מדדי פיזור

בפרק זה נדון במדדי פיזור.
מדדים אלה באים לשקף עד כמה ערכי המשתנים קרובים או רחוקים אחד מהשני.
כדי להבין את הצורך במדדים אלה, נתבונן בדוגמה שבה מוצגים ציוני התעודה של שני ילדים: עודד ודביר.

עודד: 80 , 80 , 70 , 70 , 70 , 60 , 60

דביר: 100 , 100 , 70 , 70 , 70 , 40 , 40

הממוצע של שניהם הוא 70, החציון של שניהם הוא 70, והשכיח של שניהם הוא 70.
למרות זאת, ברור שהתפלגות הציונים של עודד **שונה** מזו של דביר. השוויון במדדי המרכז מראה שהם לא משקפים היטב את ההבדלים בין ציוני התלמידים.

אפשר לתאר את ההבדלים בין הציונים על פי מידת הפיזור של הציונים סביב הממוצע.
בתעודה של עודד, מלבד שלושה ציונים של 70, יש שני ציונים של 60 ושני ציונים של 80.
בתעודה של דביר, מלבד שלושה ציונים של 70, יש שני ציונים של 40 ושני ציונים של 100.
ניתן לראות, שמלבד שלושת הציונים השווים לממוצע אצל שניהם, הציונים של עודד **קרובים** יותר לממוצע, והציונים של דביר **רחוקים** יותר מהממוצע ולכן מפוזרים יותר.
בחלק מהמקרים, כמו בדוגמה זו, אפשר לדעת באיזו קבוצת ציונים הנתונים מפוזרים יותר, גם מבלי לבצע חישובים. אבל, לעיתים זה מורכב יותר, ולכן נצטרך לבצע חישובים.
נכיר שני מדדי פיזור: האחד נקרא **טווח**, והשני נקרא **סטיית תקן.**

טווח

אחת האפשרויות לאפיין פיזור של קבוצת נתונים היא לחשב את ההפרש בין הערך הגדול ביותר בקבוצה לערך הקטן ביותר בקבוצה. מדד פיזור זה נקרא טווח. נגדיר אותו.

טווח – ההפרש בין המשתנה בעל הערך הגבוה ביותר למשתנה בעל הערך הנמוך ביותר.

בדוגמה הנ"ל, הציון הגבוה ביותר של עודד הוא 80, והציון הנמוך ביותר שלו הוא 60, לכן טווח הציונים שלו הוא $80-60$, כלומר 20.
לעומת זאת, הציון הגבוה ביותר של דביר הוא 100, והציון הנמוך ביותר שלו הוא 40, לכן טווח הציונים שלו הוא $100-40$, כלומר 60.
בדוגמה זו ראינו מראש, שהציונים של דביר מפוזרים יותר מאשר הציונים של עודד.
כאשר חישבנו את הטווח ראינו שגם על פיו פיזור הציונים של דביר גדול יותר.
במילים אחרות, **בדוגמה זו הטווח משקף את הבדל הפיזור בין שתי הקבוצות.**

נוסיף לדוגמה הנ"ל את ציוני התעודה של בועז: 100, 90 , 70 , 70 , 70 , 50 , 40

הציונים של בועז אינם זהים לציונים של עודד ודביר, אבל הממוצע, החציון והשכיח של הציונים שלו זהים לאלה של עודד ודביר. פיזור הציונים של בועז שונה מהפיזור של הציונים של עודד ודביר. טווח הציונים של בועז הוא $100-40$, כלומר 60.
הטווח של בועז שונה מהטווח של עודד, כלומר מדד הטווח אכן משקף את ההבדלים בציונים שלהם, אבל טווח הציונים של דביר ובועז שווה, למרות שפיזור הציונים שלהם שונה, כלומר מדד הטווח אינו משקף בצורה טובה את ההבדלים בין הציונים שלהם.
נלמד מכך ש**מדד הטווח לא תמיד משקף** בצורה טובה הבדלים בין התפלגויות שונות.

יתרונות הטווח:
פשוט וקל לחישוב, **מתאים למושג האינטואיטיבי שיש לנו על פיזור.**

חסרונות/מגבלות הטווח:
אינו מספיק מהימן, מכיוון שהוא מבוסס רק על שני הערכים הקיצוניים.
ניתן לחשבו רק עבור משתנה כמותי (בדיד או רציף)

תרגילים

1. שישה תלמידים נבחנו במבחן בהיסטוריה. התקבלו הציונים הבאים: 9, 8, 8, 7, 7, 6.
א. חשבו את טווח הציונים.
ב. הוסיפו ציון של תלמיד שביעי שנבחן בהיסטוריה. הטווח של שבעת התלמידים גדל ל- 4.
מהו הציון של התלמיד שהתווסף? כתבו את שתי האפשרויות.

2. להלן רשימת הציונים של חמישה תלמידים בשני מקצועות: ספרות ותנ״ך.
הציונים בספרות: 85, 80, 80, 80, 75.
הציונים בתנ״ך: 90, 90, 80, 70, 70.
א. הראו שהממוצע בכל אחד משני המקצועות הוא 80.
ב. העריכו: באיזה מקצוע פיזור הציונים גדול יותר?
ג. מהו טווח הציונים בספרות ומהו טווח הציונים בתנ״ך?
ד. **האם הטווח משקף** בצורה טובה את ההבדלים בפיזור הציונים בשני המקצועות?

3. נתון המשקל (בק״ג) של ארבעה נערים: 75, 70, 60, 55
א. חשבו את ממוצע המשקל של ארבעת הנערים.
ב. מהו טווח המשקלים של הנערים?
ג. לאחר חודשיים, נבדקו שוב המשקלים, והתברר כי המשקל של שני נבדקים השתנה.
משקלו של אחד הנערים עלה מ- 55 ק״ג ל- 60 ק״ג, ומשקלו של נער אחר ירד מ- 75 ק״ג ל- 70 ק״ג. המשקלים (בק״ג) בבדיקה השנייה: 70, 70, 60, 60
ענו ללא חישוב: האם לאחר הבדיקה הנוספת, פיזור המשקלים גדול, קטן או שווה לפיזור המשקלים לפני הבדיקה הנוספת?
ד. (1) האם בעקבות השינוי הטווח השתנה?
(2) האם מגמת השינוי בטווח מתארת את מגמת השינוי בפיזור המשקלים? נמקו.
ה. בודקים את המשקל של נער חמישי. **האם ייתכן** שטווח המשקל של חמשת הנערים (ארבעת הנערים מתחילת השאלה והנער שהתווסף) הוא 5 ק״ג?

4. בטבלה שמשמאל מתוארת התפלגות של ציונים, המקבלים ערכים שלמים בין 0 ל- 100.
נתון כי טווח הציונים הוא 70,
וחציון הציונים הוא 50. מצאו את x ואת y.

הציון	20	40	60	x
מספר התלמידים	7	y	15	5

5. קבעו עבור כל טענה האם היא נכונה:
א. בקבוצת נתונים, כאשר כל הנתונים שווים זה לזה, אז הטווח הוא אפס.
ב. בקבוצת נתונים, כאשר הטווח הוא אפס, אז כל הנתונים שווים זה לזה.

תשובות:

1. א. 3. ב. 5 או 10.
2. ב. בתנ״ך. ג. בתנ״ך - 20, בספרות - 10.
ד. כן. הפיזור בתנ״ך גדול יותר וגם טווח הציונים בתנ״ך גדול יותר.
3. א. 65 ק״ג. ב. 20 ק״ג. ג. פיזור המשקלים קטן.
ד. (1) כן. (2) כן. הפיזור קטן, וגם טווח המשקלים ירד מ- 20 ק״ג ל- 10 ק״ג. ה. לא.
4. $x = 90$, $y = 13$. **5.** א. הטענה נכונה. ב. הטענה נכונה.

סטיית תקן

נכיר עכשיו את מדד הפיזור העיקרי הנקרא **סטיית תקן**.
מדד זה מייצג את **מידת הפיזור** של הנתונים **סביב הממוצע.**
בעזרתו נלמד באיזו מידה הנתונים קרובים לממוצע (כלומר ההתפלגות יותר הומוגנית) או רחוקים מהממוצע (כלומר ההתפלגות יותר הטרוגנית).
נתחיל בדוגמה.

חמישה תלמידים נבחנו בשני מקצועות: פיסיקה ותנ"ך. התקבלו הציונים הבאים:

תנ"ך	**פיסיקה**
6 , 6 , 7 , 8 , 8	4 , 4 , 7 , 10 , 10

הממוצע בכל אחד משני המקצועות הוא 7.
ניתן לראות שקיים הבדל בין המקצועות במידת הפיזור של הציונים סביב הממוצע.
הציונים בתנ"ך קרובים יותר לממוצע 7, מאשר הציונים בפיסיקה, כלומר **מידת פיזור הציונים סביב הממוצע בפיסיקה גדולה יותר** מאשר במתמטיקה.

סטיית תקן של נתונים מייצגת את מידת הפיזור שלהם סביב הממוצע, ולכן כאשר בשני המקצועות יש אותו ממוצע, ונחשב את סטיית התקן של בשני המקצועות, נקבל ש**סטיית התקן של הציונים בפיסיקה גדולה יותר** מסטיית התקן של הציונים בתנ"ך.

נסביר עכשיו באמצעות דוגמה את תהליך החישוב של סטיית התקן של קבוצת נתונים.
אחר כך נסכם את התהליך באמצעות נוסחה לחישוב סטיית התקן.

דוגמה:
נתונה קבוצת ציונים: 10 , 10 , 8 , 7 , 5. חשבו את סטיית התקן של הציונים.
פתרון:
שלב א' – נחשב את הציון הממוצע: $\bar{x}=\frac{5+7+8+10+10}{5}=\frac{40}{5}=8$

שלב ב' – נחשב את ההפרש (הסטייה) של כל ציון מהממוצע.

הסטייה של הציון 5 מהממוצע 8 היא $5-8=-3$.
הסטייה של הציון 7 מהממוצע 8 היא $7-8=-1$, של הציון 8 היא $8-8=0$.
של הציון 8 היא $8-8=0$, של הציון 10 היא $10-8=+2$, של הציון 10 היא $10-8=+2$.

הסטיות מהממוצע הן: $2\ ,\ 2\ ,\ 0\ ,\ -1\ ,\ -3$

אנו מעוניינים במדד לפיזור הנתונים סביב הממוצע, ולכן חישוב הסטיות (ההפרשים) מהממוצע הוא הגיוני. עם זאת, **לא ניתן** להשתמש בסכום הסטיות מהממוצע ככלי לבדיקת פיזור הנתונים, מאחר ו**סכום** זה שווה תמיד לאפס (למדנו על כך בעמוד 129).
הסטיות של הנתונים מהממוצע יכולות להיות חיוביות, שליליות או אפס.
הן מבטלות זו את זו וסכומן תמיד אפס.
כדי שהסטיות לא תבטלנה זו את זו, נוכל לפעול בשתי דרכים.

דרך אחת היא למצוא את **הערכים המוחלטים** של הסטיות.
דרך שנייה היא **העלאה בריבוע** של כל סטייה.
<u>הדרך השנייה היא הדרך הנהוגה</u> בחישוב סטיית תקן, וכך גם נפעל.

שלב ג' – נחשב את ריבוע ההפרש (הסטייה) של כל ציון מהממוצע.

נרשום שוב את הסטיות מהממוצע: $2\ ,\ 2\ ,\ 0\ ,\ -1\ ,\ -3$

נעלה <u>בריבוע</u> את הסטיות מהממוצע: $2^2\ ,\ 2^2\ ,\ 0^2\ ,\ (-1)^2\ ,\ (-3)^2$

נחשב את <u>ריבועי</u> הסטיות מהממוצע: $4\ ,\ 4\ ,\ 0\ ,\ 1\ ,\ 9$

שלב ד' – נחשב את <u>סכום</u> ריבועי הסטיות מהממוצע.

נחבר את ריבועי הסטיות שחישבנו בשלב ג'.

נקבל שסכום ריבועי הסטיות הוא $18=4+4+0+1+9=2^2+2^2+0^2+(-1)^2+(-3)^2$

<u>הערה:</u> הציון 10 מופיע פעמיים. מכיוון שאנחנו מעוניינים בסכום ריבועי הסטיות, לא חייבים לחשב פעמיים את ריבוע הסטייה שלו מהממוצע.
פשוט יותר לחשב את ריבוע הסטייה שלו מהממוצע, כלומר 2^2, ולכפול ב- 2 .
נציג בצורה זו את חישוב סכום ריבועי הסטיות: $(-3)^2+(-1)^2+0^2+2\cdot 2^2=18$

שלב ה' – נחשב את <u>ממוצע</u> ריבועי הסטיות מהממוצע.

נחלק את סכום ריבועי הסטיות שחישבנו בשלב ג', השווה ל- 18, במספר הנתונים, השווה ל- 5 . נקבל: $\frac{18}{5}=3.6$.

החלוקה במספר הנתונים היא חשובה. אם נחשב רק את סכום ריבועי הסטיות, תהיה לו נטייה לגדול ככל שיש יותר משתנים, גם אם הפיזור לא גדל.

ממוצע ריבועי הסטיות שחישבנו <u>נקרא שונות</u>. בדוגמה זו השונות שווה ל- 3.6 .

שלב ו' – נחשב את <u>סטיית התקן</u>, שהיא השורש הריבועי של השונות.

בדוגמה זו השונות היא 3.6 , ולכן סטיית התקן היא $\sqrt{3.6}$, כלומר 1.897 .

הערה: השונות וסטיית התקן מודדים מידת פיזור. יחידות המדידה של השונות הן **ריבוע** יחידות המדידה של המשתנה. לדוגמה: אם המשתנה נמדד במטרים, אז השונות נמדדת במ"ר. כדי לחזור ליחידות המדידה של המשתנה, מוציאים שורש מהשונות, ומקבלים את סטיית התקן, המודדת את הפיזור ביחידות המדידה של המשתנה.

נסכם את תהליך חישוב סטיית התקן. נהוג לסמן את סטיית התקן על ידי S .

$$\text{S} \ \textbf{סטיית התקן} = \sqrt{\frac{\textbf{סכום ריבועי הסטיות מהממוצע}}{\textbf{מספר הנתונים}}}$$

הערה: הביטוי הנמצא בתוך השורש הוא השונות.
כפי שראינו, סטיית התקן היא שורש של השונות. **נהוג לסמן את השונות** על ידי S^2 .

בדוגמה הנ"ל קבוצת הציונים היא: 10 , 10 , 8 , 7 , 5

נתבונן עכשיו בקבוצת ציונים חדשה: 9 , 9 , 8 , 7 , 7

בדומה לקבוצת הציונים הנ"ל, גם בקבוצת הציונים החדשה ממוצע הציונים הוא 8 .
נחשב את סטיית התקן בקבוצה החדשה על פי השלבים שתיארנו:

נרשום את הסטיות מהממוצע: 1 , 1 , 0 , -1 , -1

נחשב את סכום ריבועי הסטיות מהממוצע: $(-1)^2+(-1)^2+0^2+1^2+1^2=4$

נחלק את סכום ריבועי הסטיות במספר הנתונים (שהוא 5), ונקבל את הממוצע של סכום ריבועי הסטיות. לממוצע זה קראנו שונות והוא שווה ל- $\frac{4}{5}=0.8$, כלומר $S^2=\frac{4}{5}=0.8$.

לחישוב סטיית התקן נוציא שורש מהשונות. נקבל: $S=\sqrt{\frac{4}{5}}=\sqrt{0.8}=0.894$

ניתן לראות שסטיית התקן של הציונים הללו, השווה ל- 0.894 , **קטנה** יותר מאשר סטיית התקן בדוגמה הקודמת, שם היא שווה ל- 1.897 . ערך גדול יותר של סטיית תקן מצביע על פיזור גדול יותר של המשתנה ביחס לממוצע, וערך קטן יותר של סטיית תקן מראה כי הערכים של המשתנה מקובצים יותר סביב הערך הממוצע.
כלומר **פיזור** הציונים כאן **קטן** לעומת קבוצת הציונים הקודמת.

באופן כללי, כדי לבדוק פיזור של משתנים, נעדיף להשתמש בסטיית התקן ולא בשונות, כך שהפיזור נמדד ביחידות המדידה של המשתנה. לעומת זאת, **השוואה של שתי קבוצות משתנים (או יותר)**, אפשר לעשות באמצעות סטיית תקן, או באמצעות שונות.

תכונות סטיית התקן:

(1) ככל שסטיית התקן גדולה יותר, הנתונים מפוזרים יותר ורחוקים יותר מהממוצע. לכן, אם נשווה שתי קבוצות נתונים בעלי **אותו ממוצע**, אז הקבוצה שבה הנתונים מפוזרים יותר, תהיה בעלת סטיית תקן גדולה יותר, ולהיפך.

(2) סטיית התקן היא גודל אי שלילי, וכך גם השונות.

(3) אם **כל הנתונים שווים** זה לזה בערכם, אז סטיית התקן היא אפס, ולהיפך, כלומר אם סטיית התקן היא אפס, אז כל הנתונים שווים זה לזה.

יתרונות סטיית התקן:

(1) מבוססת על כל הנתונים.

(2) היא באותן יחידות של המשתנה (בניגוד לשונות שאינה באותן יחידות של המשתנה).

חסרונות/מגבלות סטיית התקן:

(1) המשתנים בעלי הערך הקיצוני "תורמים" לערך של סטיית התקן תרומה גדולה יותר מאשר משקלם האמיתי.

(2) ניתן לחשבה רק עבור משתנה כמותי (בדיד או רציף), אך לא ניתנת לשימוש עבור משתנה איכותי או כאשר אחד מערכי הקיצון (או שניהם) הוא קבוצה פתוחה.

הערה: כאשר ידועים ערכי כל המשתנים, אפשר לחשב את סטיית התקן בעזרת מחשבון מתאים, או בעזרת תוכנה דוגמת אקסל (אין לעשות זאת בזמן בחינה).

תרגילים

1. שתי קבוצות משחקות כדורסל זו מול זו. לפניכם הגבהים (בס"מ) של שחקני כל קבוצה:

קבוצה א': 191, 190, 190, 190, 189

קבוצה ב': 200, 195, 190, 185, 180

א. הראו שממוצע הגובה בשתי הקבוצות זהה. מהו הגובה הממוצע?

ב. באיזו קבוצה מידת פיזור הגבהים סביב הממוצע גדולה יותר? הסבירו ללא נוסחאות.

ג. באיזו קבוצה סטיית התקן גדולה יותר? הסבירו ללא נוסחאות.

2. נתונות שלוש קבוצות של מספרים.

קבוצה א': 0, 0, 16, 16

קבוצה ב': 0, 8, 8, 16

קבוצה ג': 7, 8, 8, 9

א. הראו שממוצע המספרים בשלוש הקבוצות שווה ל- 8.

ב. קבעו ללא שימוש בנוסחאות, באיזו קבוצה מידת פיזור המספרים סביב הממוצע:

(1) היא הגדולה ביותר. (2) היא הקטנה ביותר.

ג. באיזו קבוצה סטיית התקן היא הקטנה ביותר? הסבירו ללא נוסחאות.

3. בשני בתי ספר תיכון נערך מבחן משווה. בתיכון "פסגות" התקבל ממוצע 80 וסטיית תקן 10. בתיכון "הנחל" התקבל ממוצע 80 וסטיית תקן 15.

א. באיזה תיכון פיזור הציונים גדול יותר? הסבירו.

ב. באיזה תיכון רמת הציונים אחידה יותר, ובאיזה תיכון אחידה פחות? נמקו.

4. שישה תלמידים נבחנו במבחן במדעים. התקבלו הציונים הבאים: 10, 7, 6, 5
א. חשבו את הציון הממוצע.
ב. חשבו את ההפרש (הסטייה) בין הציון של כל אחד מהתלמידים לבין הציון הממוצע.
סדרו בשורה את הסטיות מהממוצע.
ג. חשבו את ריבועי כל אחת מהסטיות שמצאתם בסעיף ב׳, וסדרו אותם בשורה.
ד. חשבו את הסכום של ריבועי הסטיות שמצאתם בסעיף ג׳.
ה. חשבו את השונות של הציונים (זהו הממוצע של סכום ריבועי הסטיות מהממוצע).
ו. חשבו את סטיית התקן של הציונים.

5. מנהלת הליגה בדקה כמה צופים הגיעו לכל אחד משמונה משחקי הכדורעף שנערכו ביום מסוים. התוצאות רשומות לפניכם: 140, 140, 100, 90, 50, 20
א. חשבו את הממוצע של מספר הצופים.
ב. חשבו את ריבועי הסטיות של כל נתון מהממוצע.
ג. חשבו את סכום ריבועי הסטיות של הנתונים מהממוצע.
ד. חשבו את השונות של מספר הצופים.
ה. חשבו את סטיית התקן של מספר הצופים.

6. חמישה תלמידים נבחנו כל אחד בשני מקצועות: ספרות ומתמטיקה.
הציונים בספרות הם: 6, 6, 7, 8, 8.
הציונים במתמטיקה הם: 4, 5, 7, 9, 10.
א. מהו הציון הממוצע בכל אחד משני המקצועות?
ב. <u>העריכו ללא חישוב</u>: באיזה משני המקצועות <u>פיזור</u> הציונים גדול יותר?
ג. חשבו את סטיית התקן של הציונים בכל מקצוע, ובדקו את תשובתכם לסעיף ב׳.

7. נבדקו הציונים של חמישה תלמידים בשני מקצועות: ביולוגיה וכימיה.
להלן הציונים בביולוגיה: 89, 89, 90, 91, 91, 91.
להלן הציונים בכימיה: 80, 80, 80, 90, 100, 100, 100.
א. מהו הציון הממוצע בכל אחד משני המקצועות?
ב. באיזה משני המקצועות <u>פיזור</u> הציונים גדול יותר להערכתכם? ענו ללא חישובים.
ג. חשבו את סטיית התקן של הציונים בכל מקצוע, ובדקו את תשובתכם לסעיף ג׳.
ד. האם טווח הציונים משקף בצורה טובה את ההבדל בפיזור הנתונים בשני המקצועות?

תשובות:

1. א. 190 ס״מ. ב. בקבוצה ב׳. ג. בקבוצה ב׳, מכיוון שמידת פיזור הנתונים בה גדולה יותר.
2. ב. (1) קבוצה א׳. (2) קבוצה ג׳. ג. קבוצה ג׳, כיוון שמידת פיזור הנתונים בה הקטנה ביותר.
3. א. פיזור הציונים גדול יותר בתיכון ״הנחל״.
ב. בתיכון ״פסגות״ סטיית התקן קטנה יותר, ולכן הציונים קרובים יותר לציון הממוצע, פחות מפוזרים, אחידים (הומוגניים) יותר. בתיכון ״הנחל״ הציונים הטרוגניים יותר.
4. א. 7. ב. 3, 0, -1, -2. ג. 9, 0, 1, 4. ד. 14. ה. 3.5. ב. 1.871.
5. א. 90 צופים. ב. 2,500, 2,500, 100, 0, 1,600, 4,900. ג. 11,600. ד. 1,933.33. ה. 43.97 צופים.
6. א. ספרות: 7, מתמטיקה: 7. ב. מתמטיקה. ג. $S_{\text{ספרות}} = 0.8944$, $S_{\text{מתמטיקה}} = 2.28$.
סטיית התקן במתמטיקה אכן גדולה יותר.
7. א. ביולוגיה: 90, כימיה: 90. ב. בכימיה.
ג. $S_{\text{ביולוגיה}} = 0.9258$, $S_{\text{כימיה}} = 9.258$. סטיית התקן בכימיה אכן גדולה יותר. ד. כן.

חישוב סטיית תקן מתוך טבלת שכיחויות

נראה עכשיו כיצד מחשבים סטיית תקן כאשר הנתונים מסודרים בטבלת שכיחויות. תהליך החישוב כבר הוצג בדוגמאות קודמות, אבל מכיוון שבטבלת שכיחויות ישנם בדרך כלל הרבה נתונים, נציג **נוסחה** בעזרתה נקצר ונפשט את החישוב.

הנוסחה לסטיית תקן היא:

$$S=\sqrt{\frac{(x_1-\bar{x})^2\cdot f_1+(x_2-\bar{x})^2\cdot f_2+(x_3-\bar{x})^2\cdot f_3+\ \ldots\ +(x_n-\bar{x})^2\cdot f_n}{N}}$$

נסביר את הסימונים:

הממוצע – $\bar{x}$, סטיית התקן – S

ערכי המשתנה x **השונים** שהתקבלו במדידה – $x_1,\ x_2,\ x_3\ldots$

השכיחויות המתאימות לערכי ה- x השונים – $f_1,\ f_2,\ f_3\ldots$

סכום השכיחויות – N . מתקיים: $N=f_1+f_2+f_3+\ldots+f_n$.

שימו לב!

(1) הסטיות של המשתנים מהממוצע הן: $x_1-\bar{x},\ x_2-\bar{x}\ ,\ \ldots\ ,\ x_n-\bar{x}$

ריבועי הסטיות של המשתנים מהממוצע הן: $(x_1-\bar{x})^2,\ (x_2-\bar{x})^2\ ,\ \ldots\ ,\ (x_n-\bar{x})^2$

סכום **ריבועי הסטיות מהממוצע** נמצא בנוסחה במונה השבר שבתוך השורש ומיוצג על ידי הביטוי הבא:

$$(x_1-\bar{x})^2\cdot f_1+(x_2-\bar{x})^2\cdot f_2+(x_3-\bar{x})^2\cdot f_3\ldots+(x_n-\bar{x})^2\cdot f_n$$

כאשר מחלקים את סכום ריבועי הסטיות במספר המשתנים N , מקבלים את השונות. נוציא שורש מהשונות ונקבל את סטיית התקן.

(2) אם נעלה בריבוע את שני אגפי הנוסחה נקבל את השונות, המסומנת S^2 .

$$S^2=\frac{(x_1-\bar{x})^2\cdot f_1+(x_2-\bar{x})^2\cdot f_2+(x_3-\bar{x})^2\cdot f_3+\ \ldots\ +(x_n-\bar{x})^2\cdot f_n}{N}$$

נפתור דוגמה שבה מחשבים סטיית תקן כאשר הנתונים מוצגים בטבלת שכיחויות.

דוגמה:

בטבלה שלפניכם נתונה התפלגות של ציונים. חשבו את סטיית התקן של הציונים.

x – ציון	50	60	70	80	90
f – מספר התלמידים	3	8	10	8	3

פתרון:

בשלב הראשון נחשב את הציון הממוצע: $\bar{x}=\frac{50\cdot3+60\cdot8+70\cdot10+80\cdot8+90\cdot3}{3+8+10+8+3}=\frac{2240}{32}=70$

נסמן בטבלה את המשתנים x_1 , x_2 , x_3 ... ואת השכיחויות f_1 , f_2 , f_3 ...

	x_1	x_2	x_3	x_4	x_5
	↑	↑	↑	↑	↑
x – ציון	50	60	70	80	90
f – מספר התלמידים	3	8	10	8	3
	↓	↓	↓	↓	↓
	f_1	f_2	f_3	f_4	f_5

בשלב השני נחשב את סטיית התקן באמצעות הנוסחה:

$$S=\sqrt{\frac{(x_1-\overline{x})^2\cdot f_1+(x_2-\overline{x})^2\cdot f_2+(x_3-\overline{x})^2\cdot f_3+...+(x_n-\overline{x})^2\cdot f_n}{N}}$$

נציב בנוסחה את הנתונים המוצגים בטבלה:

$$S=\sqrt{\frac{(50-70)^2\cdot 3+(60-70)^2\cdot 8+(70-70)^2\cdot 10+(80-70)^2\cdot 8+(90-70)^2\cdot 3}{3+8+10+8+3}}$$

$$S=\sqrt{\frac{400\cdot 3+100\cdot 8+0\cdot 10+100\cdot 8+400\cdot 3}{32}}=\sqrt{\frac{4000}{32}}=11.18$$

לסיכום: סטיית התקן של הציונים היא 11.18. משמעות המספר 11.18 היא שהסטייה הממוצעת של הציונים מהממוצע קרובה ל- 11.18 (היא לא שווה לה במדויק).

להעשרה הכוללת חישוב ממוצע וסטיית תקן באמצעות מחשבון או תוכנת Excel, **סרקו את הקוד המצורף.**

תרגילים

.8 לפניכם התפלגות מספר דוחות התנועה שקיבלו 20 נהגי מוניות במשך שנה:

מספר הדוחות שקיבל כל נהג	0	1	2	3	4
מספר הנהגים	1	4	8	6	1

א. מצאו כמה דוחות בממוצע קיבל כל נהג.
ב. חשבו את סטיית התקן של מספר הדוחות.

.9 לפניכם התפלגות הציונים של קבוצת תלמידים שנבחנה במתמטיקה:

מספר התלמידים	5	7	10	18
הציון	70	80	90	100

א. חשבו את ממוצע הציונים.
ב. חשבו את סטיית התקן של הציונים.
ג. (1) מצאו את השכיח והחציון.
(2) מבין שלושת מדדי המרכז (ממוצע, חציון ושכיח), קבעו איזה מדד מייצג את התפלגות הציונים באופן פחות טוב.

.10 בעמוד של ספר נספרו מספר המילים בכל שורה. התוצאות רשומות בטבלה שלפניכם:

מספר המילים בשורה	8	9	10	11	12
מספר השורות	6	6	4	10	4

א. חשבו את המספר הממוצע של מילים בשורה.
ב. חשבו את סטיית התקן של מספר המילים בשורה.
ג. בכמה מהשורות שנבדקו מספר המילים גדול מהממוצע ביותר מסטיית תקן אחת?

.11 א. בקבוצה תלמידים שנבחנו באזרחות התקבלו ציון ממוצע 85 וסטיית תקן השווה לאפס. האם אפשר לדעת מהו הציון שקיבל כל תלמיד?
ב. בקבוצה תלמידים שנבחנו בספרות, כל התלמידים קיבלו ציון 93.
האם סטיית התקן של הציונים היא בהכרח אפס?

12. בשתי כיתות שבכל אחת מהן יש 20 תלמידים, נערך מבחן בהיסטוריה.
בטבלה שמשמאל מתוארת התפלגות הציונים של הכיתה הראשונה.

הציון	70	80	90
מספר התלמידים	2	16	2

בטבלה שמשמאל מתוארת התפלגות הציונים של הכיתה השנייה.

הציון	60	70	80	90	100
מספר התלמידים	5	3	4	3	5

א. חשבו את ממוצע הציונים בכל כיתה והראו שהוא שווה ל- 80.
ב. באיזו כיתה פיזור הנתונים גדול יותר ביחס לממוצע? ענו ללא חישובים.
ג. חשבו את סטיית התקן בכל כיתה, והשוו להערכתכם שבסעיף ב'.

13. משרד הבריאות בחר מדגם של סיגריות מסוגים שונים, ובדק את כמות הניקוטין (במיליגרם) בכל אחת מהן. התוצאות מוצגות בטבלה הבאה:

כמות הניקוטין (במיליגרם)	10	12	14	16
מספר סיגריות	15	20	x	10

הממוצע של כמות הניקוטין לסיגריה שנמצא במדגם הוא 12.4 מיליגרם.
א. מצאו בכמה סיגריות נמצאו 14 מיליגרם ניקוטין.
ב. חשבו את סטיית התקן של כמות הניקוטין בסיגריות שנבדקו.

להעשרה בנושא השפעת הניקוטין, **סרקו את הקוד המצורף.**

14. נתונה ההתפלגות של יבול ענבים בטונות ב- 40 חלקות שדה.

מספר החלקות	4	?	18	7	?
יבול בטונות	8	9	10	11	12

השכיחות היחסית של חלקות שבהן היבול לחלקה הוא 12 טונות היא 10%.
א. בכמה חלקות היה יבול הענבים 9 טונות?
ב. חשבו את סטיית התקן של יבול הענבים.
ג. חשבו את העשירון העליון של יבול הענבים (בטונות).

15. במבחן שנערך התקבלו הציונים הבאים:

מספר התלמידים	10	6	4	3
הציון	50 – 59	60 – 69	70 – 79	80 – 89

א. חשבו את ממוצע הציונים.
ב. חשבו את סטיית התקן של הציונים.

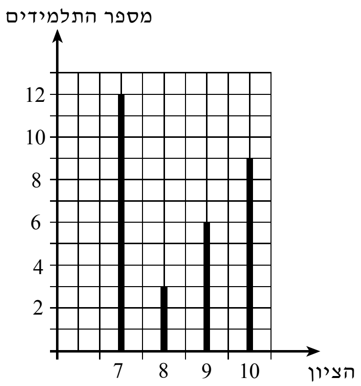
תשובות:

8. א. 2.1 דוחות. ב. 0.943 דוחות. **9.** א. 90.25. ב. 10.60. ג. (1) שכיח - 100, חציון - 90.
(2) השכיח. **10.** א. 10 מילים בשורה. ב. 1.366 מילים בשורה. ג. 4 שורות.
11. א. כן, כל אחד מהם קיבל 85. ב. כן.
12. ב. בכיתה השנייה.
ג. בכיתה הראשונה: 4.472. בכיתה השנייה: 15.17. תואם להערכה שבסעיף ב'.
13. א. 5 סיגריות. ב. 2.154 מיליגרם ניקוטין.
14. א. 7 חלקות. ב. 1.072 טונות. ג. 11.5 טונות. **15.** א. 64.5. ב. 10.63.

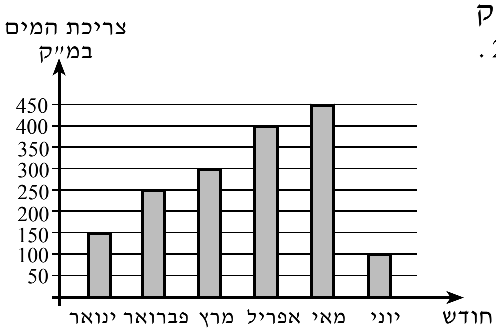
חישוב סטיית תקן מתוך ייצוגים גרפיים שונים

כדי לחשב סטיית תקן, כשהנתונים מוצגים באופן גרפי (דיאגרמת **עמודות**, דיאגרמת עיגול, היסטוגרמה או גרף של נקודות), עדיף להעביר את הנתונים לטבלת שכיחויות.

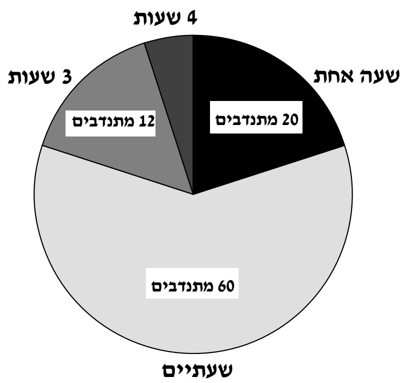
16. לפניכם דיאגרמה המתארת את התפלגות הציונים במבחן בסייבר במחלקה למדעים:
א. כמה סטודנטים נבחנו בסך הכול?
ב. סדרו את הנתונים בטבלת שכיחויות.
ג. חשבו את הממוצע, השכיח והחציון של הציונים.
ד. חשבו את סטיית התקן של הציונים.
ה. חשבו את השכיחות היחסית של הציונים, שגבוהים מהממוצע ביותר מסטיית תקן אחת.



17. הדיאגרמה שלפניכם מתארת את צריכת המים במ״ק של מפעל בששת החודשים הראשונים של שנת 2023.
א. מהי כמות המים במ״ק שצרך המפעל בחודש אפריל של שנת 2023?
ב. מה הייתה הצריכה החודשית הממוצעת של המפעל במחצית הראשונה של 2023?
ג. חשבו את סטיית התקן של הצריכה החודשית של המפעל במחצית <u>הראשונה</u> של שנת 2023.



18. דיאגרמת העיגול שלפניכם מציגה את מספר שעות ההתנדבות בשבוע של 96 תלמידים המתנדבים במוסדות ציבוריים:
א. כמה תלמידים מתנדבים במשך 4 שעות?
ב. חשבו את הממוצע של מספר שעות ההתנדבות במוסדות ציבוריים.
ג. מהו מספר שעות ההתנדבות השכיח? מה משמעותו?
ד. מהו החציון של מספר שעות ההתנדבות?
ה. חשבו את סטיית התקן של מספר שעות ההתנדבות.



תשובות:

16. א. 30 סטודנטים. ב.

x	הציון	7	8	9	10
f	מספר התלמידים	12	3	6	9

ג. הממוצע הוא 8.4, השכיח הוא 7, החציון הוא 8.5. ד. 1.281. ה. $30\% = \frac{3}{10} = \frac{9}{30}$.

17. א. 400 מ״ק. ב. 275 מ״ק. ג. 125 מ״ק.

18. א. 4 תלמידים. ב. שעתיים. ג. שעתיים. המשמעות: הכי הרבה תלמידים מתנדבים במשך שעתיים במוסדות הציבוריים. ד. שעתיים. ה. 0.707 שעות (42.42 דקות).

מציאת נעלמים כאשר נתונה סטיית התקן

דוגמה:
לפניכם התפלגות הציונים של קבוצת תלמידים במבחן בפיסיקה:

הציון	6	7	8	9	10
מספר התלמידים	3	6		6	3

א. הראו כי ממוצע הציונים במבחן היה 8.
ב. מצאו כמה תלמידים קיבלו את הציון 8, אם סטיית התקן של הציונים הייתה $\frac{3}{4}$.

פתרון:
א. **דרך א':** נסמן ב- x את מספר התלמידים שקיבלו 8, וניעזר בנוסחה לחישוב הממוצע. נקבל:
$$\bar{x}=\frac{6\cdot3+7\cdot6+8\cdot x+9\cdot6+10\cdot3}{3+6+x+6+3}=\frac{144+8x}{18+x}$$
במונה נוציא גורם משותף 8 ונצמצם ב- $(18+x)$:
$$\bar{x}=\frac{8(18+x)}{(18+x)}=8$$

דרך ב': נחשב את הממוצע ללא התלמידים שקיבלו 8:
$$\bar{x}=\frac{6\cdot3+7\cdot6+9\cdot6+10\cdot3}{3+6+6+3}=\frac{144}{18}=8$$

הוספת תלמידים שקיבלו בדיוק 8 (ולא חשוב מספרם) לשאר תלמידים (שהממוצע שלהם הוא 8) לא תשנה את הממוצע, לכן הציון הממוצע של כל התלמידים הוא 8.

ב. נסמן ב- x את מספר התלמידים שקיבלו 8 (שהוא גם הממוצע).
נביע את סטיית התקן באמצעות x, ונשווה אותו לסטיית התקן הנתונה.

נביע את סטיית התקן:
$$S=\sqrt{\frac{(6-8)^2\cdot3+(7-8)^2\cdot6+(8-8)^2\cdot x+(9-8)^2\cdot6+(10-8)^2\cdot3}{3+6+x+6+3}}=$$
$$S=\sqrt{\frac{4\cdot3+1\cdot6+0\cdot x+1\cdot6+4\cdot3}{18+x}}=\sqrt{\frac{36}{18+x}}$$

נתון כי סטיית התקן היא $\frac{3}{4}$ ולכן המשוואה היא:
$$\sqrt{\frac{36}{18+x}}=\frac{3}{4}$$

כדי לבטל את השורש הריבועי נעלה בריבוע את שני האגפים:
$$\left(\sqrt{\frac{36}{18+x}}\right)^2=\left(\frac{3}{4}\right)^2$$

באגף שמאל השורש מתבטל: $\frac{36}{18+x}=\frac{9}{16}$.

נכפול את שני האגפים במכנה המשותף. נקבל $36\cdot16=9\,(18+x)$, ומכאן $576=162+9x$.
פתרון המשוואה הוא $x=46$.
הערה: במשוואות כאלה אין צורך לבדוק את הפתרון, למרות שהעלינו בריבוע את שני האגפים, מאחר ושני האגפים הם חיוביים.

לסיכום: 46 תלמידים קיבלו את הציון 8.

19. פתרו את המשוואות הבאות:

א. $\sqrt{\frac{20}{x+2}}=\frac{1}{2}$ ב. $\sqrt{\frac{60+8x}{3+x}}=3$ ג. $\sqrt{\frac{2x+19}{3x+24}}=\frac{5}{6}$ ד. $\sqrt{\frac{20+5x}{14+x}}=2$

20. בדקו את מספר החדרים בדירות בבניין מסוים והתקבלה ההתפלגות הבאה:

מספר החדרים בדירה	2	3	4
מספר הדירות	2	x	2

א. הראו כי ממוצע מספר החדרים בדירה הוא 3. הסבירו באופן מילולי או על ידי חישובים.
ב. מצאו את הערך של x, אם ידוע כי סטיית התקן של מספר החדרים בדירה היא $\frac{2}{3}$.

21. להלן התפלגות הציונים של קבוצת תלמידים במבחן:

הציון	6	7	8	9	10
מספר התלמידים	4	9	x	9	4

א. הראו כי ממוצע הציונים במבחן הוא 8.
ב. כמה תלמידים קיבלו ציון 8, אם **סטיית התקן** במבחן זה היא $\frac{5}{6}$?

22. הציונים שהתקבלו במבחן במדעים היו 60, 70 ו- 80 בלבד.
10 תלמידים קיבלו ציון 70 והציון הממוצע של כל התלמידים היה 70.
א. הראו שמספר התלמידים שקיבלו ציון 60 שווה למספר התלמידים שקיבלו ציון 80.
ב. סטיית התקן של הציונים הייתה $5\sqrt{2}$. כמה תלמידים קיבלו ציון 60 וכמה קיבלו ציון 80?
ג. האם יש מדד מרכז (ממוצע, שכיח או חציון) המייצג את ההתפלגות באופן טוב יותר?

23. במבחן בתנ"ך הציונים שהתקבלו היו 6, 7 ו- 9 בלבד.
שישה תלמידים קיבלו ציון 6 והציון הממוצע היה 7.
א. מצאו כמה תלמידים קיבלו את הציון 9. תוכלו להסביר באופן מילולי או על ידי חישובים.
ב. **השונות** של הציונים הייתה 0.36. כמה תלמידים קיבלו ציון 7?

24. בשכונה מסוימת בדקו כמה ילדים יש בכל משפחה. התוצאות רשומות בטבלה שלפניכם:

מספר המשפחות	10	x	31	23	y	3
מספר הילדים בכל משפחה	0	1	2	3	4	5

א. ידוע כי מספר הילדים הממוצע למשפחה הוא 2 וסטיית התקן היא 1.2.
מצאו את הערכים של x ו- y. הדרכה: פתרו מערכת של שתי משוואות עם שני נעלמים.
ב. לאילו עשירונים יש 2 ילדים במשפחה?

25. במפעל המייצר ברגים בדקו כמה ברגים פגומים יש בכל קופסה.
התוצאות שהתקבלו רשומות בטבלה שלפניכם:

מספר הקופסאות	15		75	
מספר הברגים הפגומים	0	1	2	3

ידוע כי ממוצע מספר הברגים הפגומים בקופסה הוא 1.4 וסטיית התקן היא 0.7.
א. בכמה קופסאות יש בורג אחד פגום ובכמה יש 3 ברגים פגומים?
ב. חשבו את השכיחות היחסית (באחוזים) של הקופסאות שבהן מספר הגפרורים הפגומים גדול מהממוצע ביותר משתי סטיות תקן.

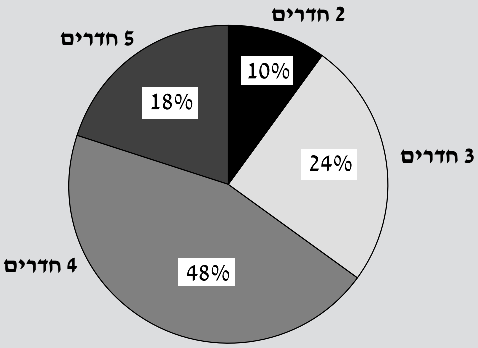
תשובות:
19. א. $x=78$. ב. $x=33$. ג. $x=28$. ד. $x=36$. **20.** ב. $x=5$. **21.** ב. 46 תלמידים.
22. ב. 5 תלמידים, 5 תלמידים. ג. לא, שלושת המדדים שווים בערכם.
23. א. 3 תלמידים. ב. 41 תלמידים. **24.** א. $x=26$, $y=7$. ב. עשירונים רביעי, חמישי ושישי.
25. א. 100 קופסאות, 10 קופסאות. ב. 5%.

סטיית תקן – שאלות עם פרמטרים ויחסים

נראה עכשיו חישובים של סטיית תקן הכוללים **פרמטרים**, וחישובים של סטיית תקן במקרים שבהם לא נתונות השכיחויות, אלא השכיחויות היחסיות (לרוב ב**אחוזים**).

דוגמה:

חברת הבנייה "גגות" בנתה פרויקט שבו היו דירות למגורים בנות 2, 3, 4 ו-5 חדרים. הדיאגרמה שלפניכם מתארת את התפלגות הדירות בפרויקט זה:



א. חשבו את מספר החדרים הממוצע בדירה בפרויקט.

ב. מהי סטיית התקן של מספר החדרים בדירה בפרויקט?

פתרון:

א. נעביר את הנתונים לטבלה של שכיחות יחסית, כאשר **התכונה הנבדקת** היא מספר החדרים ו**השכיחות היחסית** היא האחוז המתאים לכל אחד מסוגי הדירות.

נקבל:

מספר החדרים	2	3	4	5
השכיחות היחסית	10%	24%	48%	18%

דרך א' – נתונות לנו שכיחויות יחסיות, לכן נחשב למעשה ממוצע משוקלל.

נקבל: $\bar{x} = 2 \cdot \frac{10}{100} + 3 \cdot \frac{24}{100} + 4 \cdot \frac{48}{100} + 5 \cdot \frac{18}{100} = 3.74$

לסיכום: מספר החדרים הממוצע הוא 3.74 חדרים.

דרך ב' – נסמן ב- x את מספר הדירות בכל הפרויקט, ונבנה טבלת שכיחויות.
השכיחות של דירות בנות 2 חדרים היא 10% מתוך x, כלומר $\frac{10}{100}x = 0.1x$.
באופן דומה, נחשב ונקבל את השכיחויות האחרות הרשומות בטבלה.

נקבל:

מספר החדרים	2	3	4	5
השכיחות	0.1x	0.24x	0.48x	0.18x

נחשב את הממוצע: $\bar{x} = \frac{2 \cdot 0.1x + 3 \cdot 0.24x + 4 \cdot 0.48x + 5 \cdot 0.18x}{0.1x + 0.24x + 0.48x + 0.18x} = \frac{3.74x}{x} = 3.74$

ב. נסמן ב- x את מספר הדירות בכל הפרויקט, ונחשב את סטיית התקן על פי דרך ב':

$$S = \sqrt{\frac{(2-3.74)^2 \cdot 0.1x + (3-3.74)^2 \cdot 0.24x + (4-3.74)^2 \cdot 0.48x + (5-3.74)^2 \cdot 0.18x}{x}} =$$

$$S = \sqrt{\frac{0.3028x + 0.1314x + 0.0324x + 0.2858x}{x}} = \sqrt{\frac{0.7524x}{x}} = \sqrt{0.7524} = 0.8674$$

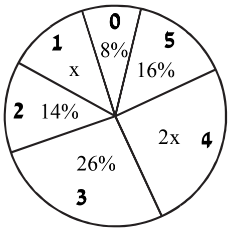
לסיכום: סטיית התקן של מספר החדרים היא 0.8674.

.26 במפעל יש 20 פועלים שכל אחד מהם משתכר k שקלים לחודש, ו- 30 פועלים שכל אחד מהם משתכר $k+1000$ שקלים לחודש.
א. הביעו באמצעות k את ממוצע השכר.
צמצמו את השבר המתקבל.
ב. חשבו את סטיית התקן של השכר החודשי (תוצאה מספרית).

.27 במבחן כיתתי חצי מהתלמידים קיבלו 70 וחצי מהתלמידים קיבלו 80 .
א. מהו ממוצע הציונים?
ב. חשבו את סטיית התקן של הציונים.

.28 במבחן שכבתי 80% מהתלמידים קיבלו ציון 80 ו- 20% מהתלמידים קיבלו ציון 60 .
א. מהו טווח הציונים?
ב. חשבו את ממוצע הציונים.
ג. חשבו את סטיית התקן של הציונים.

.29 בבית ספר תיכון הוחלט לבדוק את מספר החיסורים החודשי של כל תלמיד ותלמידה.
בדיאגרמת העיגול שלפניכם מתוארת התפלגות החיסורים באחוזים.
על כל גזרה רשום מספר האיחורים המתאים לה,
כאשר מספר האיחורים הנמוך ביותר הוא 0 ,
ומספר האיחורים הגבוה ביותר הוא 5 .
א. לאיזה אחוז מהתלמידים יש 5 איחורים?
ב. כתבו את השכיח והחציון של מספר האיחורים החודשי.
ג. סדרו את הנתונים בטבלה של שכיחויות יחסיות באחוזים.
ד. חשבו את ממוצע מספר האיחורים החודשי.
ה. חשבו את סטיית התקן של מספר האיחורים.
ו. מהי השכיחות היחסית (באחוזים) של האיחורים,
שקטנים מהממוצע ביותר מסטיית תקן אחת?



תשובות:

.26 א. $k+600$ שקלים לחודש. ב. 489.89 שקלים.

.27 א. 75 . ב. 5 .

.28 א. 20 . ב. 76 . ג. 8 .

.29 א. 24% . ב. השכיח הוא 3 איחורים, החציון הוא 3 איחורים.
ג.

מספר החיסורים	0	1	2	3	4	5
השכיחות היחסית	8%	12%	14%	26%	24%	16%

ד. 2.94 איחורים. ה. 1.489 איחורים. ו. 20% .

שינוי בערכי כל המשתנים והשפעתו על מדדי הפיזור

קיימים מקרים שבהם מבצעים פעולה של חיבור, חיסור, כפל או חילוק על **כל** הנתונים הנבדקים. למדנו בעבר כיצד שינויים אלה משפיעים על מדדי המרכז.
נלמד עכשיו כיצד הם משפיעים על מדדי הפיזור.

לדוגמה: נתונים שלושה מספרים: 24, 18, 15
אם כל אחד משלושת המספרים יגדל **ב-** 2, אז המספרים יהיו 26, 20, 17.
המרווחים ביניהם לא ישתנו, הפיזור לא ישתנה. סטיית התקן לא תשתנה, וכך גם הטווח.
לעומת זאת, אם כל מספר גדל **פי** 2, אז המספרים יהיו 48, 36, 30.
המרווחים ביניהם גדלו פי 2. סטיית התקן תגדל פי 2, וכך גם הטווח.

נסכם את הכללים העיקריים:
א. כאשר מגדילים (או מקטינים) את **כל** הנתונים במספר קבוע חיובי k,
למדנו שהממוצע, החציון והשכיח גדלים (או קטנים) גם הם ב- k.
במקרה כזה, הפיזור אינו משתנה, **ומדדי הפיזור** – טווח וסטיית תקן **אינם משתנים.**
ב. כאשר כופלים (או מחלקים) את **כל** הנתונים במספר קבוע חיובי k, למדנו שהממוצע, החציון והשכיח מוכפלים (או מחולקים) גם הם ב- k. במקרה כזה, הפיזור משתנה.
מדדי הפיזור – טווח וסטיית תקן **מוכפלים (או מחולקים) גם הם ב-** k.

דוגמה:
במבחן שכבתי סטיית התקן הייתה 8 נקודות. מאחר והמבחן היה קשה מהרגיל, הוחלט להעלות כל ציון ב- 20%, ולהוסיף לציון שמתקבל לאחר ההעלאה עוד 6 נקודות.
א. מה תהיה סטיית התקן החדשה לאחר שני השינויים?
ב. מתברר שלאחר שני השינויים, טווח הציונים הוא 30 נקודות. מהו הטווח המקורי?
פתרון:
א. סטיית התקן הייתה 8 נקודות. נבדוק את ההשפעה של כל אחד משני השינויים.
בשלב הראשון מעלים את הציון של כל תלמיד ב- 20%, לכן לאחר השינוי כל ציון גדל פי 1.2. במקרה כזה, סטיית התקן **עולה גם היא** ב- 20%, כלומר גדולה פי 1.2 מסטיית התקן המקורית. נקבל $1.2\cdot 8=9.6$, ולכן לאחר השינוי הראשון סטיית התקן גדלה ל- 9.6. בשלב השני מעלים את הציון של כל תלמיד ב- 6 נקודות.
בשינוי כזה, **הפיזור אינו משתנה**, לכן סטיית תקן **אינה משתנה ו**נשארת 9.6 נקודות.
ב. נסמן ב- x את הטווח המקורי. בשלב הראשון מעלים את הציון של כל תלמיד ב- 20%, לכן הטווח **עולה גם הוא** ב- 20%, ומהווה 120% מהטווח x, כלומר שווה ל- $1.2x$.
בשלב השני מעלים את הציון של כל תלמיד ב- 6 נקודות.
במקרה של שינוי כזה, הפיזור אינו משתנה, כלומר הטווח **אינו משתנה,** ונשאר $1.2x$.
נתון כי הטווח לאחר שני השינויים הוא 30 נקודות, ולכן המשוואה היא $1.2x=30$.
הפתרון הוא $x=25$, וזהו הטווח המקורי. לסיכום: הטווח המקורי הוא 25 נקודות.

תרגילים

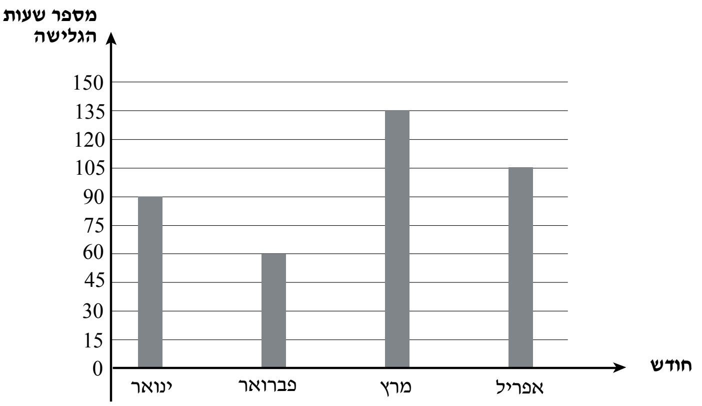
1. במבחן במדעים הציון הממוצע היה 76 וסטיית התקן הייתה 6.
הוחלט להעלות את הציונים של כל התלמידים ב- 4 נקודות.
א. (1) כיצד ישפיע שינוי זה על מדדי המרכז: ממוצע, חציון, שכיח?
(2) מה יהיה הממוצע החדש?
ב. האם השינוי ישפיע על מדדי הפיזור: טווח, סטיית תקן?
ג. בהמשך הוחלט לכפול את הציונים החדשים של כל התלמידים פי 1.1.
(1) מה יהיה הממוצע החדש?
(2) האם שינוי זה ישפיע על מדדי הפיזור: טווח, סטיית תקן?
(3) חשבו את סטיית התקן החדשה.

2. במפעל מסוים השכר הממוצע היה 8,000 שקלים וסטיית התקן הייתה 1,000 שקלים.
טווח השכר היה 3,000 שקלים. הוחלט לבצע שינוי בשכר העובדים.
א. בשלב הראשון הוחלט להעלות את השכר של כל העובדים ב- 5% .
(1) מה יהיה הממוצע של השכר לאחר השינוי?
(2) מה יהיה טווח השכר לאחר השינוי? מה תהיה סטיית התקן של השכר לאחר השינוי?
ב. בעקבות תלונות של העובדים, הוחלט לבצע העלאה נוספת בשכר,
וכל אחד מהעובדים קיבל העלאה נוספת של 200 שקלים.
(1) האם סטיית התקן של השכר תשתנה בעקבות ההעלאה הנוספת?
(2) חשבו מה יהיה טווח השכר לאחר שני השינויים.

3. בקבוצה המתאמנת לקראת תחרות ״איש הברזל״ נבדק משקלם של המתאמנים.
נמצא שממוצע המשקל שלהם הוא 85 ק״ג וסטיית התקן היא 10 ק״ג.
לקראת התחרות הוחלט שכל המתאמנים יורידו p% ממשקלם. מתברר כי לאחר <u>שכולם</u>
הורידו ממשקלם על פי התכנית, סטיית התקן של המשקל ירדה ל- 9 ק״ג.
א. חשבו את האחוז שבו ירדה סטיית התקן.
ב. חשבו את האחוז שבו ירד משקלו של כל משתתף.
ג. חשבו את ממוצע המשקל לאחר השינויים.
ד. הטווח של משקל התלמידים לאחר שני השינויים היה 21.6 ק״ג. מה היה הטווח המקורי?

4. מורה חישב ומצא שממוצע הציונים של תלמידיו הוא 50 וסטיית התקן היא 6 .
מכיוון שהממוצע היה נמוך מהרגיל, המורה העלה את הציון של כל אחד מתלמידיו באחוז
שווה, עליו החליט מראש. אחר כך הוסיף המורה לכל אחד מהציונים מספר שווה של נקודות.
לאחר שבוצעו שני השינויים, ממוצע הציונים עלה ל- 68 , וסטיית התקן עלתה ל- 7.2 .
א. בכמה אחוזים עלתה סטיית התקן בשלב הראשון?
ב. בכמה אחוזים העלה המורה כל ציון בשלב הראשון?
ג. כמה נקודות הוסיף המורה לכל ציון בשלב השני?
ד. האם לאחר שני השינויים שהמורה ביצע, הציון של כל התלמידים עלה בהכרח
<u>באותו אחוז</u> לעומת הציון המקורי שלהם?

5. לפניכם התפלגות **מספר שעות הגלישה באינטרנט** של אבי
בחודשים ינואר עד אפריל:



א. מה ממוצע שעות הגלישה
של אבי בחודשים אלה?
ב. מהי סטיית התקן של שעות
הגלישה של אבי בחודשים אלה?
ג. מסתבר שבכל אחד מארבעת
החודשים הנ״ל מספר שעות
הגלישה של שרון באינטרנט
גדול ממספר שעות הגלישה של אבי ב- 20% .
חשבו את הממוצע, החציון וסטיית התקן של שעות הגלישה של שרון.

תשובות:

1. א. (1) הם יגדלו ב- 4 נקודות. (2) 80 . ב. לא. ג. (1) 88 . (2) כן, הם יגדלו פי 1.1 . (3) 6.6 .
2. א. (1) 8,400 שקלים. (2) 3,150 שקלים, 1,050 שקלים. ב. (1) לא. (2) 3,150 שקלים.
3. א. 10% . ב. 10% . ג. 76.5 ק״ג. ד. 24 ק״ג. **4.** א. 20% . ב. 20% . ג. 8 נקודות. ד. אמיר לא צודק.
5. א. 97.5 שעות. ב. 27.04 שעות. ג. 117 שעות, 117 שעות, 32.45 שעות.

סטיית תקן – השפעה של החלפה, הוספה ומחיקת משתנים

בחלק זה נבחן את ההשפעה על סטיית התקן כאשר **מוסיפים** משתנים שהערך שלהם ידוע, או **מוחקים** משתנים שהערך שלהם ידוע, או **מחליפים** משתנים שהערך שלהם ידוע במשתנים אחרים שהערך שלהם אחר וידוע.
ננסה להבין האם **סטיית התקן** גדלה, קטנה או לא משתנה, ולעיתים אף לחשב אותה.

באופן כללי, ישנם מקרים שבהם **נוכל לחשב** את סטיית התקן החדשה, ולענות על השאלה.
לדוגמה: כאשר ידועים ערכי כל המשתנים ומוסיפים משתנים שערכם ידוע.
עם זאת, ישנם מקרים שבהם **לא** יהיו לנו מספיק נתונים כדי לחשב את סטיית התקן החדשה. במקרים כאלה נוכל לפעמים לתת תשובה עקרונית האם **סטיית התקן** גדלה, קטנה או לא השתנתה, גם מבלי שנוכל לחשב אותה. נעשה זאת על ידי כך שנעריך האם בעקבות השינויים, פיזור הנתונים סביב הממוצע גדל, קטן או לא השתנה.
הערכה כזאת נוכל לעשות בעיקר כאשר **הממוצע לא משתנה**, ובנוסף אפשר לדעת בבירור מה השפעת השינויים על פיזור הנתונים.
נראה דוגמה ולאחריה נסכם את הכללים.

דוגמה:
במבחן שכבתי חישבו ומצאו שהציון הממוצע הוא 75 .
א. הוחלט לשנות שני ציונים:
ציון אחד הוחלט להעלות מ- 75 ל- 80 וציון שני הוחלט להוריד מ- 75 ל- 70 .
(1) האם בעקבות השינויים, הממוצע גדל, קטן או נשאר ללא שינוי? נמקו.
(2) האם בעקבות השינויים, סטיית התקן גדלה, קטנה או נשארה ללא שינוי? נמקו.
ב. ענו על סעיף ב' ללא קשר לסעיף א': הוסיפו ציון נוסף, והתברר שהציון הממוצע לא השתנה. האם סטיית התקן גדלה, קטנה או לא השתנתה בעקבות צירוף הציון החדש?
פתרון:
א. (1) בוצע שינוי בשני ציונים. ציון אחד עלה ב- 5 נקודות והציון השני ירד ב- 5 נקודות, ולכן בסך הכול ממוצע הציונים נשאר ללא שינוי.
(2) שני הציונים שבהם נעשה שינוי היו שווים לממוצע 75 , והמרחק שלהם מהממוצע הוא אפס. הם הוחלפו בציונים הרחוקים מהממוצע ב- 5 נקודות, לכן השינוי מגדיל את פיזור הציונים סביב הממוצע, ומכאן שסטיית התקן גדלה.
ב. אם לאחר הוספת הציון החדש, הממוצע לא השתנה, אז הציון שהתווסף שווה לממוצע. כאשר מוסיפים ציון השווה לממוצע, פיזור הציונים סביב הממוצע קטן, ומכאן שסטיית התקן קטנה.

נסכם את הכללים העיקריים (הכללים נכונים בהנחה שהממוצע לא משתנה):

א. אם **מוסיפים** לאוכלוסייה משתנה שערכו שווה לממוצע, אז הממוצע אינו משתנה. במקרה כזה, הסטייה מהממוצע של הנתון שהתווסף שווה לאפס, כלומר הנתון הנוסף מחזק את ה"התקרבות" לממוצע, לכן פיזור הנתונים קטן ומכאן שסטיית התקן קטנה.
נסכם: כשמוסיפים נתון השווה לממוצע, סטיית התקן קטנה.

ב. כאשר **גורעים/מוחקים** משתנה שערכו שווה לממוצע, אז הממוצע לא משתנה. מחיקת נתון השווה לממוצע, מגדילה את פיזור הנתונים, ומכאן שסטיית התקן גדלה.

ג. אם מגדילים חלק מהמשתנים, ומקטינים אחרים, כך שהממוצע נשאר אותו דבר, נבדוק מה הסטייה מהממוצע של המשתנים החדשים לעומת המשתנים המוחלפים, וננסה להעריך האם הפיזור גדל או קטן, ובהתאמה האם סטיית התקן גדלה או קטנה. באופן דומה נפעל כש**מחליפים** מספר משתנים, אבל הממוצע נשמר (ראו סעיף א' בדוגמה הנ"ל).

תרגילים

1. בחוג ספורט נבדק גובהם של ארבעה משתתפים.
התקבלו הגבהים הבאים: 162 ס"מ, 165 ס"מ, 167 ס"מ ו- 170 ס"מ.
א. חשבו את הממוצע ואת סטיית התקן של הגובה.
ב. כעבור מספר שבועות נבדק שוב גובה ארבעת הספורטאים.
התברר שגובהם של שני הספורטאים הנמוכים יותר עלה ל- 166 ס"מ.
(1) האם הגובה הממוצע של ארבעת הספורטאים גדל, קטן או נשאר ללא שינוי.
נמקו ללא חישובים.
(2) חשבו את הממוצע החדש והשוו לסעיף א'.
(3) הסבירו מילולית מדוע סטיית התקן של הגובה קטנה בעקבות שינויי הגובה הנ"ל.
(4) חשבו את סטיית התקן החדשה של ארבעת הספורטאים.
הראו שהיא אכן קטנה יותר מסטיית התקן המקורית.

2. במפעל קטן ישנם חמישה עובדים.
משכורות העובדים (בשקלים) הן: 9,000 , 9,000 , 7,000 , 5,000 , 5,000
א. חשבו את הממוצע ואת סטיית התקן של שכר העובדים.
ב. עקב המצב הכלכלי המורכב, שכרם של שני העובדים בעלי השכר הגבוה הופחת מ- 9,000 שקלים ל- 8,000 שקלים.
(1) האם השכר הממוצע של חמשת העובדים גדול, קטן או שווה לשכר הממוצע המקורי שלהם? נמקו ללא חישובים.
(2) נסו לשער האם סטיית התקן של שכר חמשת העובדים גדולה, קטנה או שווה לסטיית התקן המקורית של שכרם. נמקו.
(3) חשבו את סטיית התקן החדשה של שכר חמשת העובדים.
השוו לתת סעיף ב(2)?

3. במבחן כיתתי חישבו ומצאו שהציון הממוצע הוא 80 .
לאחר בדיקה נוספת הוחלט לשנות שני ציונים:
ציון אחד הוחלט להעלות מ- 70 ל- 80 וציון שני הוחלט להוריד מ- 90 ל- 80 .
א. האם בעקבות השינויים, הממוצע גדל, קטן או נשאר ללא שינוי? נמקו.
ב. הסבירו באופן מילולי מדוע בעקבות השינויים סטיית התקן קטנה.
ג. האם ייתכן שבעקבות השינויים סטיית התקן היא אפס? נמקו.

4. בבדיקת השכר של עובדים במפעל חישבו שהשכר החודשי הממוצע הוא 6,400 שקלים.
לאחר בדיקה הוחלט להעלות את השכר לאחד העובדים מ- 6,400 ל- 6,800 שקלים,
ולהוריד את השכר לעובד אחר מ- 6,400 ל- 6,000 שקלים.
א. האם ממוצע השכר החדש גדול יותר מהממוצע המקורי, קטן ממנו או שווה לו?
ב. האם סטיית התקן החדשה גדולה יותר מהמקורית, קטנה ממנה או שווה לה? נמקו.

5. במבחן שנערך ל- 28 תלמידים הציון הממוצע הוא 80 , וסטיית התקן היא 10 .
לאחר בדיקה נוספת הוחלט לשנות שלושה ציונים:
ציון אחד הוחלט להעלות מ- 80 ל- 85 , ציון שני הוחלט להוריד מ- 80 ל- 78
וציון שלישי הוחלט להוריד מ- 80 ל- 77 .
א. הראו שממוצע הציונים נשאר ללא שינוי.
ב. האם סטיית התקן גדלה, קטנה או נשארה ללא שינוי? נמקו ללא חישובים.

6. מורה חישב שממוצע הציונים של תלמידיו הוא 70, וסטיית התקן של הציונים היא 10.
המורה הוסיף ציון של תלמיד חדש. הציון של התלמיד שהתווסף הוא 70.
א. האם חל שינוי בממוצע הכיתתי בעקבות הוספת הציון? נמקו.
ב. הסבירו באופן מילולי, מדוע כאשר הוסיפו את הציון ה- 41, סטיית התקן קטנה.

7. מורה חישב ומצא שממוצע הציונים של תלמידים הוא 60, וסטיית התקן היא 1.8.
בהמשך הוסיף המורה ציון של תלמיד נוסף, והתברר שהממוצע של **כל** התלמידים נשאר 60.
א. מהו הציון של התלמיד הנוסף? נמקו את תשובתכם.
ב. האם סטיית התקן לאחר הוספת הציון גדולה, קטנה או שווה לסטיית התקן לפני ההוספה? אין צורך בחישוב אלגברי.
ג. בחרו את התשובה הנכונה: אם נתונה קבוצת נתונים שסטיית התקן שלהם אינה אפס, ומוסיפים נתון שערכו שווה לממוצע, סטיית התקן גדלה/ קטנה / לא משתנה.
ד. (1) האם הטווח של הציונים גדל, קטן או לא השתנה בעקבות הוספת הציון?
(2) האם הטווח משקף לנו שפיזור הציונים השתנה בעקבות הצטרפות התלמיד?

8. מורה חישב ומצא שממוצע הציונים של תלמידיו הוא 60 וסטיית התקן היא 4.
א. תלמיד שהציון שלו הוא 60 עזב את הכיתה, ולכן הציון שלו נמחק.
האם חל שינוי בממוצע הכיתתי? נמקו את תשובתכם.
ב. הסבירו באופן מילולי מדוע כאשר מחקו את הציון של התלמיד, סטיית התקן גדלה.

9. ממוצע הציונים של תלמידים במבחן הוא 75. התברר שאחד התלמידים שקיבל ציון 75 אינו שייך לכיתה, לכן מחקו את הציון שלו, וחישבו מחדש את הממוצע ואת סטיית התקן.
א. מהו הציון הממוצע של כל התלמידים שנותרו? נמקו.
ב. האם בעקבות המחיקה סטיית התקן גדלה, קטנה או נותרה ללא שינוי?
(הערה: הניחו שלא כל הציונים שווים).
ג. (1) האם בעקבות מחיקת הציון הטווח של הגובה גדל, קטן או לא השתנה?
(2) האם הטווח משקף לנו במקרה זה שפיזור הציונים השתנה בעקבות מחיקת הציון?

10. לפניכם חמישה מספרים: 90, 80, 70, 60, 50
א. חשבו את הממוצע ואת סטיית התקן של המספרים.
ב. <u>נסו להעריך</u> מה יקרה לסטיית התקן של חמשת המספרים, אם נוסיף לרשימה:
(1) את המספר 90. (2) את המספר 70. (3) את המספר 75.

11. נבדקו ציונים של קבוצה הכוללת 10 תלמידים.
ממוצע הציונים שהתקבל הוא 8, וסטיית התקן היא 0.5.
לקבוצה נוספו שני תלמידים: אחד שהציון שלו 7 ושני שהציון שלו 9.
א. הסבירו מדוע הממוצע של 12 התלמידים שווה לממוצע המקורי.
ב. הסבירו ללא חישובים מדוע הוספת שני התלמידים הגדילה את סטיית התקן.
ג. תנו דוגמה לציון שאפשר להוסיף לקבוצה המקורית שיקטין את סטיית התקן המקורית.

12. בדקו את משקלם של 10 ספורטאים. המשקל הממוצע היה 75 ק"ג, וסטיית התקן 10 ק"ג.
למדידה הצטרפו ארבעה ספורטאים.
א. מה צריך להיות המשקל של כל אחד מהספורטאים שהצטרפו כדי שסטיית התקן של 14 הספורטאים תהיה <u>מינימלית</u>?
ב. תנו דוגמה למשקל של שני ספורטאים שאפשר להוסיף לקבוצה המקורית, כך שסטיית התקן המקורית תגדל.

13. להלן התפלגות הציונים של קבוצת תלמידים במבחן :

הציון	7	8	9
מספר התלמידים	11	5	11

א. חשבו את הממוצע ואת סטיית התקן של הציונים.
ב. שלושה תלמידים שנעדרו מהמבחן נבחנו במבחן חוזר. כל אחד משלושת התלמידים האלה קיבל את הציון 10 . המורה צירף ציונים אלה לציונים של שאר התלמידים.
(1) האם הציון הממוצע של התלמידים גדל, קטן או נשאר ללא שינוי? נמקו ללא חישובים.
(2) נסו לשער האם סטיית התקן גדלה, קטנה או נשארה ללא שינוי.
(3) חשבו את סטיית התקן החדשה של הציונים.
האם התוצאה תואמת את תשובתכם לתת סעיף ב(2)?

תשובות:

1. א. 166 ס״מ, 2.915 ס״מ. ב. (1) גדל. הסבר : אם גובהם של שניים גדל, ושל האחרים נותר ללא שינוי, אז סכום הגבהים גדל, וכך גם הממוצע. (2) 167.25 ס״מ.
(3) הפיזור הגבהים קטן, ולכן סטיית התקן קטנה. (4) 1.639 ס״מ.
2. א. ממוצע : 7,000 שקלים, סטיית תקן : 1,789 שקלים. ב. (1) קטן. (2) קטנה. (3) 1,356 שקלים.
3. א. נשאר ללא שינוי. ב. הציונים שהוחלפו לאחר הבדיקה היו רחוקים 10 נקודות מהממוצע. הציונים המחליפים שווים לממוצע ולכן מקטינים את הפיזור (ואת סטיית התקן).
ג. כן, אם לאחר התיקון כל הציונים שווים ל- 80 , אז סטיית התקן תהיה אפס.
4. א. נשאר ללא שינוי. ב. סטיית התקן גדולה יותר מהמקורית. הסבר : רמות השכר שהוחלפו שוות לממוצע, ולכן החלפתם מגדילה את הפיזור, כלומר מגדילה את סטיית התקן.
5. ב. גדלה.
6. א. לא. כאשר מוסיפים ציון השווה לממוצע, הממוצע לא משתנה.
ב. הציון שהתווסף שווה לממוצע. הוספה של משתנה שערכו שווה לממוצע מקטינה את פיזור הנתונים סביב הממוצע, כלומר מקטינה את סטיית התקן.
7. א. אם מוסיפים ציון שאינו משפיע על הממוצע, הרי הוא שווה לממוצע, כלומר שווה ל- 60 .
ב. הציון הנוסף שווה לממוצע, לכן הוא מקטין את הפיזור, כלומר מקטין את סטיית התקן.
ג. קטנה. ד. (1) לא השתנה. (2) לא משקף.
8. א. לא. כאשר מוחקים ציון השווה לממוצע, הממוצע לא משתנה.
ב. הציון שנמחק שווה לממוצע. מחיקה של משתנה שערכו שווה לממוצע מגדילה את פיזור הנתונים סביב הממוצע, כלומר מגדילה את סטיית התקן.
9. א. הממוצע של התלמידים שנותרו הוא 75 . הסבר : כאשר מוחקים משתנה שערכו שווה לממוצע, אז הממוצע לא משתנה. ב. הציון שנגרע שווה לממוצע, לכן הוא מגדיל את הפיזור, כלומר מגדיל את סטיית התקן. ג. (1) לא השתנה. (2) לא משקף.
10. א. ממוצע - 70 , סטיית תקן - 14.14 . ב. (1) הממוצע יגדל, סטיית התקן תגדל.
(2) הממוצע לא ישתנה, סטיית התקן תקטן. (3) הממוצע יגדל, סטיית התקן תקטן.
11. א. הממוצע של שני המצטרפים שווה לממוצע המקורי. ב. הציונים שהתווספו רחוקים מהממוצע מרחק הגדול מסטיית התקן 0.5 , ולכן סטיית התקן קטנה. ג. ציון 8 .
12. א. 75 ק״ג. ב. 90 ק״ג ו- 60 ק״ג.
13. א. הממוצע הוא 8 , סטיית התקן היא 0.903 . ב. (1) הממוצע של התלמידים שהצטרפו הוא 10 , והוא גדול יותר מהממוצע המקורי, ולכן הממוצע גדל. (2) גדלה. (3) 1.046 .

