקה תיאורית- מדדי מיקום יחסי-ציון תקן 32
8. סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה 34
9. סטטיסטיקה תיאורית - טרנספורמציה לינארית 36
10. סטטיסטיקה תיאורית שאלות אמריקאיות- מסכמות 38
11. מדדי קשר - מדד הקשר של קרמר 44
12. מדד הקשר ספירמן 46
13. מקדם המתאם הלינארי (פירסון) 49
14. רגרסיה 65
15. מדדי קשר - תרגול בחירת מדד מתאים 68

סטטיסטיקה בניהול

פרק 1 - סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה

תוכן העניינים

1. סולמות מדידה 1

סטטיסטיקה תיאורית – סיווג משתנים וסולמות מדידה:

רקע:

סטטיסטיקה תיאורית הוא ענף בו לומדים כיצד לאסוף נתונים, להציג אותם ולנתח אותם.

בסטטיסטיקה תיאורית אנו פונים לקבוצה מסוימת. באותה קבוצה אנו אוספים נתונים על הישויות באותה קבוצה.

משתנה – תכונה שיכולה לקבל מספר ערכים : דעה פוליטית, מקום מגורים, גובה של אדם וכדומה.

כל ישות בקבוצה שאנו צופים בה ואוספים לגביה נתונים נקראת תצפית.

הנתונים שנאספים בדרך כלל מרוכזים במסד נתונים. במסד הנתונים כל שורה היא תצפית וכל עמודה מייצגת משתנה.

דוגמה (פתרון בהקלטה):

למחלקת טראומה הגיעו 5 פצועים מתאונה שקרתה בכביש החוף.
אספו נתונים לגבי אותם פצועים, הנתונים מרוכזים בטבלה הבאה :

מין	גיל	מצב הפצוע	דופק
גבר	26.6	אנוש	40
גבר	24.5	קשה	38
אישה	32.1	קשה	50
גבר	34.9	בינוני	65
אישה	23.1	קל	89

ענו על השאלות הבאות :
הגדירו את הקבוצה שבדוגמה.
כמה תצפיות בקבוצה?
כמה משתנים בקבוצה?
כמה ערכים יש למשתנה "מין"?

© כל הזכויות שמורות לברק קנדל

את המשתנים במחקר אנו מסווגים ל-"סולמות מדידה" הדבר חשוב לדרך שבה ננתח את הנתונים בהמשך.

מיון משתנים לפי סולמות המדידה:

1. סולם שמי (nominal) – משתנה שלערכיו יש משמעות רק מבחינת הזהות ואין עניין של יותר או פחות לערכים שלהם לדוגמה: צבע מועדף.

   משתנה דיכוטומי (הינו מסולם שמי) אותם משתנים שיש להם רק שני ערכים אפשריים. למשל: אזרחות ישראלית: יש או אין.

2. סולם סדר (ordinal) – כאשר לערכים של המשתנה בנוסף לשם ישנה גם משמעות לסדר מי יותר או מי פחות אבל אין משמעות לגודל.
   משתנה מסולם סדר יכול לקבל ערכים מילוליים או מספריים.
   למשל, דרגה בצבא.

3. סולם כמותי (scale) – משתנה שחייב להיות מספרי, לערכים שלו בנוסף לשם ולסדר בניהם יש משמעות לערך המספרי. משתנה כמותי הוא משתנה שניתן בדרך כלל לספור או למדוד על ידי מכשיר מדידה. למשל, מספר המחשבים בדירה, שטח הדירה במ"ר.

את המשתנה הכמותי אנו מסווגים לשני סוגים:

משתנה בדיד:
משתנה שערכיו מתקבלים מתוך סידרה של ערכים אפשריים כמו: מספר המחשבים בדירה.

משתנה רציף:
משתנה שערכיו מתקבלים מתוך אינסוף ערכים בתחום מסוים, הערכים מתקבלים ברצף וללא קפיצות של ערכים. למשל, שטח הדירה במ"ר.

דוגמה (פתרון בהקלטה):

למחלקת טראומה הגיעו 5 פצועים מתאונה שקרתה בכביש החוף.
אספו נתונים לגבי אותם פצועים, הנתונים מרוכזים בטבלה הבאה:

מין	גיל	מצב הפצוע	דופק
גבר	26.6	אנוש	40
גבר	24.5	קשה	38
אישה	32.1	קשה	50
גבר	34.9	בינוני	65
אישה	23.1	קל	89

סווגו כל משתנה במסד הנתונים: שמי, סדר, כמותי רציף, כמותי בדיד.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) לפניכם טבלה המסכמת נתונים לגבי סקר שנעשה היום:

שם משפחה	מצב משפחתי	מידת דתיות	מספר רכבים
כהן	רווק	חילוני	0
חדד	נשוי	חילוני	1
לביא	גרוש	מסורתי	1
פיינגולד	אלמן	חילוני	2
אבו שוקרא	נשוי	דתי	1
בן חיים	נשוי	מסורתי	0
רוטשילד	רווק	חילוני	0

א. כמה תצפיות בדוגמה זו?

ב. כמה משתנים בדוגמה זו?

ג. כמה ערכים ישנם ל-"מידת דתיות"?

ד. מהם הערכים האפשריים למשתנה "מצב משפחתי"?

2) סניף מספר 543 של בנק "רווה" בדק ל-80 לקוחות את מספר הפעמים שכל לקוח נכנס לסניף הבנק במשך שבוע. התוצאות שהתקבלו הן: 50 אנשים נכנסו 0 פעמים לסניף, 20 אנשים נכנסו פעם אחת לסניף, 5 אנשים נכנסו פעמיים לסניף, 5 אנשים נכנסו יותר מפעמיים.

א. הגדירו את הקבוצה בדוגמה זו.

ב. כמה תצפיות בדוגמה זו?

ג. הגדירו את המשתנה בדוגמה זו. מהו סולם המדידה שלו?

3) במחקר רפואי התעניינו לדעת כיצד מינון תרופת "קופקס" משפיע על מספר שעות השינה של אדם. במחקר השתתף אדם אחד בשם דני שבכל יום ניתן לו מינון שונה של התרופה. הטבלה שלהלן מתארת בכל יום את מינון התרופה במ"ג שקיבל האדם וכמו כן את מספר שעות השינה שלו באותו הלילה:

מספר היום	מינון התרופה	מספר שעות שינה
1	12	6
2	14	7
3	16	7.5
4	18	6.5
5	20	8

א. כמה תצפיות נאספו במחקר?

ב. סווגו את סולם המדידה של "מינון התרופה"?

© כל הזכויות שמורות לברק קנדל

4) לפניכם רשימה של משתנים.
ציינו באיזה סולם מדידה מדובר (שמי, סדר, כמותי בדיד, כמותי רציף):

א. גובה אדם בס"מ.

ב. מספר ילדים למשפחה.

ג. מידת חרדה לפני מבחן.

ד. שביעות רצון משירות לקוחות בסקלה מ-1 עד 7 (1 כלל לא מרוצה עד 7 מרוצה מאד).

ה. השכלה.

ו. מספר אוטובוס.

ז. מקום מגורים.

ח. מין (1=גבר ו-2=אישה).

ט. מידת נעליים.

5) לפניכם רשימה של משתנים כמותיים.
ציין ליד כל משתנה אם הוא רציף או בדיד:

א. שכר עובד ב-₪.

ב. ציון בחינת בגרות.

ג. תוצאה בהטלת קובייה.

ד. מהירות ריצה במטר לשנייה בתחרות 100 מטר.

ה. שיעור התמיכה בממשלה בעיר.

6) גברת לוי החליטה לדגום 25 ימים של נסיעה לעבודה, כאשר בדרך לעבודתה יש 3 צמתים מרומזרים.
ב-9 ימים הגיעה גברת לוי לעבודה מבלי לעצור באף צומת.
ב-9 ימים נוספים היא הצליחה לעבור בשני רמזורים ירוקים.
ב-5 ימים נוספים היא הצליח לעבור רק בירוק אחד.
בשאר הימים, היא לא עברה באף רמזור ירוק.
מעוניינים לחקור את מספר הרמזורים האדומים בהם עצרה גברת לוי.

א. מהו המשתנה הנחקר בדוגמה זו?

ב. מהם הערכים האפשריים של משתנה זה?

ג. כמה ערכים אפשריים יש למשתנה?

ד. מהו סולם המדידה של המשתנה?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. $n=7$.
ב. 4.
ג. 3.
ד. רווק, נשוי, גרוש, אלמן.

2) א. לקוחות סניף 543 של בנק "רווה".
ב. $n=80$.
ג. X = מספר הפעמים בשבוע שלקוח נכנס לסניף. כמותי בדיד.

3) א. $n=5$.
ב. כמותי רציף.

4) א. כמותי רציף.
ב. כמותי בדיד.
ג. אין מספיק נתונים.
ד. סדר.
ה. אין מספיק נתונים.
ו. שמי.
ז. שמי.
ח. שמי.
ט. סדר.

5) א. רציף.
ב. בדיד.
ג. בדיד.
ד. רציף.
ה. רציף.

6) א. מספר הרמזורים בהם עוצרת גברת לוי ביום בדרך לעבודה.
ב. 0, 1, 2, 3.
ג. 4.
ד. כמותי בדיד.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 2 - סטטיסטיקה תיאורית- הצגה של נתונים

תוכן העניינים

1. כללי 6

סטטיסטיקה תיאורית – הצגה של נתונים:

רקע:

דרכים להצגת נתונים שנאספו:

רשימה של תצפיות:

התצפית היא הערך שנצפה עבור ישות מסוימת בקבוצה. רושמים את התצפיות שהתקבלו כרשומה, יעיל שיש מספר מועט של תצפיות. ההצגה הזו רלבנטית לכל סוגי המשתנים. למשל, להלן מספר החדרים בבניין בן 5 דירות: 4, 5, 3, 4, 3.

טבלת שכיחויות בדידה:

שם המשתנה- X	שכיחות – $f(x)$	שכיחות יחסית באחוזים
X_1	f_1	$\frac{f_1}{N}\cdot 100$
X_2	f_2	$\frac{f_2}{N}\cdot 100$
X_3	f_3	$\frac{f_3}{N}\cdot 100$
⋮	⋮	⋮
X_k	f_k	$\frac{f_x}{N}\cdot 100$
סה"כ	$N=\sum_{i=1}^{k} f_i$	100%

רושמים את התצפיות בטבלה שבה עמודה אחת מבטאת את ערכי המשתנה והשנייה את השכיחות. יעיל עבור משתנה איכותי וכמותי בדיד וכשיש מספר רב של תצפיות. לא יעיל למשתנה כמותי רציף.

© כל הזכויות שמורות לברק קנדל

דוגמה:

להלן התפלגות הציונים בכיתה מסוימת:

הציון- X	מספר התלמידים – השכיחות - f	F_i	$\frac{f_i}{n}$
5	2	2	0.08=2/25
6	4	6	0.16=4/25
7	8	14	0.32=8/25
8	5	19	0.2=5/25
9	4	23	0.16=4/25
10	2	25	0.08=2/25

שכיחות מצטברת – צבירה של השכיחויות.

השכיחויות F_i – השכיחות המצטברת נותנת כמה תצפיות קטנות או שוות לערך.

שכיחות יחסית (פרופורציה) – השכיחות מחולקת לכמות התצפיות הכללי:

$\frac{f_i}{n}$ – איזה חלק מהתצפיות בקבוצה שוות לערך.

טבלת שכיחויות במחלקות:

משתמשים שהמשתנה כמותי רציף או כאשר יש מספר ערכים רב במשתנה הבדיד וטבלת שכיחויות תהיה ארוכה מידי.

דוגמה:

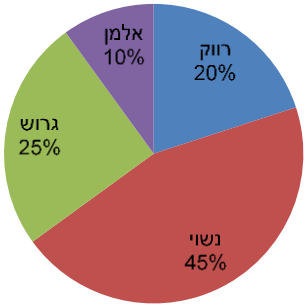
נתנו לקבוצת ילדים לבצע משימה, בדקו את התפלגות זמן הביצוע, בדקות.
להלן ההתפלגות שהתקבלה:

זמן בדקות	מספר הילדים
0.5-3.5	20
3.5-9.5	18
9.5-19.5	14
19.5-29.5	8

© כל הזכויות שמורות לברק קנדל

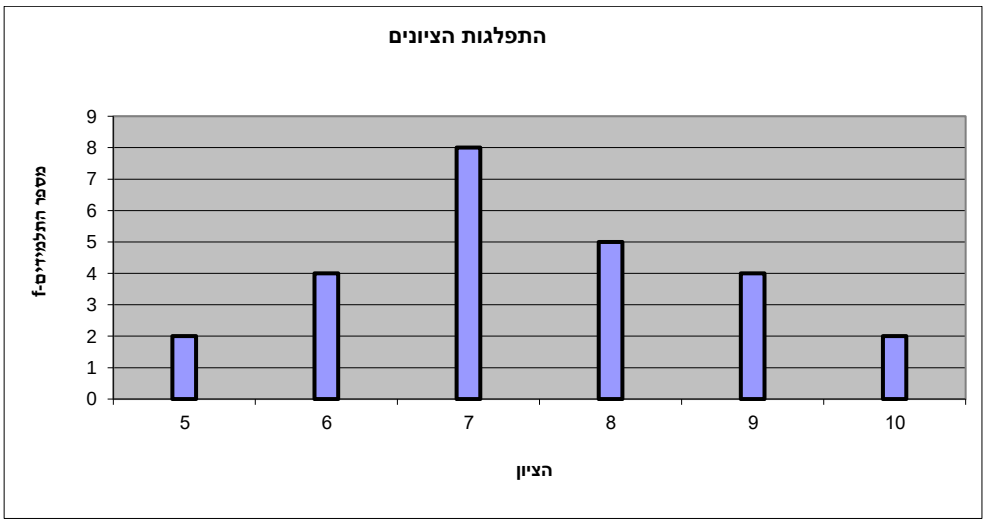
דיאגרמת עוגה:

זהו התיאור הגרפי של משתנה איכותי. בדיאגראמת עוגה כל ערך במשתנה מקבל "נתח", שהוא פרופורציונלי לשכיחות היחסית של ערך המשתנה בנתונים.



דיאגרמת מקלות:

הציר האופקי הוא הציר של המשתנה והציר האנכי של השכיחות, כך שהגובה של המקל מעיד על השכיחות. רלבנטי למשתנה כמותי בדיד. לא נהוג להשתמש בתיאור למשתנה איכותי וכמו כן לא למשתנה כמותי רציף, וכן בסולמות מדידה עבור משתנה מסולם סדר.

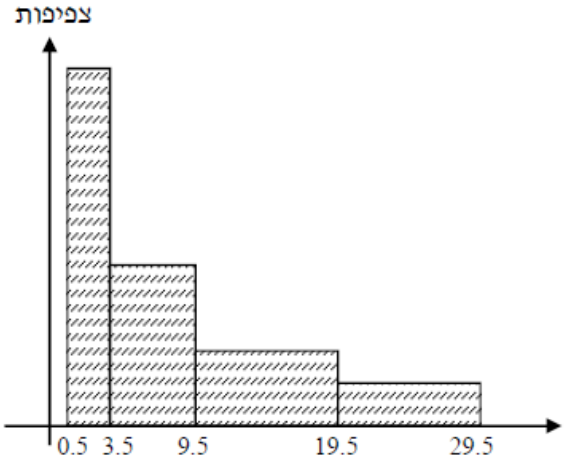


© כל הזכויות שמורות לברק קנדל

היסטוגרמה:

היסטוגרמה היא הדרך הגרפית כדי לתאר טבלת שכיחויות במחלקות, והיא רלוונטית למשתנה כמותי רציף.
בהיסטוגרמה הציר האופקי הוא הציר של המשתנה והציר האנכי הוא הציר של הצפיפות.
הצפיפות מחושבת בכל מחלקה על ידי חלוקת השכיחות ברוחב של כל המחלקה, והיא נותנת את מספר התצפיות הממוצע בכל מחלקה ליחדה. אם המחלקות הן שוות ברוחב, ניתן לשרטט את ההיסטוגרמה לפי השכיחות ואין צורך בצפיפות.

X	רוחב	אמצע	שכיחות	מצטברת	צפיפות
0.5 - 3.5	3	2	20	20	6.6667
3.5 - 9.5	6	6.5	18	38	3
9.5 - 19.5	10	14.5	14	52	1.4
19.5 - 29.5	10	24.5	8	60	0.8



פוליגון – מצולעון:

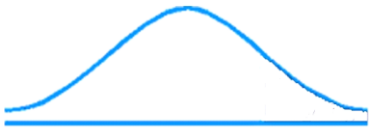
אם נחבר את אמצע קצה כל מלבן בקווים ישרים.
נותן מראה חזותי לצורה של התפלגות המשתנה.

© כל הזכויות שמורות לברק קנדל

צורות התפלגות נפוצות:

התפלגות סימטרית פעמונית

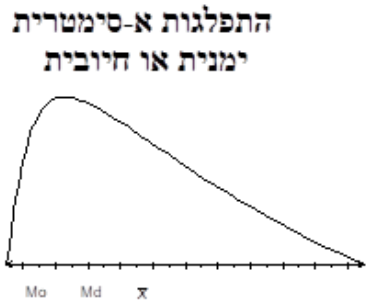
רוב התצפיות במרכז, וככל שנתרחק מהמרכז יהיו פחות תצפיות באופן סימטרי. לדוגמה, ציוני IQ.



ישנן התפלגויות סימטריות שאינן פעמוניות, כגון:

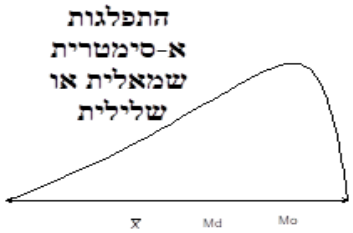
התפלגות אסימטרית ימנית (חיובית)

רוב התצפיות מקבלות ערכים נמוכים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים גבוהים קיצוניים. לדוגמה, שכר במשק.



התפלגות אסימטרית שמאלית (שלילית)

רוב התצפיות מקבלות ערכים גבוהים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים נמוכים קיצוניים. לדוגמה, אורך חיים.



© כל הזכויות שמורות לברק קנדל

שאלות:

1) בסקר צפייה בטלוויזיה התקבלו התוצאות הבאות: 25 צפו בערוץ הראשון, 25 צפו בערוץ 10, 75 צפו בערוץ השני, 50 צפו באחד מערוצי הכבלים ו-25 לא צפו בטלוויזיה בזמן הסקר.

א. רשמו את טבלת השכיחות ואת השכיחות היחסית.

ב. תארו את הנתונים באופן גרפי.

2) להלן נתונים על התפלגות המקצוע המועדף של תלמידי שכבה ו' בבית הספר "מעוף":

מספר התלמידים	המקצוע
44	מתמטיקה
20	תנ"ך
12	אנגלית
26	היסטוריה

א. מהו המשתנה הנחקר?

ב. מהי פרופורציית התלמידים שמעדיפים תנ"ך?

3) להלן התפלגות ההשכלה במקום עבודה מסוים:

מספר העובדים	השכלה
60	נמוכה
120	תיכונית
20	אקדמאית

א. מהו המשתנה הנחקר?
מאיזה סולם הוא?

ב. תארו את הנתונים באופן גרפי.

4) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
6, 5, 8, 7, 6, 9, 8, 6, 7, 6, 7, 8, 5, 4, 6, 10, 9, 8, 6, 7.

א. מהו המשתנה? האם הוא בדיד או רציף?

ב. תארו את הרשימה בטבלת שכיחויות.

ג. הוסיפו שכיחויות יחסיות לטבלה.

ד. תארו את הנתונים באופן גרפי.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. להלן טבלה:

x	$f(x)$	$\frac{f(x)}{n}$	%
ערוץ 1	25	$\frac{25}{200}$	12.5%
ערוץ 10	25	$\frac{25}{200}$	12.5%
ערוץ 2	75	$\frac{75}{200}$	37.5%
כבלים	50	$\frac{50}{200}$	25%
לא צפו	25	$\frac{25}{200}$	12.5%
סה"כ	200	1	100%

ב. עיין גרף מלא בסרטון הוידאו.

2) א. מקצוע מועדף.

ב. 19.6%.

3) א. משתנה נחקר: השכלה, סוג: סדר.

ב. עיין גרף מלא בסרטון הוידאו.

© כל הזכויות שמורות לברק קנדל

4) א. המשתנה: ציון, משתנה בדיד.
ד. עיין גרף מלא בסרטון הוידאו.

ב+ג. להלן טבלה:

x	$f(x)$	$\frac{f(x)}{n}$	%
4	1	$\frac{1}{20}$	5%
5	2	$\frac{2}{20}$	10%
6	6	$\frac{6}{20}$	30%
7	4	$\frac{4}{20}$	20%
8	4	$\frac{4}{20}$	20%
9	2	$\frac{2}{20}$	10%
10	1	$\frac{1}{20}$	5%
סה"כ	20	20	100%

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 3 - סטטיסטיקה תיאורית- סכימה

תוכן העניינים

1. כללי 14

סטטיסטיקה תיאורית – סכימה:

רקע:

בסטטיסטיקה ישנה צורת רישום מקובלת לסכום של תצפיות: $\sum_{i=1}^{n} X_i$.

נסביר את צורת הרישום על ידי הדוגמה הבאה:

i	X_i
1	5
2	0
3	1
4	3
5	2

(הסבר מלא מופיע בסרטונים באתר).

© כל הזכויות שמורות לברק קנדל

שאלות:

1) בבניין 5 דירות. לכל דירה רשמו את מספר החדרים שיש בדירה (X),
ומספר הנפשות החיות בדירה (Y) חשבו:

מספר דירה	X	Y
1	2	1
2	3	1
3	2	2
4	4	3
5	3	2

א. $\sum_{i=1}^{3} X_i$.

ב. $\sum_{i=1}^{5} Y_i$.

ג. $\sum_{i=1}^{4} X_i$.

ד. $\left(\sum_{i=1}^{4} X_i\right)^2$.

ה. $\sum X_i$.

ו. $\sum X_i Y_i$.

ז. $\sum(X_i)\sum(Y_i)$.

© כל הזכויות שמורות לברק קנדל

2) נתון לוח ערכי המשתנים X_i ו- Y_i, כאשר: $i=1,2,...,6$, ונתונים הקבועים: $a=2$, $b=5$. חשבו את הנוסחאות הבאות:

i	1	2	3	4	5	6
X_i	3	2	4	-2	1	4
Y_i	2	0	0	1	-5	2

א. $\sum_{i=1}^{4} y_i$.

ב. $\sum_{i=1}^{6} a$.

ג. $\sum_{i=1}^{6} x_i y_i$.

ד. $\sum_{i=1}^{6} (x_i + y_i)$.

ה. $\sum_{i=1}^{6} x_i + a$.

3) קבעו לכל זהות האם היא נכונה:

א. $\sum_{i=1}^{n} bX_i = b \cdot \sum_{i=1}^{n} X_i$.

ב. $\sum_{i=1}^{n} a = a \cdot n$.

ג. $\left(\sum_{i=1}^{n} X_i \right)^2 = \sum_{i=1}^{n} X_i^{\,2}$.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. 7. ב. 9. ג. 11. ד. 121.
ה. 14. ו. 27. ז. 126.

2) א. 3. ב. 12. ג. 7. ד. 12.
ה. 14.

3) א. נכונה. ב. נכונה. ג. לא נכונה.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 4 - סטטיסטיקה תיאורית - מדדי מיקום מרכזי

תוכן העניינים

1. כללי 18

סטטיסטיקה תיאורית – מדדי מיקום מרכזי:

רקע:

המטרה במדדי המיקום המרכזי היא למדוד את מרכז ההתפלגות של התצפיות.

השכיח – Mode:

השכיח הוא הערך הנפוץ ביותר בהתפלגות.

ברשימה

הערך החוזר על עצמו הכי הרבה פעמים: 6, 10, 4, 8, 4, 9, 7.

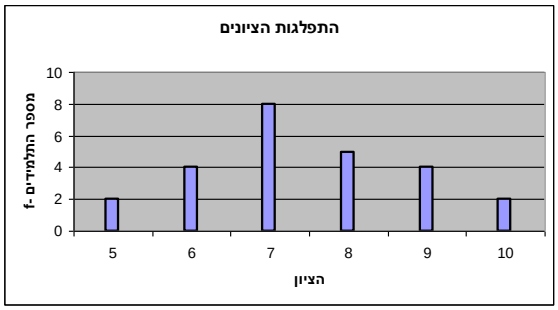
בטבלת שכיחויות בדידה

הערך שהשכיחות שלו היא הגבוהה ביותר.

# תוכניות החיסכון	$f(x)$
0	100
1	75
2	25
3	25
4	25

בדיאגרמת מקלות

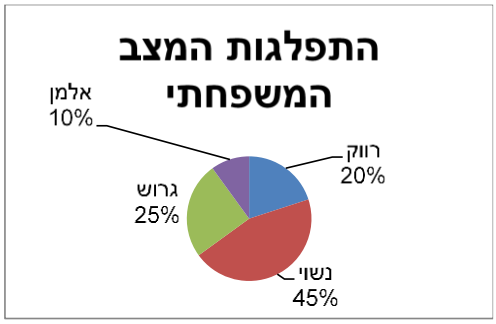
שיעור ה- X של המקל הגבוה ביותר.



© כל הזכויות שמורות לברק קנדל

בעוגה

הערך של הפלח הגדול ביותר.



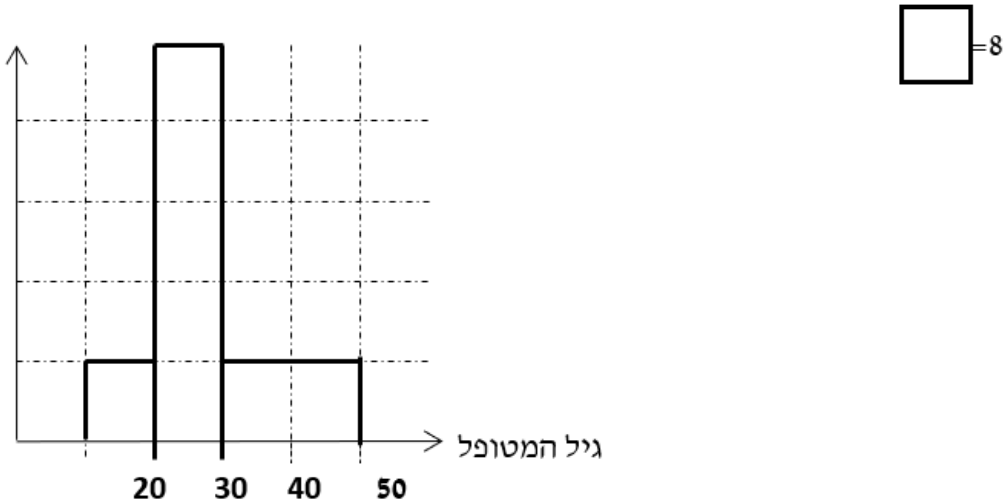
בטבלת שכיחויות במחלקות

אמצע המחלקה עם הצפיפות הגבוהה ביותר.
לדוגמה, התפלגות הציונים בכיתה:

$f(x)$	X
20	0-60
10	60-70
18	70-80
15	80-90
15	90-100

בהיסטוגרמה

שיעור ה- X של אמצע המחלקה הגבוהה ביותר.
לדוגמה, גיל המטופלים של ד"ר שוורץ בשנים:



© כל הזכויות שמורות לברק קנדל

כללי

יתכן שלהתפלגות יותר משכיח אחד.

השכיח הוא מדד הרלבנטי לכל סוגי המשתנים.

אמצע תחום (טווח) – Midrange:

הממוצע בין התצפית הגבוהה ביותר לתצפית הנמוכה ביותר:

$$MR = \frac{X_{\min} + X_{\max}}{2}.$$

החציון – Median:

החציון הוא ערך שמחצית מהתצפיות קטנות או שוות לו ומחצית מהתצפיות גדולות או שוות לו.

ברשימה

נסדר את התצפיות בסדר עולה.

אם יש מספר אי זוגי של איברים, מקומו של החציון יהיה התצפית שמיקומה: $\frac{n+1}{2}$.

אם יש מספר זוגי של איברים – החציון הוא ממוצע של האיבר ה-$\frac{n}{2}$,

והאיבר ה-$\frac{n}{2}+1$, כלומר שיש מספר אי-זוגי של תצפיות החציון יהיה: $md = X_{\frac{n+1}{2}}$,

וכשיש מספר זוגי של תצפיות החציון יהיה: $md = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}$.

בטבלת שכיחויות בדידה

נעשה תהליך דומה אך נעזר בשכיחות המצטברת.

דיאגרמת מקלות

נמיר לטבלת שכיחויות בדידה במטרה למצוא את החציון.

© כל הזכויות שמורות לברק קנדל

בטבלת שכיחויות במחלקות

שלב א: נמצא את המחלקה החציונית שמיקומה יהיה $\frac{n}{2}$.

שלב ב: נציב בנוסחה הבאה: $Md = L_0 + \frac{\frac{n}{2} - F\left(x_{m-1}\right)}{f\left(x_m\right)} \cdot \left(L_1 - L_0\right)$.

$F\left(x_{m-1}\right)$ - שכיחות מצטברת של מחלקה אחת לפני המחלקה החציונית.

$f\left(x_m\right)$ - השכיחות של המחלקה החציונית.

L_0 - גבול התחתון של המחלקה.

L_1 - גבול העליון של המחלקה.

היסטוגרמה

החציון הוא הערך על ציר ה- X שמחלק את ההיסטוגרמה לשני חלקים שווים בשטח.

כללי

החציון אינו רלבנטי למשתנה מסולם שמי ולא רלבנטי למשתנה איכותי.

הממוצע – Average:

הממוצע הוא מרכז הכובד של ההתפלגות.

ברשימה

$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$.

בטבלת שכיחויות

$\bar{x} = \frac{\sum x \cdot f}{n}$.

© כל הזכויות שמורות לברק קנדל

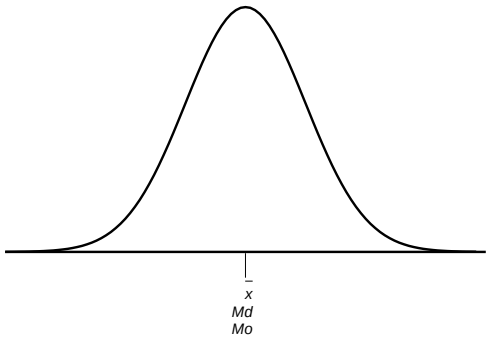
כללי

הממוצע רלבנטי רק למשתנה כמותי.

מדדי המיקום המרכזי בהתפלגויות המיוחדות:

בהתפלגות סימטרית פעמונית כל מדדי המרכז שווים זה לזה:

התפלגות סימטרית



בהתפלגות סימטרית השכיח לא חייב להיות במרכז:

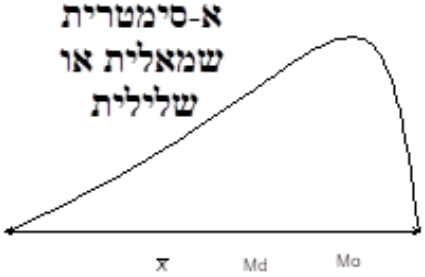
התפלגות U



התפלגות א-סימטרית ימנית או חיובית



התפלגות א-סימטרית שמאלית או שלילית



© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
6, 5, 8, 7, 6, 9, 8, 6, 7, 6, 7, 8, 5, 4, 6, 10, 9, 8, 6, 7.
חשבו את החציון, השכיח, והממוצע של הציונים.

2) בדקו את מספר החדרים לדירה בבניין בן 5 דירות והתקבל ממוצע 3.8.
לגבי 4 דירות נמצא מספר חדרים: 5, 4, 3, 4.

א. כמה חדרים יש בדירה החמישית?

ב. מהו השכיח ומהו החציון?

3) להלן התפלגות מספר מקלטי הטלוויזיה שנספרו עבור כל משפחה בישוב מסוים:

מספר משפחות	מספר מקלטים
22	0
28	1
18	2
22	3
10	4

א. חשבו את הממוצע, החציון והשכיח של ההתפלגות.

ב. הסבירו ללא חישוב כיצד כל מדד שחישבת בסעיף א' היה משתנה אם חלק מהמשפחות (לא כולן) שלא היה להם עד היום טלוויזיה היו רוכשים מקלט אחד.

4) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

מספר מכוניות למשפחה	1	2	3	4	5
שכיחות	65	150	220	140	55

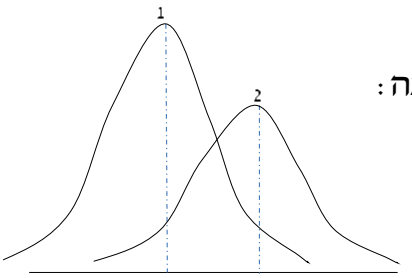
א. כמה משפחות יש בישוב?

ב. מה אחוז המשפחות בישוב עם לכל היותר 2 מכוניות?

ג. חשבו את הממוצע, החציון והשכיח.

הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם!

© כל הזכויות שמורות לברק קנדל



5) מורה לימד 2 כיתות, הוא תיאר באותה מערכת צירים את התפלגות הציונים בכל כיתה. בחרו בתשובה הנכונה:

א. בכיתה 1 השכיח גבוה יותר מכיתה 2.

ב. בכיתה 2 השכיח גבוה יותר מכיתה 1.

ג. בשתי הכיתות אותו שכיח.

ד. לא ניתן לדעת באיזו כיתה השכיח גדול יותר.

תשובות סופיות:

1) חציון: 7, שכיח: 6, ממוצע: 6.9.

2) א. 3. ב. שכיח: 3.4, חציון: 4.

3) א. ממוצע: 1.7, חציון: 1.5, שכיח: 1.

ב. הממוצע יגדל ויתר המדדים לא ישתנו.

4) א. 630. ב. 34.13%. ג. שכיח וחציון: 3, ממוצע: 2.952.

5) ב'.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 5 - סטטיסטיקה תיאורית - מדדי פיזור - הטווח, השונות וסטיית התקן

תוכן העניינים

1. כללי 25

סטטיסטיקה תיאורית – מדדי פיזור – הטווח, השונות וסטיית התקן:

רקע:

המטרה: למדוד את הפיזור של הנתונים, כלומר כמה הם רחוקים זה מזה ושונים זה מזה.

הטווח / תחום (RANGE):
ההפרש בין התצפית הגבוהה ביותר לנמוכה ביותר: $R = X_{\max} - X_{\min}$.

שונות וסטיית תקן:

שונות היא ממוצע ריבועי של הסטיות מהממוצע וסטיית התקן היא שורש של השונות.

עבור סדרת נתונים: $S_x^{\,2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2$.

דוגמאות:

1) נחשב את השונות של סדרת המספרים הבאה: 9, 4, 5.

עבור טבלת שכיחויות: $S_x^{\,2} = \frac{\sum(x-\bar{x})^2 f}{n} = \frac{\sum x^2 \cdot f}{n} - \bar{x}^2$.

2) להלן התפלגות הציונים בכיתה מסוימת בה ממוצע הציונים הוא 7.44.

הציון X	השכיחות F	$x^2 \cdot F$
5	2	50
6	4	144
7	8	392
8	5	320
9	4	324
10	2	200
סה"כ		**1430**

$$S_x^{\,2} = \frac{\sum_x x^2 f(x)}{n} - \bar{x}^2 = \frac{1430}{25} - 7.44^2 = 1.8464$$

$$S = \sqrt{S_x^{\,2}} = \sqrt{1.8464} = 1.3588$$

כשיש מחלקות נעזר באמצע המחלקה כדי לחשב את השונות.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
6, 5, 8, 7, 6, 9, 8, 6, 7, 6, 7, 8, 5, 4, 6, 10, 9, 8, 6, 7.
חשבו את השונות, סטיית התקן והטווח של הציונים.

2) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

מספר מכוניות למשפחה	1	2	3	4	5
שכיחות	65	150	220	140	55

א. חשבו סטיית התקן.

ב. חשבו את הטווח של הנתונים.

הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם.

3) בחברה העוסקת בטלמרקטינג בדקו עבור כל עובד את מספר שנות הוותק שלו. התקבל שממוצע שנות הוותק הוא 4 שנים וסטיית התקן היא שנתיים.

א. האם הממוצע יגדל/יקטן/לא ישתנה וסטיית התקן תגדל/תקטן/לא תשנה כאשר יתווספו שני עובדים עם וותק של 4 שנים להתפלגות?

ב. האם הממוצע יגדל/יקטן/לא ישתנה וסטיית התקן תגדל/תקטן/לא תשנה כאשר יתווספו שני עובדים אשר אחד עם וותק של 0 שנים והשני עם וותק של 8 שנים להתפלגות?

4) נתונה רשימה של 5 תצפיות, אך רק עבור 4 מהן נרשמו הסטיות שלהן מהממוצע: 2, 3, 2, -1. חשבו את השונות של חמש התצפיות.

5) בשכונה בדקו בכל דירה את מספר החדרים לדירה. בשכונה 200 דירות.

מספר חדרים	פרופורציה
1	0.1
2	0.2
3	0.4
4	0.15
5	

א. מה הממוצע של מספר החדרים לשכונה בדירה?

ב. חשבו את סטיית התקן של מספר החדרים לדירה.

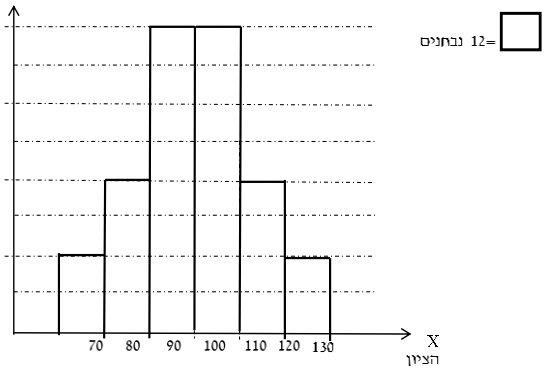
ג. חלק מבעלי הדירות בנות 2 החדרים הפכו את דירתם לדירת חדר. כיצד הדבר ישפיע (יקטין, יגדל, לא ישנה) על כל מדד שחישבתם בסעיפים הקודמים.

© כל הזכויות שמורות לברק קנדל

6) להלן התפלגות המשקל של קבוצה מסוימת בק"ג:
מהי סטיית התקן של התפלגות המשקל?

משקל	מספר מקרים
40-45	10
45-50	20
50-60	30
60-65	20
65-70	10

7) להלן התפלגות הציונים במבחן אינטליגנציה:



א. מה הממוצע ומה החציון של ההתפלגות?

ב. חשבו את סטיית התקן של הציונים.

ג. מסתבר שיש להוסיף 20 תצפיות לכל אחת משתי המחלקות 90-100 ו- 100-110. כיצד הדבר ישתנה את כל אחד מהמדדים של הסעיפים הקודמים?

תשובות סופיות:

1) שונות: 2.19, סטיית תקן: 1.48, טווח: 6.

2) א. סטיית תקן: 1.106. ב. טווח: 4.

3) א. ממוצע לא ישתנה, סטיית התקן תקטן.

ב. ממוצע לא ישתנה, סטיית התקן תגדל.

4) 10.8.

5) א. 3.05. ב. 1.16. ג. ממוצע: יקטן, סטיית התקן: תגדל.

6) 7.73.

7) א. 100. ב. 12.96. ג. ממוצע: לא ישתנה, סטיית תקן: תקטן.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 6 - סטטיסטיקה תיאורית - מדדי פיזור - טווח בין רבעוני

תוכן העניינים

1. טווח בינרבעוני 28

סטטיסטיקה תיאורית – מדדי פיזור – טווח בין רבעוני:

רקע:

הטווח הבין-רבעוני (יש הקוראים לו התחום הבין-רבעוני) נותן את הטווח בין הרבעונים בו נמצאים 50% מהתצפיות המרכזיות. הרעיון ליצור מדד פיזורי שלא רגיש לתצפיות החריגות ביותר. כדי לחשב את הטווח הבין-רבעוני יש למצוא את הרבעון התחתון והעליון של התפלגות התצפיות.

רבעון תחתון – ערך שמחלק את ההתפלגות לשניים.
25% מהמקרים נמוכים ממנו או שווים לו ו75% מהמקרים גבוהים או שווים לו.
סימון: Q_1.

רבעון עליון – ערך שמחלק את ההתפלגות לשניים.
75% מהמקרים נמוכים ממנו או שווים לו ו25% מהמקרים גבוהים או שווים לו.
סימון: Q_3.

הטווח הבין רבעוני הוא הפער בין שני הרבעונים: $IQR = Q_3 - Q_1$.

שלבים במציאת טווח בין-רבעוני בטבלת שכיחויות:

שלב א: נמצא את הרבעון התחתון: הוא הערך שהשכיחות היחסית המצטברת באחוזים עברה לראשונה את 25%.

שלב ב: נמצא את הרבעון העליון: הוא הערך שהשכיחות היחסית המצטברת באחוזים עברה לראשונה את 75%.

שלב ג: נמצא את הטווח הבין-רבעוני: נחסר את הרבעונים: $IQR = Q_3 - Q_1$.

דוגמה (פתרון בהקלטה):

בסניף בנק 250 לקוחות. ספרו לכל לקוח את מספר תוכניות החיסכון שלו.
מהו הטווח הבין-רבעוני של מספר תוכניות החיסכון בסניף?

# תוכניות החיסכון	$f(x)$	שכיחות מצטברת	שכיחות יחסית מצטברת
0	100		
1	75		
2	25		
3	25		
4	25		

© כל הזכויות שמורות לברק קנדל

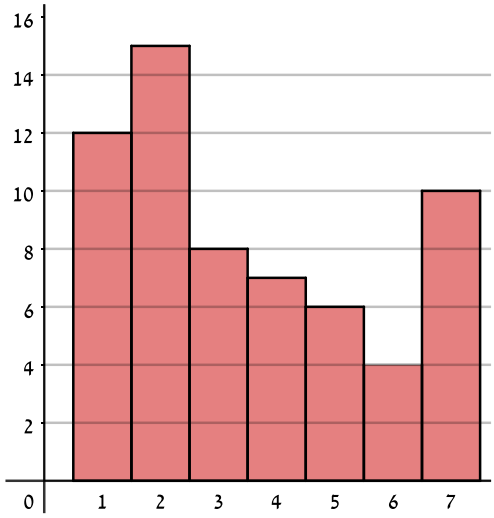
שאלות:

1) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

מספר מכוניות למשפחה	1	2	3	4	5
שכיחות	65	150	220	140	55

מהו הטווח הבין-רבעוני של מספר המכוניות למשפחה בישוב "הגורן"?

2) בסקר שנעשה בדקו את מספר ימי המחלה השנתיים של מורים בארץ.



א. מה מייצגים הערכים בציר האופקי?

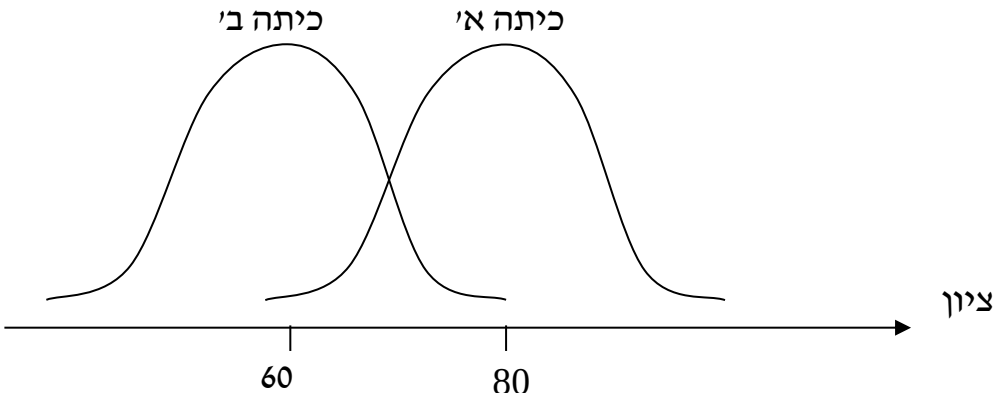
ב. מהו הטווח הבין-רבעוני של מספר ימי המחלה של המורים

ג. אם נוסיף 25 מורים אשר הצהירו שמספר ימי המחלה השנתיים שלהם הוא 4 ימים, כיצד הדבר ישנה את הטווח הבין-רבעוני? הסבירו.

ד. אם מסתבר שחלק מהמורים בסקר הצהירו שהם חלו 7 ימים בשנה אבל בפועל הם חלו 8 ימים, כיצד הדבר ישנה את הטווח הבין-רבעוני? הסבירו.

© כל הזכויות שמורות לברק קנדל

3) לפניך שתי עקומות המתארות את התפלגות הציונים בכל כיתה.
באיזו כיתה הטווח הבין-רבעוני גדול יותר?



א. כיתה א.
ב. כיתה ב'.
ג. לשתיהן אותו טווח בין-רבעוני.
ד. לא ניתן לדעת, אין מספיק נתונים.

4) הוספת גודל קבוע לכל תצפיות סדרת נתונים:
א. תגדיל את הטווח הבין-רבעוני.
ב. תקטין את הטווח הבין-רבעוני.
ג. לא תשנה הטווח הבין-רבעוני.
ד. לא ניתן לדעת מה יקרה לטווח הבין-רבעוני.

5) חושב הטווח הבין-רבעוני עבור התפלגות מסוימת והתקבלה התוצאה אפס.
לכן:
א. לפחות 50% מהתצפיות זהות.
ב. סטיית התקן היא אפס.
ג. ההתפלגות היא סימטרית.
ד. מצב זה כלל לא יתכן.

© כל הזכויות שמורות לברק קנדל

6) סניף מספר 543 של בנק "רווה" בדק ל-80 לקוחות את מספר הפעמים שכל לקוח נכנס לסניף הבנק במשך שבוע. התוצאות שהתקבלו הן:
50 אנשים נכנסו 0 פעמים לסניף.
20 אנשים נכנסו פעם אחת לסניף.
5 אנשים נכנסו פעמיים לסניף.
5 אנשים נכנסו יותר מפעמיים.
מהו הטווח הבין-רבעוני?

א. 60.
ב. 2.
ג. 50.
ד. 1.

7) התפלגות הציונים במבחן ווקסלר היא סימטרית לכן:

א. טווח הציונים הוא אפס.
ב. הטווח הבין רבעוני של הציונים אפס.
ג. סעיפים א ו-ב הם נכונים.
ד. אף סעיף אינו נכון.

תשובות סופיות:

1) 2.
2) א. מספר ימי המחלה השנתיים. ב. 3. ג. יקטן. ד. לא ישתנה.
3) ג'.
4) ג'.
5) א'.
6) ד'.
7) ד'.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 7 - סטטיסטיקה תיאורית- מדדי מיקום יחסי-ציון תקן

תוכן העניינים

1. כללי 32

סטטיסטיקה תיאורית – מדדי מיקום יחסי – ציון תקן:

רקע:

המטרה למדוד איך תצפית ממוקמת ביחס לשאר התצפיות בהתפלגות.

ציון תקן:

הנוסחה לציון תקן של תצפית היא: $Z = \frac{X - \overline{X}}{S}$.

ציון התקן נותן כמה סטיות תקן סוטה התצפית מהממוצע. כלומר, ציון התקן מעיד על כמה סטיות תקן התצפית מעל או מתחת לממוצע:

- ציון תקן חיובי אומר שהתצפית מעל הממוצע.
- ציון תקן שלילי אומר שהתצפית מתחת לממוצע.
- ציון תקן אפס אומר שהתצפית בדיוק בממוצע.

דוגמה (פתרון בהקלטה):

במקום עבודה מסוים, ממוצע המשכורות הוא 8 אלף ₪, עם סטית תקן של אלפיים ₪.
באותו מקום עבודה ההשכלה הממוצעת של העובדים הנה 14 שנים, עם סטית תקן של 1.5 שנים. ערן מרוויח במקום עבודה זה 11 אלף ₪ והשכלתו 16 שנים.
מה ערן יותר, באופן יחסי, משכיל או משתכר?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) תלמידי כיתה ח' ניגשו למבחן בלשון ולמבחן במתמטיקה.
להלן התוצאות שהתקבלו:

המקצוע	ממוצע	סטיית תקן
לשון	74	12
מתמטיקה	80	16

עודד קיבל: 68 בלשון ו-70 במתמטיקה.

א. באיזה מקצוע עודד טוב יותר באופן יחסי לשכבה שלו?

ב. איזה ציון עודד צריך לקבל במתמטיקה כדי שיהיה שקול לציונו בלשון?

2) במפעל לייצור מצברים לרכב בדקו במשך 40 ימים את התפוקה היומית (מספר מצברים במאות) ואת מספר הפועלים שעבדו באותו היום.
להלן טבלה המסכמת את המידע שנאסף על שני המשתנים:

	תפוקה	מספר פועלים
ממוצע	48	15
סטיית תקן	10	2

באחד הימים מתוך כלל הימים שנבדקו התפוקה הייתה 50 מאות מצברים ובאותו היום עבדו 13 פועלים.
מה יותר חריג באותו היום, יחסית לשאר הימים שנבדקו: נתוני התפוקה או כמות הפועלים?

א. התפוקה.

ב. כמות הפועלים.

ג. חריגים באותה מידה.

ד. חסרים נתונים כדי לדעת זאת.

3) הגובה הממוצע של המתגייסים לצבא הוא 175 סנטימטר עם סטיית תקן של 10 סנטימטר. המשקל הממוצע הוא 66 ק"ג עם סטיית תקן של 8 ק"ג.
ערן התגייס כשגובהו 180 ס"מ ומשקלו 59 ק"ג.

א. במה ערן חריג יותר ביחס לשאר המתגייסים, גובהו או משקלו?

ב. כמה ערן אמור לשקול כדי שמשקלו יהיה שקול לגובהו?

תשובות סופיות:

1) א. לשון. ב. 72.

2) ב'.

3) א. משקל. ב. 70.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 8 - סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה

תוכן העניינים

1. כללי 34

סטטיסטיקה תיאורית – מדדי מיקום יחסי – אחוזונים בטבלה בדידה:

רקע:

האחוזון (המאון) ה- p הוא הערך בנתונים המחלק את הנתונים בצורה כזאת, שעד אליו (כולל) יש p % מהנתונים. מסמנים את האחוזון ה- p ב- X_p .

חישוב האחוזון מתוך נתונים בטבלת שכיחויות בדידה:

האחוזון הוא הערך שבו בפעם הראשונה השכיחות היחסית המצטברת (באחוזים) גדולה או שווה ל- p % .

דוגמה (פתרון בהקלטה):

בסניף בנק 250 לקוחות. ספרו לכל לקוח את מספר תוכניות החיסכון שלו:

# תוכניות החיסכון	$F(x)$	שכיחות מצטברת	שכיחות יחסית מצטברת
0	100		
1	75		
2	25		
3	25		
4	25		

א. מצאו את האחוזון ה-25.

ב. מצאו את הערך ש-20% מהמקרים מעליו.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן התפלגות של משתנה כלשהו:

X	$F(x)$
0	10
1	40
2	30
3	15
4	5

מצאו להתפלגות את:

א. האחוזון ה-60.

ב. המאון ה-40.

ג. העשירון העליון.

ד. הטווח בין הרבעונים.

2) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

מספר מכוניות למשפחה	1	2	3	4	5
שכיחות	65	150	220	140	55

חשבו את:

א. העשירון התחתון.

ב. האחוזון ה-30.

ג. הערך ש-20% מהתצפית גדולות ממנו.

ד. רבעון עליון.

תשובות סופיות:

1) א. 2 ב. 1. ג.3. ד. 1.

2) א. 1. ב. 2. ג. 4. ד. 4.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 9 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

תוכן העניינים

1. כללי 36

סטטיסטיקה תיאורית – טרנספורמציה לינארית:

רקע:

מצב שבו מבצעים שינוי מסוג הוספה (או החסרה) של קבוע, והכפלה (או חילוק) של קבוע, לכל התצפיות: $y = a \cdot x + b$. כך יושפעו המדדים השונים:

מדדי המרכז:

$$MR_y = a \cdot MR + b$$
$$MO_y = a \cdot MO + b$$
$$\bar{y} = a \cdot \bar{x} + b$$
$$Md_y = a \cdot \leq Md_x + b$$

מדדי הפיזור:

$$R_y = |a| R_x$$
$$S_y = |a| S_x$$
$$S_y^{\,2} = a^2 S_x^{\,2}$$

מדדי המיקום היחסי:

$$Y_p = a \cdot X_p + b$$
$$Z_Y = \frac{a}{|a|} Z_x$$

שלבי העבודה:

1. נזהה שמדובר בטרנספורמציה לינארית (שינוי קבוע לכל התצפיות).
2. נרשום את כלל הטרנספורמציה לפי נתוני השאלה.
3. נפשט את הכלל ונזהה את ערכי a ו-b.
4. נציב בנוסחאות שלעיל בהתאם למדדים שנשאלים.

דוגמה (פתרון בהקלטה):

השכר הממוצע של עובדים הינו 9000 ₪ וטווח 6000 ₪. חשבו את המדדים הללו לאחר שהעלו את כל המשכורות ב-10% ואחר כך קנסו אותם ב-100 ₪.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) עבור סדרת נתונים התקבל: $\bar{x}=80, S=15,\ MO=70$.
הוחלט להכפיל את כל התצפיות ב-4 ולהחסיר מהתוצאה 5.
חשבו את המדדים הללו לאחר השינוי.

2) בחברה מסוימת השכר הממוצע הוא 40 ₪ לשעה עם סטיית תקן של 5 ₪ לשעה.
הוחלט להעלות את כל המשכורות ב-10%, אך זה לא סיפק את העובדים ולכן הם קיבלו לאחר מכן תוספת של 2 ₪ לשעה.
מה הממוצע ומהי השונות של השכר לשעה לאחר כל השינויים.

3) במבחן מסוים הציון החציוני היה 73, טווח הציונים היה 40 נקודות והעשירון העליון היה הציון 87. כיוון שהציונים בבחינה היו נמוכים, המורה החליט לתת פקטור של 4 נק' לכל התלמידים.
חשבו את המדדים לאחר הפקטור.

4) דגמו מקו ייצור 50 קופסאות של גפרורים. בדקו בכל קופסא בה יש 40 גפרורים את כמות הגפרורים הפגומים. התקבל שבממוצע יש 3 גפרורים פגומים בקופסא, עם סטיית תקן של 1.5 גפרורים.
מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקינים בקופסא?

5) חברת בזק הציעה את ההצעה הבאה:
שלושים שקלים דמי מנוי חודשיים קבועים וכן 10 אגורות לכל דקה של שיחה יוצאת. אדם בדק במשך שנה את דקות השיחות היוצאות שלו, וקיבל שבממוצע חודשי יש לו 600 דקות שיחות יוצאות עם שונות של 2500 דקות רבועות, כמו כן בחודש ינואר ציון התקן היה 2.
חשבו את המדדים הללו עבור חשבון הטלפון החודשי של אותו אדם בשקלים אם היה משתמש בחבילה המוצעת לו על ידי בזק.

תשובות סופיות:

1) ממוצע: 315, סטיית תקן: 60, שכיח: 275.
2) ממוצע: 46, שונות: 30.25.
3) טווח: 40, חציון: 77, עשירון עליון: 91.
4) ממוצע: 37, סטיית תקן: 1.5.
5) ממוצע: 90, שונות: 25, ציון תקן: 2.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 10 - סטטיסטיקה תיאורית שאלות אמריקאיות- מסכמות

תוכן העניינים

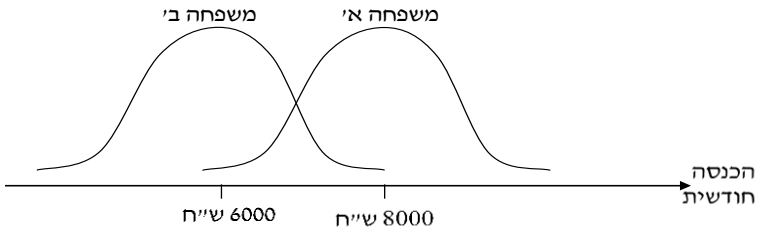
1. כללי 38

סטטיסטיקה תיאורית – שאלות אמריקאיות:

שאלות:

שאלות 1-3 מתייחסות לקטע הבא:

להלן שתי עקומות המתארות את התפלגות ההכנסות החודשיות של שתי משפחות שנבחרו באקראי:



1) לאיזו משפחה הכנסה שכיחה גבוהה יותר?
א. משפחה א'.
ב. משפחה ב'.
ג. לשתיהן אותה הכנסה שכיחה.
ד. לא ניתן לדעת – אין מספיק נתונים.

2) באיזו משפחה ההכנסה החציונית שווה להכנסה הממוצעת?
א. משפחה א'.
ב. משפחה ב'.
ג. בשתיהן ההכנסה החציונית שווה להכנסה הממוצעת.
ד. לא ניתן לדעת – אין מספיק נתונים.

3) באיזו משפחה סטית התקן של ההכנסה החודשית גבוהה יותר?
א. משפחה א'.
ב. משפחה ב'.
ג. לשתיהן אותה סטית תקן.
ד. לא ניתן לדעת – אין מספיק נתונים.

© כל הזכויות שמורות לברק קנדל

הנתונים הבאים מתייחסים לשאלות 4-6:

להלן נתונים חלקיים של טבלת שכיחויות:
כמו כן, נתון כי הממוצע הוא 1.66.

x	$F(x)$
0	?
1	10
2	6
3	15
4	?
סה"כ	**50**

4) השכיח של הנתונים הוא:
א. 0.
ב. 15.
ג. ישנם שני שכיחים: 0 ו-3.
ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של השכיח.

5) חציון הנתונים הוא:
א. 2.
ב. 1.5.
ג. 25.5.
ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של החציון.

6) הטווח של הנתונים:
א. 11.
ב. 3.
ג. 4.
ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של החציון.

7) בהתפלגות אסימטרית ימנית של משתנה כמותי רציף, הערך המתאים למאון ה-30, ציון התקן שלו הוא בהכרח:
א. שלילי.
ב. חיובי.
ג. אפס.
ד. לא ניתן לדעת ללא הנתונים.

© כל הזכויות שמורות לברק קנדל

8) סדרת נתונים סטטיסטיים מונה 10 תצפיות. נתון כי סדרת הנתונים סימטרית סביב הממוצע. ממוצע הסדרה-40 ושונות הסדרה-100. בשלב מאוחר יותר נוספו שתי תצפיות נוספות לסדרה: 50 ו-30. השונות של 12 התצפיות:

א. תקטן.
ב. תגדל.
ג. לא תשתנה.
ד. לא ניתן לחשב את השונות ללא ידיעת התצפיות.

9) הוספת גודל קבוע לכל תצפיות סדרת נתונים:

א. תגדיל את סטיית התקן.
ב. תקטין את סטיית התקן.
ג. לא תשנה את סטיית התקן.
ד. לא ניתן לדעת.

הנתונים הבאים מתייחסים לשאלות 10-11:

להלן נתונים על ציוני תלמידים שנבחנו במועדים שונים בסטטיסטיקה:

שם התלמיד	ציון	ממוצע הציונים במועד בו נבחן	סטיית התקן של הציונים במועד בו נבחן
צבי	50	50	12
סטף	82	80	5
שרית	65	60	15
לובה	60	63	1.5
מיטב	70	70	10

10) התלמיד הטוב ביותר ביחס לנבחנים באותו מועד בו נבחן הוא:

א. מיטב.
ב. צבי.
ג. לובה.
ד. שרית.
ה. סטף.

11) פנינה נבחנה עם סטף וציון התקן שלה שווה לציון התקן של שרית לכן ציונה הוא:

א. 80.55.
ב. 65.
ג. 80.
ד. 81.66.

© כל הזכויות שמורות לברק קנדל

הנתונים הבאים מתייחסים לשאלות 12-15:

בבדיקת פתע של משרד הבריאות במפעל שוקולד, נמצא ש:

שוקולד פגום	0	1	2	3	4	5	6	7
מס' קופסאות	35	63	48	12	13	11	10	8

12) מהו החציון של מספר הפגומים בקופסא:

א. 1.
ב. 2.
ג. 4.
ד. לא ניתן לדעת.

13) מהו הרבעון התחתון של מספר הפגומים בקופסא?

א. 1.
ב. 2.
ג. 3.
ד. 4.
ה. לא ניתן לדעת.

14) מספר הפגומים בקופסא הוא משתנה:

א. סדר.
ב. שמי.
ג. כמותי בדיד.
ד. כמותי רציף.

15) השכיח של מספר הפגומים בקופסא:

א. 63.
ב. 1.
ג. 200.
ד. לא ניתן לדעת.

16) ביחס לציר המספרים, רוב הערכים בהתפלגות א-סימטרית ימנית נמצאים:

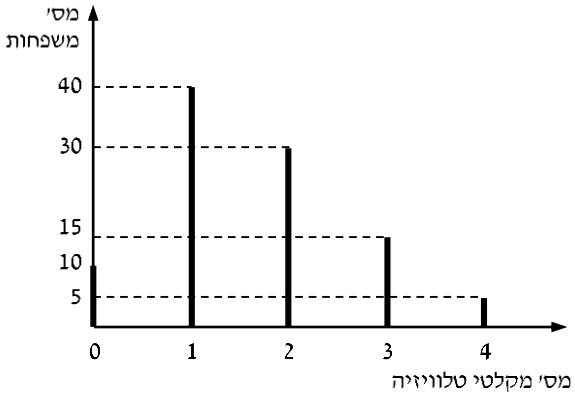
א. בערכים הגבוהים.
ב. בחלוקה זהה בין הערכים הגבוהים והנמוכים.
ג. בערכים הנמוכים.
ד. לא ניתן לדעת.
ה. אף לא תשובה מהנ"ל נכונה.

© כל הזכויות שמורות לברק קנדל

17) בוצע מחקר על מספר העובדים בחברות מזון לעומת חברות תקשורת.
החציון והממוצע בשתיהן שווה 8.
איזה מהטענות הבאות היא הנכונה והמלאה ביותר:
א. השכיחות ב-2 החברות זהה אך שונה מ-8.
ב. השכיח ב-2 החברות זהה אך לא ניתן לדעת מהו.
ג. השכיח בשתי חברות הינו בהכרח 8.
ד. שכיח בחברה אחת שונה מ-8 ובשנייה הוא 8.
ה. אף תשובה אינה נכונה.

הנתונים הבאים מתייחסים לשאלות 18 עד 22:

נערך סקר על מספר מקלטי הטלוויזיה הנמצאים בבית.
תוצאות הסקר נתונות בדיאגרמת מקלות הבאה:



18) המשתנה הנחקר כאן הוא:
א. משתנה שמי.
ב. משתנה מסולם סדר.
ג. משתנה כמותי בדיד.
ד. משתנה כמותי רציף.

19) הטווח של ההתפלגות הוא:
א. 35.
ב. 4.
ג. 3.
ד. 2.

© כל הזכויות שמורות לברק קנדל

20) ממוצע מספר מקלטי הטלוויזיה למשפחה הוא:

א. 1.65.
ב. 1.5.
ג. 1.
ד. 2.

21) השכיח של התפלגות זו היא:

א. 40.
ב. 1.5.
ג. 1.
ד. 2.

22) מסתבר שיש בין 2 ל-5 משפחות נוספות שאין להם מקלטי טלוויזיה ויש לצרף את המשפחות הללו להתפלגות. כיצד הנתון זה ישפיע על סטיית התקן?

א. יקטין אותו.
ב. יגדיל אותו.
ג. לא ישנה אותו.
ד. אין לדעת.

תשובות סופיות:

1) א'	**2)** ג'	**3)** ג'	**4)** ג'	**5)** ב'
6) ג'	**7)** א'	**8)** ג'	**9)** ב'	**10)** ה
11) ד'	**12)** ג'	**13)** א'	**14)** ג'	**15)** ב'
16) ג'	**17)** ה'	**18)** ג'	**19)** ב'	**20)** א'
21) ג'	**22)** ב'			

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 11 - מדדי קשר - מדד הקשר של קרמר

תוכן העניינים

1. כללי 44

מדדי קשר – מדד הקשר של קרמר:

רקע:

משתמשים במדד זה כאשר אחד המשתנים הוא מסולם שמי והשני מכל סולם אפשרי.
מדד הקשר מקבל ערכים בין 0 ל-1. ככל שהמדד יותר קרוב לאחד קיים קשר בעוצמה יותר חזקה בין המשתנים.

דוגמה (פתרון בהקלטה):

במחקר רוצים לבדוק את הקשר בין מין לדעה בנושא מסוים. שאלו 100 גברים ו-100 נשים על דעתם באיזשהו נושא. להלן טבלת השכיחויות המשותפת שהתקבלה:

Y/X	בעד	נגד	נמנע	$F(x)$
גבר	50	40	10	100
אישה	30	60	10	100
$F(y)$	80	100	20	$n=200$

הטבלה נקראת טבלת O (observe):

X - מין (גבר/אישה) – סולם שמי.

Y - דעה (בעד/נמנע/נגד) – סולם שמי/סדר.

שלבים בחישוב r_c:

שלב א': נבנה את טבלת E (Expected).

נעתיק את המסגרת של טבלת O ואז כל: $E_i=\left(F(x)\cdot F(y)\right)/n$.

Y/X	בעד	נגד	נמנע	$f(x)$
גבר				100
אישה				100
$f(y)$	80	100	20	$n=200$

שלב ב': נחשב $\chi^2=\sum_i \frac{(O_i-E_i)^2}{E_i}$.

שלב ג': נחשב: $r_c=\sqrt{\frac{1}{n(L-1)}\chi^2}$,

כאשר L מבטא את המספר הקטן מבין מספר השורות או העמודות.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן תוצאות מחקר שבדק את הקשר בין מין להשכלה.
לגבי כל נחקר נבדק המין שלו והשכלתו.
האם קיים קשר בין מין להשכלה? נמקו!

מין/ השכלה	נמוכה	תיכונית	גבוהה
גבר	120	40	20
אישה	20	20	80

2) נלקחו 200 אנשים שמתוכם 60 הצהירו שהם עוסקים בפעילות גופנית סדירה.
מתוך אלו שעוסקים בפעילות גופנית סדירה 50 נמצאו במצב בריאותי תקין.
מתוך אלו שלא עוסקים בפעילות גופנית סדירה 90 נמצאו במצב בריאותי תקין.

א. בנו טבלת שכיחות משותפת לנתונים שהוצגו בשאלה.

ב. האם קיים קשר בין פעילות גופנית למצב בריאותי?
חשבו לפי מדד הקשר של קרמר.

תשובות סופיות:

1) ישנו קשר בעוצמה בינונית, מקדם המתאם של קרמר: 0.595.

2) א. להלן טבלה: ב. 0.19.

y/x	תקין	לא תקין	$f(x)$
כן	50	10	60
לא	90	50	140
$f(y)$	140	60	$n=200$

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 12 - מדד הקשר ספירמן

תוכן העניינים

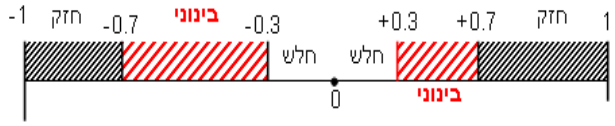
1. כללי 46

מדדי קשר – מדד הקשר של ספירמן:

רקע:

מתי נשתמש במדד ספירמן?
כאשר אחד המשתנים מסולם סדר והשני מסולם סדר ומעלה.
הקשר שהמדד בודק הוא קשר דירוגי.
מדד הקשר בודק את:

1. כיוון הקשר.
2. עצמת הקשר.



המדד מקבל ערכים בסקלה מ-(1-) ועד 1.

קשר דירוגי חיובי מלא:
מדד הקשר של ספירמן יוצא 1.
ככל שמשתנה אחד עולה, השני עולה ללא יוצא מן הכלל.

קשר דירוגי חיובי חלקי:
מקדם המתאם בין 0 ל-1.
ככל שמשתנה אחד עולה, לשני יש נטייה לעלות אך לא באופן מוחלט.

קשר דירוגי שלילי מלא:
מדד הקשר של ספירמן יוצא 1-.
ככל שמשתנה אחד עולה השני יורד ללא יוצא מן הכלל.

קשר דירוגי שלילי חלקי:
מקדם המתאים הוא בין 0 ל-(1-).
ככל שמשתנה אחד עולה לשני יש נטייה לרדת אך לא באופן מוחלט.
על מנת לחשב את הקשר יש לבצע פעולת דירוג (RANK).
כאשר מדרגים, אם יש כמה תצפיות שתופסות את אותו הערך אז הדירוג שלהם הוא הממוצע של המקומות שהן תופסות.

הנוסחה של מדד הקשר: $r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2-1\right)}$.

© כל הזכויות שמורות לברק קנדל

דוגמה (פתרון בהקלטה):

בתחרות רוקדים עם כוכבים השתתפו 7 זוגות, 2 שופטים נתנו את ציוניהם לריקוד של כל זוג. להלן התוצאות שהתקבלו:

מספר הזוג	ציון שופט א'	R_x	ציון שופט ב'	R_y	$d=r_x-r_y$	d^2
1	4		5			
2	5		5			
3	6		7			
4	5		7			
5	8		9			
6	7		9			
7	3		7			

מהי מידת ההתאמה בין ציוני השופטים?

X - ציון שופט א (סולם סדר).

Y - ציון שופט ב (סולם סדר).

© כל הזכויות שמורות לברק קנדל

שאלות:

1) בתחרות יופי חילקו שני שופטים ציונים למועמדות:

מספר מועמדת	1	2	3	4	5	6	7
ציון שופט א'	7	8	6	8	9	5	6
ציון שופט ב'	8	8	7	8	9	5	7

האם קיים קשר בין שתי הערכות השופטים? נמקו והסבירו!

2) משרד רצה לבחון האם קיים קשר בין מידת המוטיבציה של העובדים שלו לבין מספר החיסורים של העובדים בחודש עבודה. להלן התוצאות שהתקבלו:

מספר חיסורים	מידת מוטיבציה
0	גבוהה
4	נמוכה
2	בינונית
5	נמוכה
1	גבוהה

האם קיים קשר בין רמת המוטיבציה של העובד ומספר החיסורים שלו?
חשבו באמצעות מדד הקשר המתאים והסבירו.

3) אם: $r_S = 1$, הדבר אומר שערכי X תמיד שווים לערכי Y.
האם הטענה נכונה? הסבר.

תשובות סופיות:

1) קיים קשר דירוג חיובי חזק בין הערכת שופט א' להערכת שופט ב'.
מדד הקשר: 0.973.

2) קיים קשר שלילי בעוצמה חזקה בין רמת המוטיבציה של העובד למס' החיסורים שלו.
מדד הקשר: -0.85.

3) לא נכון.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 13 - מקדם המתאם הלינארי (פירסון)

תוכן העניינים

1. מקדם המתאם הלנארי (פירסון) 49
2. חישוב מקדם המתאם הלנארי (פירסון) 60

מקדם המתאם (מדד קשר) הלינארי ומובהקותו

מדד הקשר הלינארי (פירסון) – מבוא

מעוניינים לבדוק עד כמה קיים קשר מסוג קשר לינארי (קו ישר) בין שני משתנים.
שני המשתנים שאנו בודקים לגביהם קשר צריכים להיות משתנים כמותיים.
מבחינת סולמות מדידה כל משתנה נחקר צריך להיות מסולם רווחים או מנה.
בדרך כלל המשתנה המוצג כ- Y הוא המשתנה התלוי והמשתנה המוצג ב- X הוא המשתנה הבלתי תלוי.
תיאור גרפי לנתונים נעשה על ידי דיאגרמת פיזור. בדיאגרמת פיזור אנחנו מסמנים כל תצפית בנקודה לפי שיעור ה- X ושיעור ה- Y שלה.
דיאגרמת הפיזור נותנת אינדיקציה גרפית על הקשר בין שני המשתנים.

דוגמה (פתרון בהקלטה) :

בבניין 8 דירות בדקו לכל דירה את מספר החדרים שלה וכמו כן את מספר הנפשות הגרות בדירה. להלן התוצאות שהתקבלו :

מספר חדרים בדירה	2	2	3	2	3	3	4	4
מספר הנפשות בדירה	0	1	2	2	2	3	4	5

1) כמה תצפיות ישנן בדוגמה?

2) כמה משתנים ישנם בדוגמה, מי הם?

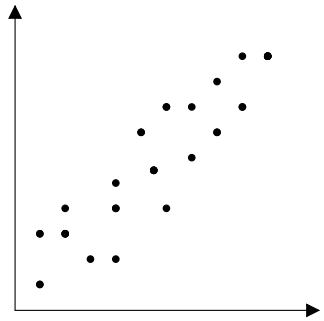
3) שרטטו לנתונים דיאגרמת פיזור.

4) מי המשתנה התלוי ומיהו המשתנה הבלתי תלוי?

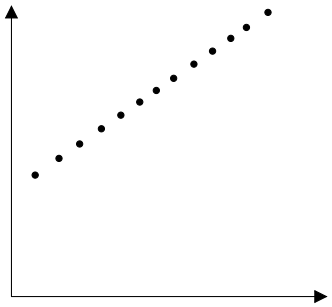
© כל הזכויות שמורות לברק קנדל

דיאגרמות פיזור לקשר בין משתנים וניתוחם

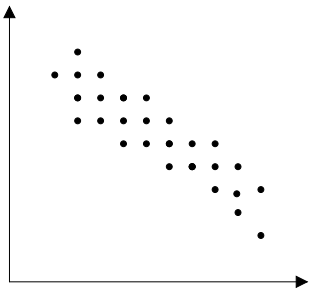
קשר לנארי חיובי חלקי



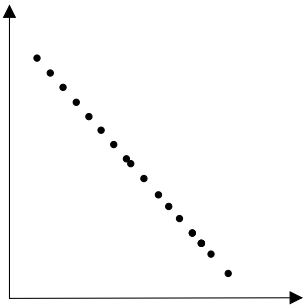
קשר לנארי חיובי מלא



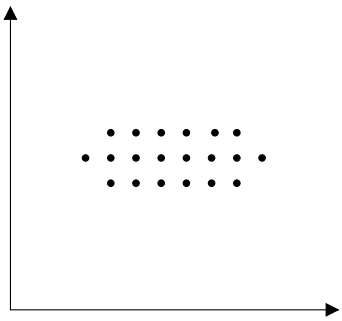
קשר לינארי שלילי חלקי



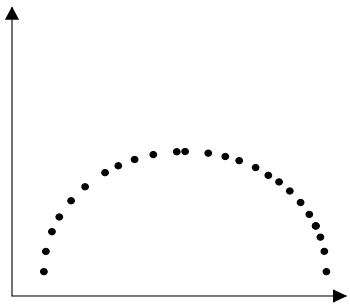
קשר לינארי שלילי מלא



אין קשר



אין קשר לינארי



משמעות מקדם המתאם:

כדי לבדוק עד כמה קיים קשר לינארי בין שני המשתנים ישנו מדד קשר שנקרא גם מקדם המתאם הלינארי הידוע גם בשם מקדם המתאם של פירסון.

מקדם מתאם זה מקבל ערכים בין 1 ל-1.

-1 0 1

© כל הזכויות שמורות לברק קנדל

מקדם מתאם 1-או 1 אומר שקיים קשר לינארי מלא בין המשתנים שניתן לבטאו על ידי נוסחה של קו ישר: $y = ax + b$.

מתאם חיובי מלא (מקדם מתאם 1):

קיים קשר לנארי מלא בו השיפוע a יהיה חיובי ואילו מתאם שלילי (מקדם מתאם-1) מלא אומר שקיים קשר לנארי מלא בו השיפוע a שלילי.

מתאם חיובי חלקי:

ככל שמשתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט ואילו מתאם שלילי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לרדת אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט. ככל שמקדם המתאם קרוב לאפס עוצמת הקשר יותר חלשה וככל שהמדד רחוק יותר מהאפס העוצמה יותר חזקה. לסיכום, מקדם המתאם בודק את עוצמת הקשר הלנארי, ואת כיוון הקשר.

מקדם המתאם הלינארי אינו מושפע מיחידות המדידה. כל שינוי ביחידות המדידה של המשתנים, לא ישנה את מקדם המתאם.

מדד הקשר הלינארי באוכלוסייה, שנקרא גם מקדם המתאם של פירסון או מדד הקשר של פירסון באוכלוסייה מסומן ב: ρ - פרמטר המאפיין את עוצמת הקשר הלינארי באוכלוסייה וכיוונו בין שני המשתנים הנחקרים. כאשר:

r - מדד הקשר הלינארי במדגם שמהווה אומד לפרמטר ρ.

קיומו של מתאם בין שני משתנים אינו מצביע על סיבתיות בהכרח. למשל, אם נמצא מתאם חיובי בין כמות הסוכרזית שאדם אוכל לבין במשקל שלו אין זה אומר שהסיבה להשמנה היא הסוכרזית. מדד הקשר של פירסון הוא מדד קשר סימטרי, כלומר אם נחליף את X ב- Y התוצאה תהיה זהה.

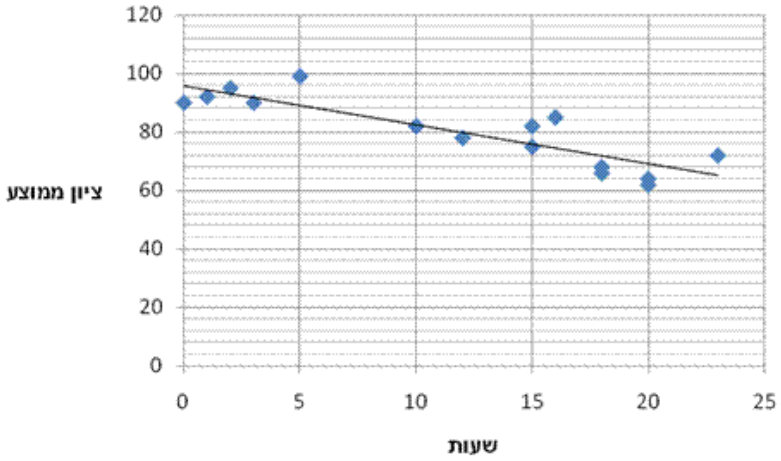
דוגמה (פתרון בהקלטה):

- מה ניתן להגיד על מקדם המתאם של שני המשתנים על סמך דיאגרמת הפיזור ששרטטנו?
- אם היינו משנים את השרטוט כך שבציר האנכי היה המשתנה "מספר החדרים" ובציר האופקי היה "מספר הנפשות", האם הדבר היה משפיע על מדד הקשר של פירסון?

© כל הזכויות שמורות לברק קנדל

שאלות

1) חוקר רצה לאפיין את הקשר בין מספר השעות בשבוע שסטודנט מקדיש לבילויים לבין הציון הממוצע שלו בסוף הסמסטר. לשם כך הוא אסף נתונים של 15 סטודנטים ויצר דיאגרמת פיזור:



א. מיהו המשתנה הבלתי תלוי?

ב. מה ניתן לומר על כיוון הקשר בין מספר שעות הבילוי השבועיות לבין הציון הממוצע של הסמסטר? מה ניתן להגיד על עוצמת הקשר?

2) להלן טבלה המסכמת את מקדמי המתאם הלינארי בין ציוני מבחנים שונים שהתקבלו עבור תלמידים בכיתה מסוימת:

	ספורט	לשון	מתמטיקה
ספורט	?	-0.7	?
לשון	?	?	0.6
מתמטיקה	-0.1	?	?

א. השלימו את מקדמי המתאם שמסומנים בסימן שאלה בטבלה.

ב. בין אילו שני ציוני מקצועות שונים קיים מתאם בעל העוצמה החזקה ביותר?

3) במחקר נתבקשו לבדוק את הקשר בין מספר שעות התרגול של קורס לבין הציון הסופי שלו. להלן תוצאות מדגם שהתקבל:

שעות תרגול	ציון סופי
20	90
25	90
30	95
15	60
30	90
20	85
10	50

א. מיהו המשתנה התלוי ומיהו המשתנה הבלתי תלוי בדוגמה זו?

ב. שרטטו דיאגרמת פיזור לנתונים.

ג. מה ניתן לומר על הקשר בין המשתנים במדגם?

ד. מסתבר שבסופו של דבר נתנו פקטור של 5 נקודות לציון הסופי. כיצד הדבר היה משנה את מקדם המתאם של המדגם?

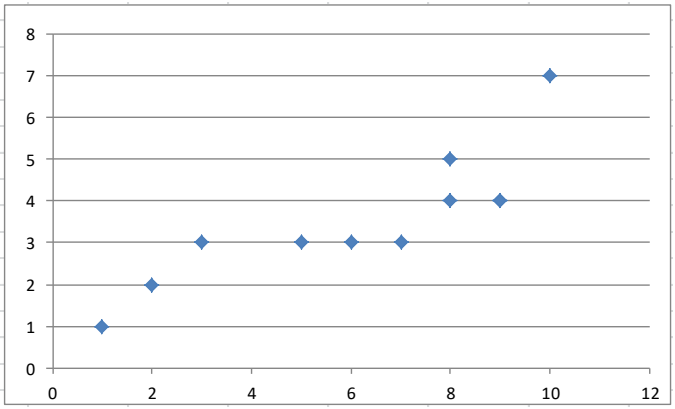
© כל הזכויות שמורות לברק קנדל

4) בתחנה המטאורולוגית רצו לבדוק את הקשר שבין הטמפרטורה במעלות צלזיוס לכמות המשקעים במ״מ. הם אספו נתונים על 10 ימים במהלך חודש ינואר. המתאם שהתקבל היה 0.8.

א. השלימו את המשפט:
בחודש ינואר ככל שהטמפרטורה היומית נוטה לרדת, כך כמות המשקעים נוטה _____.

ב. הוחלט להעביר את הטמפרטורה למעלות פרנהייט על מנת שיוכלו להשוות אותה לנתונים מארה״ב. נוסחת המעבר היא $F^0 = 32 + \frac{9}{5}C^0$.
כיצד הדבר ישפיע על מקדם המתאם בין הטמפרטורה במעלות פרנייט לכמות המשקעים במ״מ?

5) להלן דיאגרמת פיזור המראה קשר בין שני משנים:



א. השלימו: ניתן לראות שהקשר הוא לינארי _____ (מלא\ חלקי) כיוון הקשר הוא (חיובי \שלילי).

ב. השלימו: אם היינו מוסיפים תצפית שערך ה- X שלה הוא 4 וערך ה- Y שלה הוא 7, מקדם המתאם של פירסון היה _____ (גדל\ קטן\ לא משתנה).

שאלות רב ברירה (יש לבחור את התשובה הנכונה):

6) חוקר אקלים דגם כמה ימים בשנה ומדד את הטמפרטורה בטורונטו שבקנדה ואת הטמפרטורה בסידני שבאוסטרליה באותו היום. הוא חישב ומצא מקדם מתאם שלילי בין הטמפרטורה היומית בטורונטו לבין הטמפרטורה היומית בסידני. משמעות מקדם המתאם השלילי במדגם:

א. אין קשר בין הטמפרטורה בטורונטו לבין הטמפרטורה בסידני בימים שנדגמו.

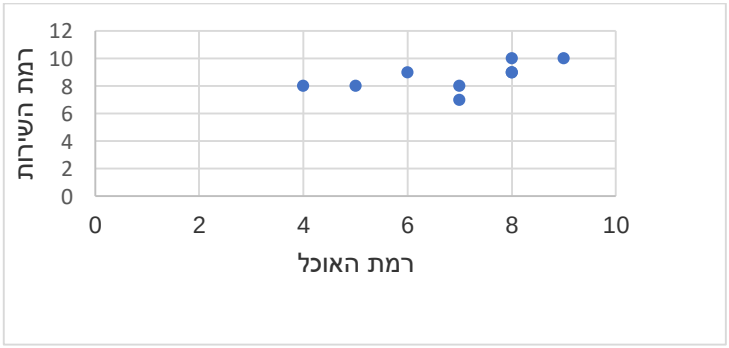
ב. במדגם, רוב הטמפרטורות בטורונטו היו שליליות.

ג. ההפרש בין הטמפרטורה בטורונטו לבין הטמפרטורה באוסטרליה, במדגם זה, הוא שלילי.

ד. במדגם יש נטייה שהטמפרטורה יורדת בטורונטו לטמפרטורה לעלות בסידני.

© כל הזכויות שמורות לברק קנדל

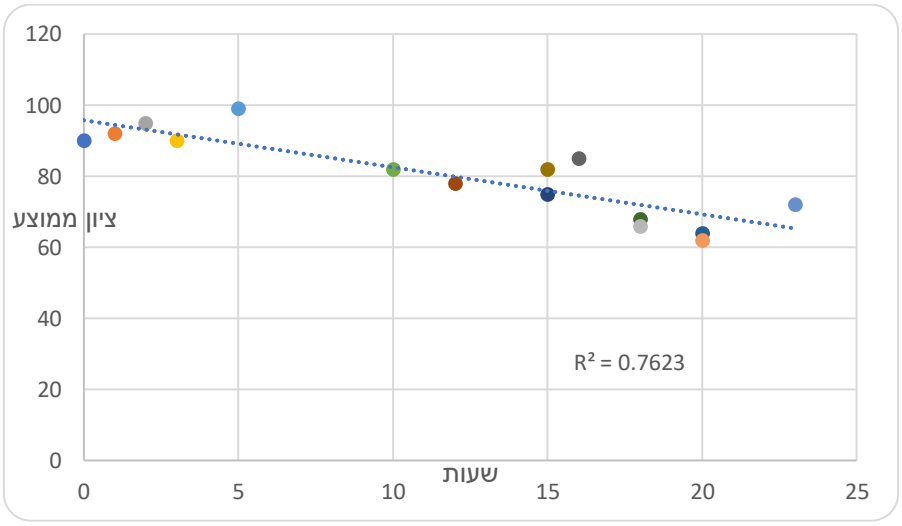
7) בסקר שביעות רצון שנערך בבית הקפה "פת לחם" התבקשו הלקוחות לדרג את מידת שביעות הרצון שלהם (בסולם 1-10) בשני נושאים: רמת האוכל ורמת השירות.



מה יהיה ערכו של מקדם המתאם (r)?

א. $r=-0.3$
ב. $r=0$
ג. $r=1.125$
ד. $r=0.593$

8) חוקר רצה לאפיין את הקשר בין מספר השעות בשבוע שסטודנט מקדיש לבילויים לבין הציון הממוצע שלו בסוף הסמסטר. לשם כך הוא אסף נתונים של 15 סטודנטים ויצר דיאגרמת פיזור.

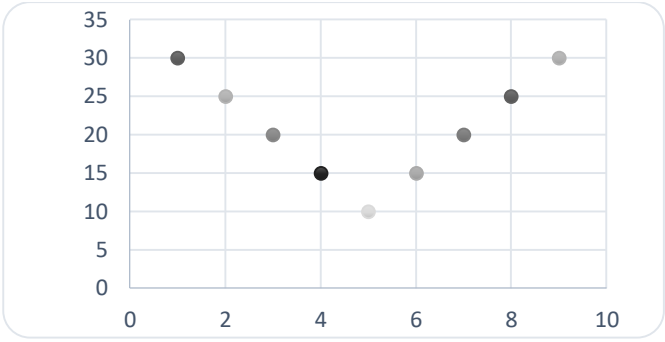


מה ניתן לומר על כיוון הקשר במדגם בין מספר שעות הבילוי השבועיות לבין הציון הממוצע של הסמסטר?

א. ככל שמבלים יותר הציון נוטה לרדת.
ב. אין קשר בין שעות הבילוי לציון.
ג. ככל שמבלים פחות הציון נוטה לרדת.
ד. ככל שהציון נוטה לרדת הסטודנט מבלה פחות.

© כל הזכויות שמורות לברק קנדל

9) התרשים הבא מתאר קשר בין שני משתנים, איזה מהמתאמים הבאים הוא המתאים ביותר לתיאור הקשר בין שני המשתנים?



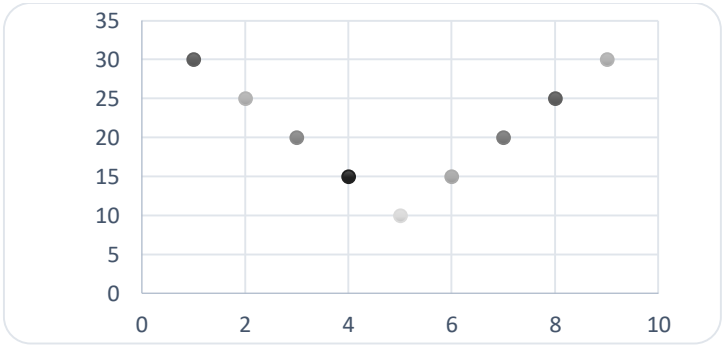
א. $r = 1$ היות ושני המשתנים יוצרים קוים ישרים.

ב. $r = 2$ היות ויש שני קוים בעלי קשר מושלם.

ג. $r = 0$ היות והקו יורד ואחר כך עולה באותו האופן.

ד. $r = \pm 1$ היות ויש קו עולה וגם קו יורד.

10) התרשים הבא מתאר דיאגרמת פיזור.



איזו טענה נכונה?

א. בתרשים מוצג הקשר בין שני משתנים.

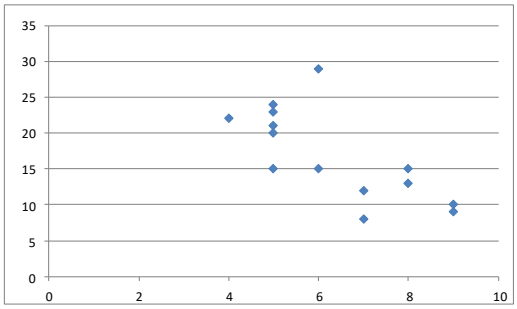
ב. בתרשים מוצג הקשר בין 9 משתנים.

ג. בתרשים מוצג הקשר בין 10 משתנים.

ד. אין לדעת כמה משתנים מוצגים בתרשים.

© כל הזכויות שמורות לברק קנדל

בגרף הבא מתוארת דיאגרמת פיזור של שני משתנים:



X - (משתנה בלתי תלוי בציר האופקי)
ו- Y (משתנה תלוי).
במדגם התקבל $r^2 = 0.52$.

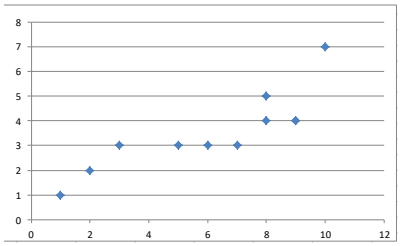
11) לאור הנתונים המופיעים בדיאגרמה, איזה מבין הערכים הבאים מתאים להיות התוצאה של r?

א. -0.52
ב. 0.72
ג. -0.72
ד. 0.52

12) אם מקדם המתאם בין שני משתנים הוא 1, אזי:

א. הערכים של המשתנים הם חיוביים.
ב. עבור כל תצפית ערך של משתנה אחד שווה לערך של המשתנה השני.
ג. הקשר הלינארי הוא בעוצמה חזקה.
ד. אף אחת מהתשובות לא בהכרח נכונה.

13) להלן דיאגרמת פיזור:



מה יהיה מקדם המתאם בין שני המשתנים?

א. 1
ב. 0.85
ג. 0.15
ד. 0

14) בבדיקת קשר בין שני משתנים התקבל: $r = -1$.

א. קיימת נוסחה לינארית הקושרת בין כל התצפיות.
ב. לא קיים קשר בין שני המשתנים.
ג. ככל שמשתנה אחד נוטה לרדת גם לשני יש נטייה לרדת.
ד. קיים קשר בין שני המשתנים, אך לא ניתן לדעת מאיזה סוג.

© כל הזכויות שמורות לברק קנדל

15) לפי הפתגם "רחוק מהעין, רחוק מהלב", יש קשר ____ בין קרבה פיזית לקרבה נפשית.

א. חיובי

ב. שלילי

ג. אפסי

ד. לא ניתן לדעת.

16) מבחן אמי"ר הינו מבחן מיון באנגלית של המרכז הארצי לבחינות והערכה. הציון המינימלי בבחינה הינו 150 והמקסימלי הינו 250. בקורס הכנה למבחן השתתפו 19 תלמידים. להלן הציונים שלהם על פי פלט שהתקבל:

		159
		170
		180
		185
		204
		224
		236
		212
		168
		189
		195
		163
		187
		206
		201
		223
		242
		203
		205
197.47	AVERAGE	
536.25	VARPA	

יש להוסיף עמודה נוספת לצד עמודת הציונים שתראה לכל תלמיד כמה נקודות חסרות לו כדי להשלים לציון המקסימלי בבחינה.

מה יהיה מקדם המתאם בין שתי העמודות (כלומר, מקדם המתאם בין הציון לבין הנקודות החסרות)?

א. -1

ב. 1

ג. -0.5

ד. 0.5

17) מקדם המתאם בין שטחי דירה למחיר שלהם חושב ונמצא 1.2. מה נובע מכך?

א. ככל שהדירה גדולה יותר בשטחה כך היא יקרה יותר.

ב. ככל שהדירה קטנה יותר בשטחה כך היא זולה יותר.

ג. לא קיים קשר בין שטח הדירה למחיר הדירה.

ד. מצב כזה שמתואר הנתונים לא אפשרי.

18) אם ניקח 10 אנשים ונרשום לכל אדם את הגובה במטר וכמו כן את הגובה בס"מ. מה יהיה מקדם המתאם בין גובה האדם במטר לגובה האדם בס"מ?

א. 1

ב. 0

ג. -1

ד. לא ניתן לדעת.

© כל הזכויות שמורות לברק קנדל

19) נמצא מתאם חיובי בעוצמה גבוהה בין X – ציון בבגרות בלשון ל Y – ציון בבגרות במתמטיקה. אילו מהמשפטים הבאים נכון?

א. ניתן לומר שאחת מהסיבות להבדלים שיש לסטודנטים במתמטיקה נובעים מההבדלים שיש להם בלשון.

ב. קיימת נוסחה של קו ישר שקושרת בין ציון בבגרות במתמטיקה לציון בבגרות בלשון.

ג. ללא יוצא מן הכלל, ניתן להגיד שכל תלמיד שמצליח יותר מתלמיד אחר בלשון גם יצליח יותר מאותו תלמיד במתמטיקה.

ד. אף אחד מהטענות שהוצגו אינה בהכרח נכונה.

20) עבור סדרה של תצפיות מדדו את X ואת Y. נמצא שעבור כל התצפיות שהערך של Y ירד הערך של X בהכרח ירד ללא יוצא מן הכלל. מקדם המתאם של פירסון יהיה בהכרח:

א. 1

ב. -1

ג. 0

ד. אף אחת מהתשובות.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות

1) א. שעות בילוי. ב. הקשר חלקי, כיוון הקשר שלילי .

2) א. להלן טבלה:

	ספורט	לשון	מתמטיקה
ספורט	1	-0.7	0.1
לשון	-0.7	1	0.6
מתמטיקה	-0.1	0.6	1

ב. ספורט ולשון.

3) א. ב"ת- מס' שעות התרגול, תלוי- ציון. ב. ראה גרף בפתרון וידאו.
ג. קשר לינארי חיובי חלקי. ד. מקדם המתאם לא היה משתנה.

4) א. לעלות. ב. לא ישפיע על מקדם המתאם.

5) א. חלקי, חיובי. ב. קטן.

6) ד'. **7)** ד'. **8)** א'. **9)** ג'. **10)** א'.

11) ג'. **12)** ד'. **13)** ב'. **14)** א'. **15)** א'.

16) א'. **17)** ד'. **18)** א'. **19)** ד'. **20)** ד'.

© כל הזכויות שמורות לברק קנדל

מדדי קשר – מדד הקשר הלינארי (פירסון) – רקע

המטרה היא לבדוק האם קיים קשר (קורלציה, מתאם) של קו ישר בין שני משתנים כמותיים. מבחינת סולמות המדידה קשר בין סולמות רווחים ומנה.
בדרך כלל, X הוא המשתנה המסביר (הבלתי תלוי) ו- Y הוא המשתנה המוסבר (התלוי).

דוגמה:

נרצה להסביר כיצד השכלה של אדם הנמדדת בשנות לימוד – X מסבירה את ההכנסה שלו Y. במקרה זה שנות ההשכלה זהו המשתנה המסביר (או הבלתי תלוי) ואנחנו מעוניינים לבדוק כיצד שינויים בשנות ההשכלה של אדם יכולים להסביר את השינויים שלו בהכנסה, ולכן רמת ההכנסה זהו המשתנה המוסבר התלוי במשתנה המסביר אותו.

שלב ראשון: נהוג לשרטט דיאגרמת פיזור. זו דיאגרמה שנותנת אינדיקציה ויזואלית על טיב הקשר בין שני המשתנים.

דוגמה:

בבניין של 5 דירות בדקו את הנתונים הבאים:
X -מס' חדרים בדירה. Y-מס' נפשות הגרות בדירה.
להלן התוצאות שהתקבלו:

מס' דירה	X	Y
1	3	2
2	2	2
3	4	3
4	3	3
5	5	4

נשרטט מנתונים אלה דיאגרמת פיזור (הדיאגרמה המלאה בסרטון). נתבונן בכמה מקרים של דיאגרמות פיזור וננתח אותן (הדיאגרמות המלאות בסרטון).

שלב שני: מחשבים את מקדם המתאם (מדד הקשר) שבודק עד כמה קיים קשר לינארי בין שני המשתנים. המדד (ניקרא גם מדד הקשר של פירסון) מכמת את מה שניראה בשלב הראשון רק בעין.
המדד בודק את כיוון הקשר (חיובי או שלילי) ואת עוצמת הקשר (חלש עד חזק).
מקדם מתאם זה מקבל ערכים בין 1- ל-1.
מקדם מתאם 1- או 1 אומר שקיים קשר לינארי מוחלט ומלא בין המשתנים שניתן לבטאו על ידי הנוסחה: $y = bx + a$.

מתאם חיובי מלא (מקדם מתאם 1):

קיים קשר לנארי מלא בו השיפוע b יהיה חיובי ואילו מתאם שלילי מלא אומר שקיים קשר לנארי מלא בו השיפוע b שלילי (מקדם מתאם 1-).

© כל הזכויות שמורות לברק קנדל

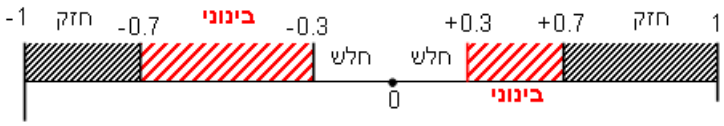
מתאם חיובי חלקי:

ככל שמשתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט.

מתאם שלילי חלקי:

ככל שמשתנה אחד עולה לשני יש נטייה לרדת אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט.

ככל שערך מקדם המתאם קרוב לאפס נאמר שעוצמת הקשר חלשה יותר וככל שמקדם המתאם רחוק מהאפס נאמר שעוצמת הקשר חזקה יותר:



מקדם המתאם יסומן באות r.

כדי לחשב את מקדם המתאם, יש לחשב את סטיות התקן של כל משתנה ואת השונות המשותפת.

שונות משותפת: $COV(x,y)=\frac{\sum(x-\bar{x})(y-\bar{y})}{n}=\frac{\sum xy}{n}-\bar{x}\cdot\bar{y}$

שונות של המשתנה X: $s_x^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}=\frac{\sum_{i=1}^{n}x_i^2}{n}-\bar{x}^2$

שונות המשתנה Y: $S_Y^2=\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n}=\frac{\sum_{i=1}^{n}y_i^2}{n}-\bar{y}^2$

מקדם המתאם הלינארי: $r_{xy}=\frac{COV(x,y)}{S_x\cdot S_y}$

© כל הזכויות שמורות לברק קנדל

שאלות

1) להלן נתונים לגבי שישה תלמידים שנגשו למבחן. בדקו לגבי כל תלמיד את הציון שלו בסוף הקורס וכמו כן את מספר החיסורים שלו מהקורס.

מספר חיסורים	2	1	0	2	3	4
ציון	80	90	90	70	70	50

א. שרטטו דיאגראמת פיזור לנתונים. מה ניתן להסיק מהדיאגרמה על טיב הקשר ביו מספר החיסורים של תלמיד לציונו? מיהו המשתנה הבלתי תלוי ומיהו המשתנה התלוי?

ב. חשבו את מדד הקשר של פירסון. האם התוצאה מתיישבת עם תשובתך לסעיף א'?

ג. הסבירו, ללא חישוב, כיצד מקדם המתאם היה משתנה אם היה מתווסף תלמיד שהחסיר 4 פעמים וקיבל ציון 80?

2) במחקר רפואי רצו לבדוק האם קיים קשר בין רמת ההורמון X בדם החולה לרמת ההורמון Y שלו. לצורך כך מדדו את רמת ההורמונים ההלו עבור חמישה חולים. להלן התוצאות שהתקבלו:

X	Y
10	12
14	15
15	15
18	17
20	21

א. מה הממוצע של כל רמת הורמון?

ב. מהו מקדם המתאם בין ההורמונים? ומה משמעות התוצאה?

3) נסמן ב- X את ההכנסה של משפחה באלפי ₪. נסמן ב- Y את ההוצאות של משפחה באלפי ₪. נלקחו 20 משפחות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{20} X_i = 240 \qquad \sum_{i=1}^{20} Y_i = 200$$

$$\sum_{i-1}^{20} (X_i - \bar{X})^2 = 76 \qquad \sum_{i-1}^{20} (Y_i - \bar{Y})^2 = 76$$

$$\sum_{i=1}^{20} (X_i - \bar{X})(Y - \bar{Y}) = 60.8$$

א. חשב את מדד הקשר הלינארי בין X ל- Y. מיהו המשתנה התלוי?

ב. מה המשמעות של התוצאה שקיבלת בסעיף א?

© כל הזכויות שמורות לברק קנדל

4) נסמן ב- X את ההכנסה של משפחה באלפי ₪. נסמן ב- Y את ההוצאות של משפחה באלפי ₪. נלקחו 20 משפחות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{20} X_i = 240 \qquad \sum_{i=1}^{20} Y_i = 200$$

$$\sum_{i-1}^{20} X_i^2 = 2960 \qquad \sum_{i=1}^{20} Y_i^2 = 2080$$

$$\sum_{i=1}^{20} X_i Y_i = 2464$$

חשבו את מדד הקשר הלינארי בין X ל- Y.

5) במוסד אקדמי ציון ההתאמה מחושב כך: מכפילים את הציון הממוצע בבגרות ב-3 ומפחיתים 2 נקודות. ידוע שעבור 40 מועמדים סטיית התקן של ממוצע הציון בבגרות הייתה 2.
מה מקדם המתאם בין ציון ההתאמה לציון הממוצע בבגרות שלהם?

6) להלן רשימת טענות, לגבי כל טענה קבעו נכון/לא נכון ונמקו.

א. מתווך דירות המיר מחירי דירות מדולר לשקל. נניח שדולר אחד הוא 3.5₪. אם מתווך הדירות יחשב את מדד הקשר של פירסון בין מחיר הדירה בשקלים למחיר הדירה בדולרים הוא יקבל 1.

ב. לסדרה של נתונים התקבל $\bar{X} = \bar{Y} = 6$, $S_x = S_y = 1$.
לכן, מדד הקשר של פירסון יהיה 1.

ג. אם השונות המשותפת של X ושל Y הינה 0 אז בהכרח גם מקדם המתאם של פירסון יהיה 0.

שאלות רב-ברירה:

7) נמצא שקיים מקדם מתאם שלילי בין הציון בעברית לציון בחשבון בבחינה לכן:

א. הדבר מעיד שהציונים בכתה היו שליליים.

ב. ככל שהציון של תלמיד יורד בחשבון יש לו נטייה לרדת בעברית.

ג. ככל שהציון של תלמיד עולה בחשבון יש לו נטייה לרדת בעברית.

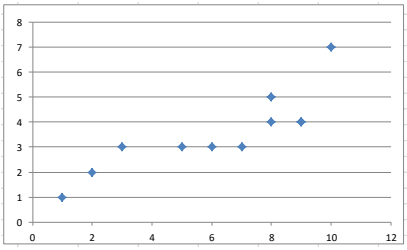
ד. אף אחת מהתשובות לא נכונה.

© כל הזכויות שמורות לברק קנדל

8) נלקחו 20 מוצרים ונבדק ביום מסוים המחיר שלהם בדולרים והמחיר שלהם בש"ח (באותו היום ערך הדולר היה-4.2₪).
מהו מקדם המתאם בין המחיר בדולר למחיר בש"ח?

א. 1
ב. 0
ג. 4.2
ד. לא ניתן לדעת.

9) להלן דיאגראמת פיזור:
מה יהיה מקדם המתאם בין שני המשתנים?



א. 1
ב. 0.85
ג. 0.15
ד. 0

תשובות סופיות

1) א. משתנה תלוי: ציון, משתנה ב"ת: מס' חיסורים. ראה דיאגרמה בוידאו. ניתן להסיק שקיים קשר לינארי שלילי וחלקי בין מספר החיסורים לציון התלמיד.
ב. 0.9325-. ג. הקשר יישאר לינארי שלילי חלקי אך עוצמתו תחלש.

2) א. $\bar{y}=16$, $\bar{x}=15.4$ ב. $r_{xy}=0.96$.

3) א. 0.8.

4) 0.8.

5) 1.

6) א. נכון. ב. לא נכון. ג. נכון.

7) ג'.

8) א'.

9) ב'.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 14 - רגרסיה

תוכן העניינים

1. כללי 65

מדדי קשר – רגרסיה ליניארית:

רקע:

במידה וקיים קשר חזק בין שני המשתנים הכמותיים נהוג לבצע ניבוי. לבנות קו ניבויים הנקרא גם קו רגרסיה המנבא משתנה אחד על סמך האחר. מדובר בקו שמנבא את Y על סמך X.
השיטה למציאת הקו הנ"ל נקראת שיטת הריבועים הפחותים והקו המתקבל נקרא קו הרגרסיה או קו הניבויים או קו הריבועים הפחותים.

a - נותן את ערך Y כאשר X הנו אפס על גבי קו הניבויים. הוא נקרא החותך של הקו.

b - הוא שיפוע הקו נותן בכמה בעצם Y משתנה כאשר X גדל ביחידה אחת על גבי קו הניבויים.

להלן המשוואות למציאת הפרמטרים של קו הרגרסיה: $b = r\frac{S_r}{S_x}$, $Y = bX + a$.

לצורך בניית קו ניבויים לניבוי X על סמך Y נצטרך לעדכן את הנוסחאות בהתאם.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) נסמן ב- X את ההכנסה של משפחה באלפי ₪. נסמן ב- Y את ההוצאות של משפחה באלפי ₪. נלקחו 20 משפחות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{20} Y_i = 200 \ , \ \sum_{i=1}^{20} X_i = 240$$

$$\sum_{i=1}^{20} \left(X_i - \overline{X} \right)^2 = 76 \ , \ \sum_{i=1}^{20} \left(Y_i - \overline{Y} \right) = 76 \ .$$

$$\sum_{i=1}^{20} \left(X_i - \overline{X} \right)\left(Y_i - \overline{Y} \right) = 60.8$$

א. חשבו את מדד הקשר הלינארי בין X ל- Y. מיהו המשתנה התלוי?
ב. מצאו את קו הרגרסיה לניבוי ההוצאה של משפחה על סמך הכנסה שלה. הסבירו את משמעות הפרמטרים של קו הרגרסיה.
ג. משפחת כהן הכניסה 15,000₪. מה ההוצאה הצפויה שלה?

2) נסמן ב- X את ההשכלה של אדם בשנות לימוד.
נסמן ב- Y את הכנסתו באלפי ₪. במחקר התקבלו התוצאות הבאות:

$$\mathrm{COV}\left(X,Y\right) = 7.5 \ , \ \overline{Y} = 8 \ , \ \overline{X} = 14 \ , \ S_Y = 5 \ , \ S_X = 2 \ .$$

א. חשבו את מדד הקשר של פירסון בין ההשכלה להכנסה.
ב. מה ההכנסה הצפויה לאדם שהשכלתו 12 שנים?
ג. מה ההשכלה הצפויה לאדם שהכנסתו 10,000₪?

3) חוקר רצה לחקור את הקשר הקווי שבין הציון המבחן בסטטיסטיקה לבין מספר שעות ההכנה של הסטודנטים למבחן. במדגם של 100 סטודנטים שנבחנו בקורס נרשמו התוצאות הבאות: הציון הממוצע של הסטודנטים היה 65 עם סטיית תקן של 27. מספר שעות ההכנה הממוצע היה 30 עם סטיית תקן של 18. מקדם המתאם בין הציון לשעות ההכנה היה 0.8.

א. על פי משוואת הרגרסיה, שעת הכנה נוספת משפרת את ציון המבחן ב-?
ב. על פי משוואת הרגרסיה, תלמיד שייגש למבחן ללא שעות הכנה כלל יקבל ציון?
ג. מהו קו הרגרסיה לניבוי הציון לפי שעות ההכנה?

4) נתונים 2 משתנים X ו- Y. כמו כן נתון: $\overline{X} = 1.5$, $S_x = S_y = 4$,
וכן שקו הרגרסיה של Y על בסיס X הינו: $Y = -0.2X + 0.5$.
חשבו מהו מקדם המתאם בין X ל- Y.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. 0.8. ב. $Y=0.8X+0.4$. ג. 12.4 אלפי ₪.

2) א. 0.75. ב. 4.25 אלפי ₪. ג. 14.6 שנים.

3) א. 1.2. ב. 29. ג. $Y=1.2X+29$.

4) -0.2.

© כל הזכויות שמורות לברק קנדל

סטטיסטיקה בניהול

פרק 15 - מדדי קשר - תרגול בחירת מדד מתאים

תוכן העניינים

1. בחירת מדד מתאים 68

מדדי קשר – בחירת מדד מתאים:

רקע:

בפרק זה נתרגל את התהליך של בחירת מדד הקשר (מקדם המתאם) המתאים.
נתרכז בשלושת מדדי הקשר הנפוצים ביותר:

- מדד הקשר של קרמר.
- מדד הקשר של ספירמן.
- מדד הקשר של פירסון (מדד הקשר הלינארי).

בחירת מדד הקשר נעשה לפי סולמות המדידה של שני המשתנים שאנחנו רוצים לבדוק את הקשר בינם. הנושא של סולמות מדידה נלמד כבר בפרק אחר, כמו כן כל מדד קשר נלמד בפרק נפרד. אנו מתרכזים ב 3 סולמות מדידה:

- סולם שמי/ זהות (nominal).
- סולם סדר (ordinal).
- סולם כמותי (scale): לכאן אנו מאחדים את סולם רווחים ומנה יחד.

שלושת מדדי הקשר שלעיל דנים בקשר בין שני משתנים. מדדי הקשר הם סימטריים, כלומר אין זה משנה איזה משתנה נגדיר בתור משתנה X ואיזה יוגדר בתור משתנה Y.

להלן טבלה שמסכמת את בחירת המדד המתאים:

X / Y	שמי	סדר	כמותי
שמי	קרמר	קרמר	קרמר
סדר	קרמר	ספירמן	ספירמן
כמותי	קרמר	ספירמן	פירסון

© כל הזכויות שמורות לברק קנדל

דוגמה (פתרון בהקלטה):

במפעל לייצור מצברים לרכב בדקו במשך 40 ימים את התפוקה היומית (מספר מצברים במאות) ואת מספר הפועלים שעבדו באותו היום.
איזה מדד קשר מתאים כדי לבדוק האם קיים קשר בין התפוקה היומית לכמות עובדים באותו היום במפעל?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. חסרים נתונים כדי לדעת זאת.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) בקרב תלמידי כיתות א' בבית הספר גבריאלי אשר בתל אביב בדקו לכל תלמיד את גובהו בס"מ ואת משקלו בק"ג. מהו מדד הקשר המתאים כדי לבדוק האם קיים קשר בין גובה התלמיד למשקלו?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

2) בסקר שנעשה על אזרחים במדינה בדקו לכל אזרח את השכלתו ואת שכרו. מהו מדד הקשר המתאים כדי לבדוק האם קיים קשר בין השכלה לשכר?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

3) בגן הילדים של שולה אספו נתונים על 25 הילדים שבגן. על כל ילד בדקו את רמת הביטחון העצמי שלו (X = מדד שמקבל ערכים בין 1 - נמוך ועד 5 - גבוה), ואת אוצר המילים שלו (Y = לפי מבחן שנעשה לכל ילד בו ספרו את מספר המילים שידע מתוך רשימה של 20 מילים). איסוף הנתונים נעשה על ידי איש מקצוע שצפה בילדים ובחן אותן.
מהו מקדם המתאם המתאים לבדיקת התלות בין X לבין Y?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

4) הועלתה השערה בבית הספר למדעי ההתנהגות שיש קשר בין המרצה להצלחת הסטודנט. לצורך בדיקת הטענה בדקו לגבי כל סטודנט שלמד סטטיסטיקה אצל איזה מרצה הוא למד (היו 3 מרצים שונים) והאם הוא עבר את הבחינה.
מהו מדד הקשר המתאים במקרה זה?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

© כל הזכויות שמורות לברק קנדל

5) בבורסה בתל אביב רצו לבדוק את הקשר בין גובה הריבית במשק בסוף החודש (באחוזים), לבין תשואת מניית אקטר (באחוזים) בסוף החודש.
מהו מדד הקשר המתאים?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

6) בעל מסעדה ביצע סקר על לקוחותיו, בין השאלות שנשאלו בסקר:

- מה מידת שביעות הרצון של הלקוח מאדיבות השירות של המלצר בסקלה של 1 עד 5.
- מה גילו של הלקוח בשנים.
- מה גובה התשר (טיפ) ב-₪ אשר נתן הלקוח למלצר בלכתו מהמסעדה.

מהו המדד המתאים כדי לבדוק האם קיים מתאם חיובי בין מידת שביעות הרצון של הלקוח מאדיבות השירות לבין גובה התשר שהוא נתן למלצר?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

7) בעל מסעדה ביצע סקר על לקוחותיו, בין השאלות שנשאלו בסקר:

- מה מידת שביעות הרצון של הלקוח מאדיבות השירות של המלצר בסקלה של 1 עד 5.
- מה גילו של הלקוח בשנים.
- מה גובה התשר (טיפ) ב-₪ אשר נתן הלקוח למלצר בלכתו מהמסעדה.

מהו המדד המתאים כדי לבדוק האם קיים מתאם בין גיל הלקוח לגובה התשר שהעניק לשירות?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

© כל הזכויות שמורות לברק קנדל

הנתונים הבאים מתאימים ל-3 השאלות הבאות:

חוקרים ערכו מדגם של ילדים מכיתות ב' ו-ג' מ-4 בתי ספר שונים. הועבר לילדים שאלון בו תואר מצב מסוים והילדים התבקשו לציין את רמת החרדה שלהם באשר לאותו מצב. המשתנים שלגביהם נאספו נתונים:

- מגדר (1 - בן, 2 - בת).
- כיתה (0 - ג', 1 - ב').
- בית ספר (A, B, C, D).
- רמת חרדה (ציון שהילד היה צריך לתת בסקלה של 1 עד 10).
- גיל התלמיד בחודשים.

8) מהו מדד הקשר המתאים כדי לבדוק את הקשר בין גיל התלמיד לבין רמת החרדה שלו מהמצב?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

9) מהו מדד הקשר המתאים כדי לבדוק את הקשר בין המגדר לבין רמת החרדה שלו מהמצב?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

10) מהו מדד הקשר המתאים כדי לבדוק את הקשר בין המגדר לבין בית הספר?

א. פירסון.
ב. ספירמן.
ג. קרמר.
ד. אין מספיק נתונים כדי לדעת.

© כל הזכויות שמורות לברק קנדל

11) בטבלה שלהלן נתונות שכיחויות ההצלחה והכישלון של 150 חולים:

תוצאה/ התרופה	A	B	C
נרפא	35	13	45
לא נרפא	15	37	5

החולים קיבלו 3 תרופות שונות ובדקו עבור כל חולה אם התרופה הצליחה בריפוי. מהו מדד הקשר המתאים?

א. פירסון.

ב. ספירמן.

ג. קרמר.

ד. אין מספיק נתונים כדי לדעת.

12) שני מוסיקאים מפורסמים נתנו ציון בסולם של 1-10 לקולם של 8 מתמודדים בתוכנית ראליטי ידועה. ציון 10 ניתן לקול שמצא חן ביותר בעיני המוסיקאי. מפיק התוכנית רצה לבדוק האם יש קורלציה בין המוסיקאים מבחינת הטעם. בטבלה הבאה נתונים הציונים של כל אחד מהמוסיקאים את שמונת המתמודדים:

	1	2	3	4	5	6	7	8
מוסיקאי א'	6	5	7	4	3	1	1	4
מוסיקאי ב'	5	7	5	2	3	3	2	7

מהו מדד הקשר המתאים?

א. פירסון.

ב. ספירמן.

ג. קרמר.

ד. אין מספיק נתונים כדי לדעת.

13) להלן טבלה המסכמת את השכר באלפי ₪ של עובדים בחברה ואת רמת המוטיבציה שלהם מ-1 עד 5:

שכר	12	18	20	15	30
מוטיבציה	4	4	5	3	5

מהו מקדם המתאם המתאים לבדיקת רמת ההתאמה בין המוטיבציה לשכר של העובד?

א. פירסון.

ב. ספירמן.

ג. קרמר.

ד. אין מספיק נתונים כדי לדעת.

© כל הזכויות שמורות לברק קנדל

14) להלן טבלה על נתונים שנאספו על מספר תצפיות:

X	1	2	3	4	5
Y	12	14	17	17	20

אם מעוניינים לבדוק עד כמה קיים קשר לנארי בין שני המשתנים.
מהו המדד המתאים?

א. פירסון.

ב. ספירמן.

ג. קרמר.

ד. אין מספיק נתונים כדי לדעת.

תשובות סופיות:

1) א'	**2)** ד'	**3)** ב'	**4)** ג'	**5)** א'
6) ב'	**7)** א'	**8)** ב'	**9)** ג'	**10)** ג'
11) ג'	**12)** ב'	**13)** ב'	**14)** א'	

© כל הזכויות שמורות לברק קנדל