

סטודנטים יקרים

לפניכם ספר תרגילים בקורס מבוא לסטטיסטיקה והסתברות א'.
הספר הוא חלק מקורס חדשני וראשון מסוגו בארץ בנושא זה,
המועבר ברשת האינטרנט On-line.

הקורס באתר כולל פתרונות מלאים לספר התרגילים, וכן את
התיאוריה הרלוונטית לכל נושא ונושא.

**הקורס כולו מוגש בסרטוני וידאו המלווים בהסבר קולי, כך שאתם
רואים את התהליכים בצורה מובנית, שיטתית ופשוטה, ממש כפי
שנעשה בשיעור פרטי, לדוגמה [לחצו כאן](#).**

את הקורס בנה מר ברק קנדל, מרצה מבוקש במוסדות אקדמיים
שונים ובעל ניסיון עתיר בהוראת המקצוע.

אז אם אתם עסוקים מידי בעבודה, סובלים מלקויות למידה, רוצים
להצטיין או פשוט אוהבים ללמוד בשקט בבית, אנחנו מזמינים אתכם
לחויית לימודים יוצאת דופן וחדשה לחלוטין, היכנסו עכשיו לאתר
www.gool.co.il.



אנו מאחלים לכם הצלחה מלאה בבחינות

צוות האתר GooL

תוכן

3.....	פרק 1 - סטטיסטיקה תיאורית - הקדמה
8.....	פרק 2 - סטטיסטיקה תיאורית - הצגה של נתונים
16.....	פרק 3 - סטטיסטיקה תיאורית - גבולות מדומים וגבולות אמתיים
18.....	פרק 4 - סטטיסטיקה תיאורית - סכימה
21.....	פרק 5 - סטטיסטיקה תיאורית - מדדי מיקום מרכזי
31.....	פרק 6 - סטטיסטיקה תיאורית - מדדי פיזור: הטווח, השונות וסטיית התקן
36.....	פרק 7 - סטטיסטיקה תיאורית - מדדי פיזור - טווח בין-רבעוני
39.....	פרק 8 - סטטיסטיקה תיאורית - מדדי מיקום יחסי - ציון תקן
42.....	פרק 9 - סטטיסטיקה תיאורית - מדדי מיקום יחסי - אחוזונים במחלקות
47.....	פרק 10 - סטטיסטיקה תיאורית - טרנספורמציה לינארית
50.....	פרק 11 - סטטיסטיקה תיאורית - מקדם ההשתנות
53.....	פרק 12 - סטטיסטיקה תיאורית - שאלות מסכמות
62.....	פרק 13 - מדדי קשר - מדד הקשר של קרמר
65.....	פרק 14 - מדדי קשר - מדד הקשר של ספירמן
69.....	פרק 15 - מדדי קשר - מדד הקשר הלינארי (פירסון)
77.....	פרק 16 - מדדי קשר - רגרסיה ליניארית
80.....	פרק 17 - בעיות בסיסיות בהסתברות
85.....	פרק 18 - פעולות בין מאורעות (חיתוך ואיחוד), מאורעות זרים ומכילים
95.....	פרק 19 - קומבינטוריקה - כלל המכפלה
99.....	פרק 20 - קומבינטוריקה - תמורה - סידור עצמים בשורה
103.....	פרק 21 - קומבינטוריקה - דגימה סידורית ללא החזרה ועם החזרה
106.....	פרק 22 - קומבינטוריקה - דגימה ללא סדר וללא החזרה
110.....	פרק 23 - הסתברות מותנית - במרחב מדגם אחיד
113.....	פרק 24 - הסתברות מותנית - מרחב לא אחיד
116.....	פרק 25 - תלות ואי תלות בין מאורעות
120.....	פרק 26 - התפלגויות בדידות מיוחדות - התפלגות בינומית
125.....	פרק 27 - התפלגויות רציפות מיוחדות - התפלגות נורמלית
133.....	פרק 28 - התפלגות הדגימה

פרק 1 - סטטיסטיקה תיאורית - הקדמה

רקע:

בסטטיסטיקה תיאורית אנו חוקרים קבוצה מסוימת. הקבוצה יכולה להיות קבוצת ילדים בגן, קבוצת מניות בתיק, כלל התושבים בעיר מסוימת וכולי. בין ישות לישות בקבוצה ישנם גורמים היכולים לקבל מספר ערכים. גורמים אלה נקראים משתנים. למשל, בין מניה למניה בתיק משתנה התשואה היומית של המניה, הוותק של המניה, תחום המניה וכדומה.

בסטטיסטיקה תיאורית אנחנו נתבונן בקבוצה מסוימת ובתוך הקבוצה הזו נאסוף נתונים לגבי משתנה מסוים ונלמד להציג את הנתונים ולנתח אותם מכל מיני אספקטים.

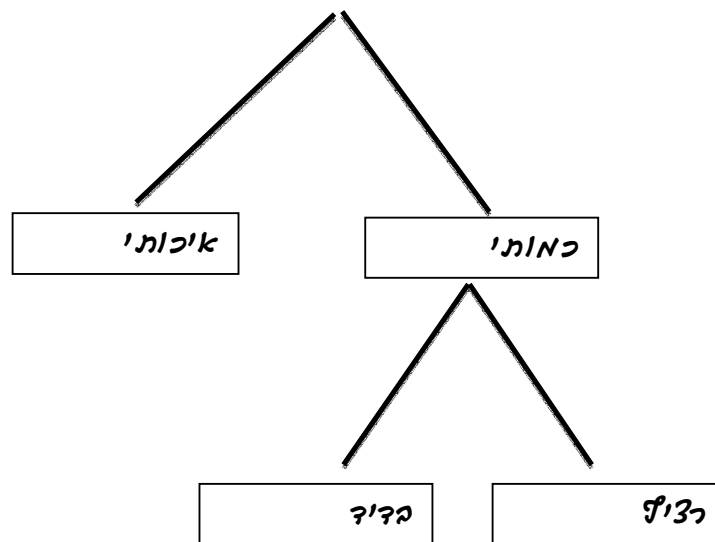
דוגמה:

בתיק מניות 10 מניות. מנהל התיק פרסם את התשואה של כל מניה בשנת 2011.

מי הקבוצה הנחקרת?

מה גודל הקבוצה?

מה המשתנה הנחקר?

סוגי משתנים:

משתנה איכותי הוא משתנה שלערכיו אין משמעות של יותר או פחות, אין עניין כמותי לערכים המתקבלים.

כמו: מקום מגורים של אדם (רעננה, תל אביב, אשדוד..)

מין האדם (זכר, נקבה)

מצב משפחתי (רווק, נשוי, גרוש, אלמן)

משתנה כמותי הוא משתנה שערכיו הם מספרים להם יש משמעות כמותית כמו: גובה אדם בס"מ, ציון בבחינה וכדומה.

את המשתנה הכמותי נסווג לשני סוגים:

משתנה בדידי: משתנה שערכיו מתקבלים מתוך סידרה של ערכים אפשריים. כמו: מספר ילדים למשפחה (1,2,3..)

ציון בבחינה (מ 0 ועד 100 בקפיצות של 1)

הערה:

משתנה רציף: משתנה שערכיו מתקבלים מתוך אינסוף ערכים בתחום מסוים, הערכים מתקבלים ברצף וללא קפיצות של ערכים.

כמו: גובה בס"מ – אם למשל, הגובה הנמוך ביותר הוא 150 ועד 190 ס"מ בקבוצה הגבהים הם ברצף. גם בין 160 ל 161 ס"מ יש רצף אינסופי של ערכים אפשריים לגובה (160.33 ס"מ הוא גם גובה אפשרי)

משקל בק"ג, מהירות בקמ"ש וכולי.

תרגילים:

1. סווג את המשתנים הבאים לפי: איכותי / כמותי בדיד / כמותי רציף:
- מספר הדירות בבניין.
 - גיל אדם בשנים.
 - אחוז האבטלה בעיר.
 - מקצוע לימוד מועדף.
2. להלן התפלגות מספר האיחורים לעבודה בחודש של העובדים בחברת "סטאר".
בחברה 200 עובדים.

מספר העובדים	מספר האיחורים
17	0
23	1
85	2
50	3
25	4

- מהו המשתנה הנחקר כאן?
- האם מדובר במשתנה איכותי או כמותי? אם הוא כמותי האם הוא בדיד או רציף?
- לפניכם רשימה של משתנים כמותיים. ציין ליד כל משתנה אם הוא רציף או בדיד.
 - שכר עובד בש"ח.
 - ציון בחינת בגרות.
 - תוצאה בהטלת קובייה.
 - מהירות ריצה בתחרות.
 - שיעור התמיכה בממשלה.

1. איזו טרנספורמציה שומרת על סולם מספרי האוטובוסים של "אגד" ?
- א. טרנספורמציה שומרת סדר.
 - ב. טרנספורמציה לינארית חיובית.
 - ג. טרנספורמציה שומרת יחס.
 - ד. כל התשובות נכונות.

פתרונות:

ד.2

ג.3

ד.4

פרק 2 - סטטיסטיקה תיאורית - הצגה של נתונים

רקע:

דרכים להצגת נתונים שנאספו :

א. 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 :

3 4 3 5 4

ב. טבלת שכיחויות בדידה :

שם המשתנה X	שכיחות $f(X)$	שכיחות יחסית באחוזים
X_1	f_1	$\frac{f_1}{N} \times 100$
X_2	f_2	$\frac{f_2}{N} \times 100$
X_3	f_3	$\frac{f_3}{N} \times 100$
\vdots	\vdots	\vdots
X_k	f_k	$\frac{f_k}{N} \times 100$
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.

למשל, להלן התפלגות הציונים בכיתה מסוימת:

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

שכיחות מצטברת – צבירה של השכיחותות: 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.

ג. **טבלת שכיחותות במחלקות:**

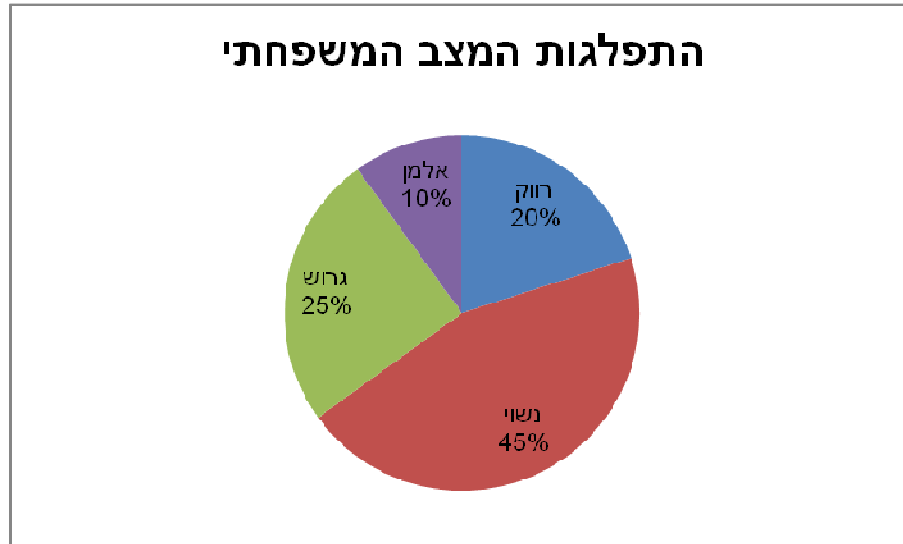
משתמשים שהמשתנה כמותי רציף או כאשר יש מספר ערכים רב במשתנה הבדיד וטבלת שכיחותות תהיה ארוכה מידי.

למשל, נתנו לקבוצת ילדים לבצע משימה מסוימת ובדקו את התפלגות זמן ביצוע המשימה בדקות. להלן ההתפלגות שהתקבלה:

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

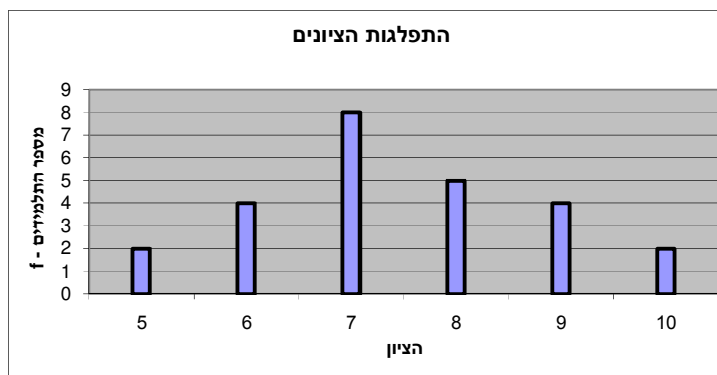
ד. דיאגרמת עוגה :

זהו התיאור הגרפי של משתנה איכותי. בדיאגרמת עוגה כל ערך במשתנה מקבל "נתח" יחסי מהעוגה. הנתח בעוגה פרופורציוני לשכיחות היחסית של ערך המשתנה בנתונים.



ה. דיאגרמת מקלות :

הציר האופקי הוא הציר של המשתנה הציר האנכי של השכיחות – הגובה של המקל מעיד על השכיחות. רלבנטי למשתנה כמותי בדיד. לא נהוג להשתמש בתיאור למשתנה איכותי וכמו כן לא למשתנה כמותי רציף. כמו כן בסולמות מדידה עבור משתנה מסולם סדר.

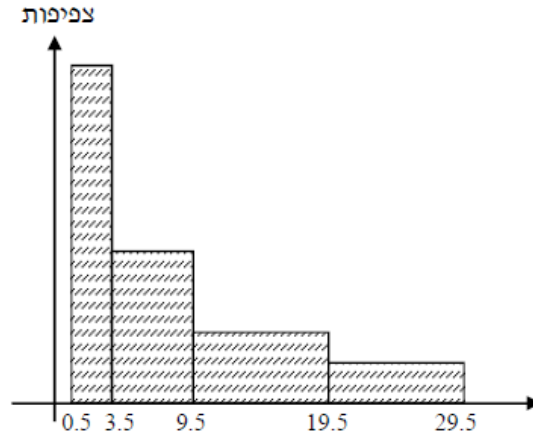


1. **היסטוגרמה:**

ההיסטוגרמה היא הדרך הגרפית כדי לתאר טבלת שכיחויות במחלקות. רלבנטית למשתנה כמותי רציף.

בהיסטוגרמה ציר האופקי הוא הציר של המשתנה וציר 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 ליחידה. אם המחלקות הן שוות ברוחב, ניתן לשרטט את ההיסטוגרמה לפי השכיחות ואין צורך בצפיפות.

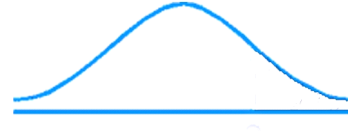
צפיפות	מצטברת	שכיחות	רוחב		X
			אמצע	רוחב	
6.6667	20	20	2	3	0.5 - 3.5
3	38	18	6.5	6	3.5 - 9.5
1.4	52	14	14.5	10	9.5 - 19.5
0.8	60	8	24.5	10	19.5 - 29.5



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

צורות התפלגות נפוצות

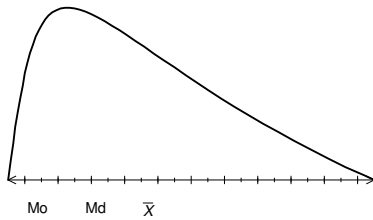
התפלגות סימטרית פעמונית- רוב התצפיות במרכז וככל שנתרחק מהמרכז יהיו פחות תצפיות באופן סימטרי. למשל, ציוני IQ.



ישנן התפלגויות סימטריות שאינן פעמוניות:

התפלגות אסימטרית ימנית (חיובית) –רוב התצפיות מקבלות ערכים נמוכים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים גבוהים קיצוניים. למשל, שכר במשק.

**התפלגות א-סימטרית
ימנית או חיובית**



התפלגות אסימטרית שמאלית (שלילית) רוב התצפיות מקבלות ערכים גבוהים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים נמוכים קיצוניים. למשל, אורך חיים.

**התפלגות
א-סימטרית
שמאלית או
שלילית**



תרגילים:

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

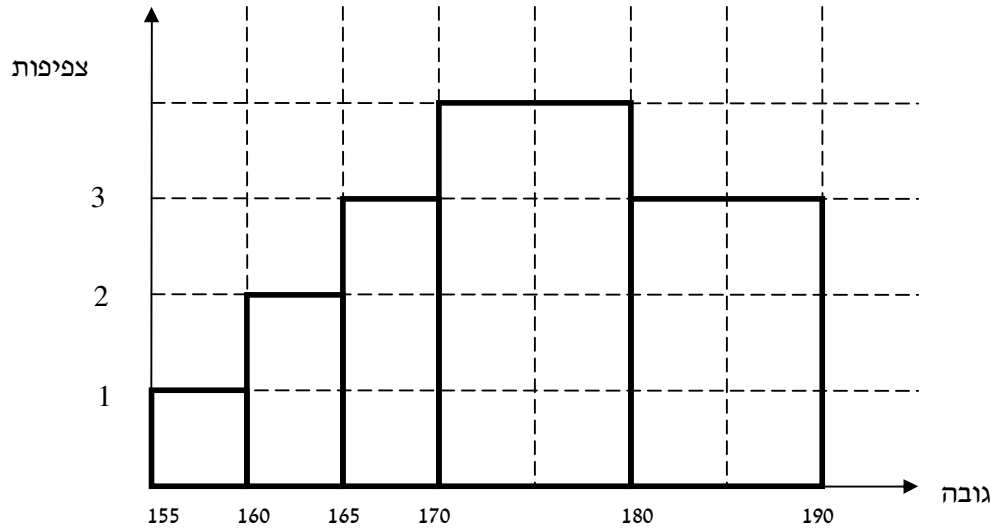
- א. מהו המשתנה הנחקר?
 ב. מהי פרופורציית התלמידים שמעדיפים תנ"ך?
 3. להלן התפלגות ההשכלה במקום עבודה מסוים:

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:
 7, 6, 8, 9, 10, 6, 4, 5, 8, 7, 6, 7, 6, 8, 9, 6, 7, 8, 5, 6

- א. מהו המשתנה? האם הוא בדיד או רציף?
 ב. תאר את הרשימה בטבלת שכיחות.
 ג. הוסף שכיחות יחסיות לטבלה.
 ד. תאר את הנתונים באופן גרפי.

5. להלן היסטוגרמה המתארת את התפלגות הגבהים בס"מ של קבוצה מסוימת:



- מהו המשתנה הנחקר? האם הוא בדיד או רציף?
- תאר את הנתונים בטבלת שכיחויות במחלקות.
- הוסף שכיחות יחסית לטבלה.
- הוסף את הצפיפות של כל מחלקה לטבלה.
- מהי צורת ההתפלגות של הגבהים?

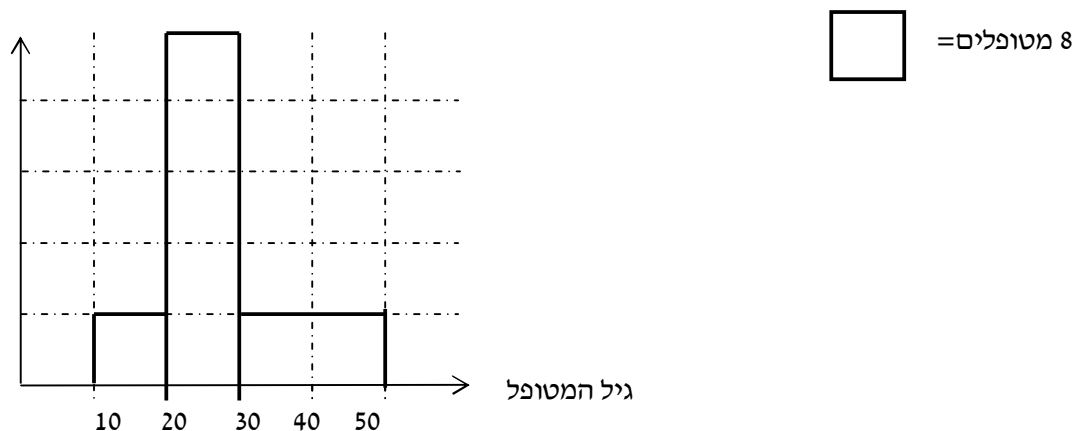
6. להלן התפלגות המשקל של קבוצה מסוימת בק"ג:

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

- תאר את ההתפלגות באופן גרפי.
- מה ניתן להגיד על צורת ההתפלגות?

7. להלן גיל המטופלים של ד"ר שוורץ בשנים :

קנה מידה :



א. מה המשתנה הנחקר? האם הוא בדיד או רציף?

ב. מהי הקבוצה הנחקרת?

ג. תרגמו את ההסיטוגרמה לטבלת שכיחות.

ד. מהי הפרופורציה של המטופלים של ד"ר שוורץ בגילאים 20-30?

פרק 3 - סטטיסטיקה תיאורית - גבולות מדומים וגבולות אמתיים

רקע:

עבור משתנה רציף נהוג לתאר את הנתונים בטבלת שכיחויות במחלקות. הנתונים שנאספים הם ברמת דיוק מסוימת. לדוגמא משקל של בני אדם או משקל של יהלומים ישקלו ברמת דיוק שונה. **גבולות מדומים:** כאשר גבול עליון של מחלקה אחת שונה מגבול תחתון של המחלקה הבאה אז הגבולות הם גבולות מדומים. כשהגבולות מדומים ההפרש בין גבול תחתון של מחלקה לבין גבול עליון של המחלקה הקודמת יהיה רמת הדיוק. **רמת הדיוק חייבת להיות קבועה** אין אפשרות שחלק מהאנשים נדייק ברמה אחת ואת השאר ברמה אחרת. בגלל שהמשתנה הוא משתנה רציף כשננתח את הנתונים נעבור מגבולות מדומים לגבולות אמתיים. אם הנתונים יינתנו בגבולות מדומים נהפוך אותם תמיד לגבולות אמתיים. כיצד עוברים מגבולות מדומים לגבולות אמתיים? לוקחים את רמת הדיוק ומחלקים אותה ב-2 את התוצאה המתקבלת מוסיפים לגבולות העליונים ומפחיתים מהגבולות התחתונים. אם יתנו נתונים בגבולות מדומים אנחנו מוכרחים לעבור לגבולות אמתיים על מנת להמשיך ולנתח, אך אם הנתונים כבר יינתנו בגבולות אמתיים נשאיר אותם כמו שהם.

דוגמה: (פתרון בהקלטה)

להלן התפלגות הגבהים בס"מ של תלמידי כיתה ח'. יש להעביר את הנתונים לגבולות אמתיים.

$f(x)$	X
20	130-139
25	140-149
30	150-159
20	160-169
10	170-189

תרגילים:

1. להלן התפלגות של משתנה בהצגה של מחלקות. יש להעביר את הנתונים לגבולות אמתיים:

f(x)	X
542	500-590
32	600-690
154	700-790
254	800-890

2. להלן התפלגות המשקלים בק"ג של קבוצת אנשים מסוימת. יש לרשום את הנתונים בגבולות אמתיים.

משקל בק"ג	מספר אנשים
60-64	18
65-69	24
70-79	52
80-89	19

פרק 4 - סטטיסטיקה תיאורית - סכימה

רקע:

בסטטיסטיקה ישנה צורת רישום מקובלת כדי לרשום סכום 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

תרגילים:

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

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

חשבו:

$$\sum_{i=1}^3 X_i$$

$$\sum_{i=1}^5 Y_i$$

$$\sum_{i=1}^4 X_i$$

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

$$\sum X_i$$

$$\sum X_i Y_i$$

$$\sum(X_i) \sum(Y_i)$$

2. נתון לוח ערכי המשתנים x_i ו- y_i כאשר: $i=1,2,\dots,6$

i	1	2	3	4	5	6
x_i	3	2	4	-2	1	4
y_i	2	0	0	1	-5	2

ונתונים הקבועים: $b=5$ $a=2$ חשבו את הנוסחאות הבאות:

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. קבע לכל זהות אם 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. $(\sum_{i=1}^n X_i)^2 = \sum_{i=1}^n X_i^2$

פרק 5 - סטטיסטיקה תיאורית - מדדי מיקום מרכזי

רקע:

המטרה במדדי המיקום המרכזי למדוד את מרכז ההתפלגות של התצפיות.

השכיח – MODE

השכיח הוא הערך הנפוץ ביותר בהתפלגות.

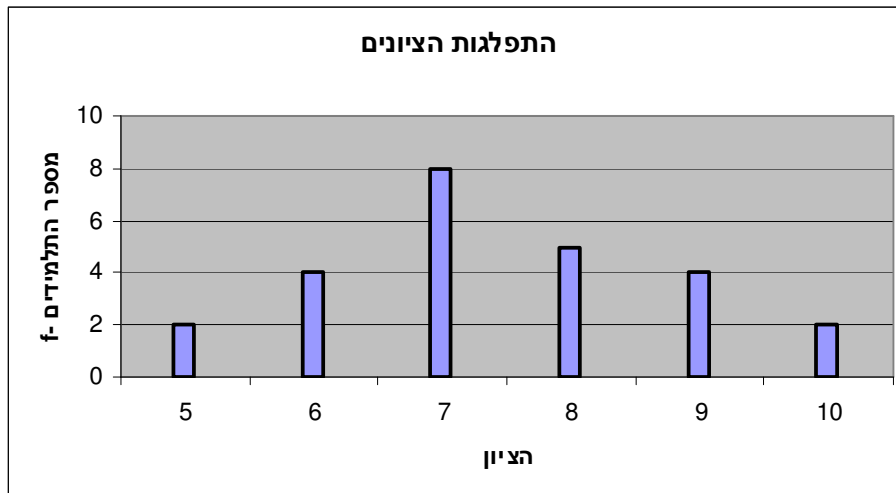
ברשימה : הערך החוזר על עצמו הכי הרבה פעמים .

7 9 4 8 4 10 6

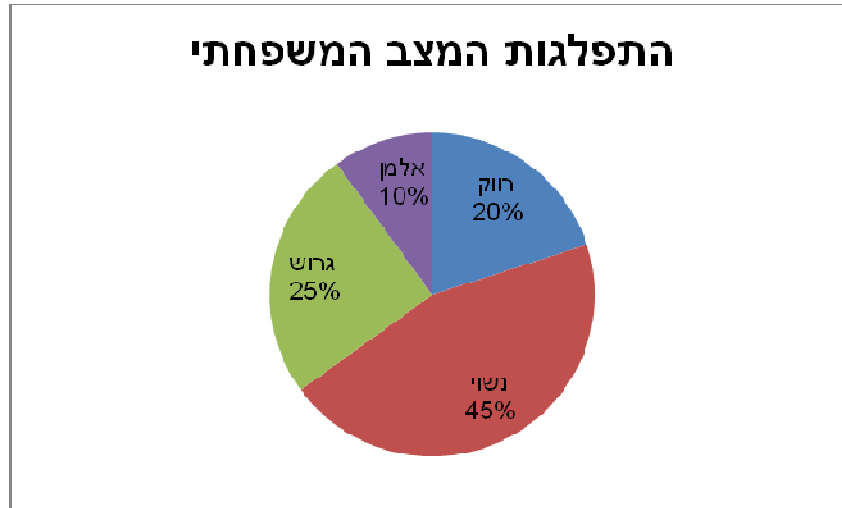
בטבלת שכיחויות בדידה : הערך שהשכיחות שלו היא הגבוהה ביותר.

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

בדיאגרמת מקלות : שיעור ה- X של המקל הגבוה ביותר.



בעוגה: הערך של הפלח הגדול ביותר.

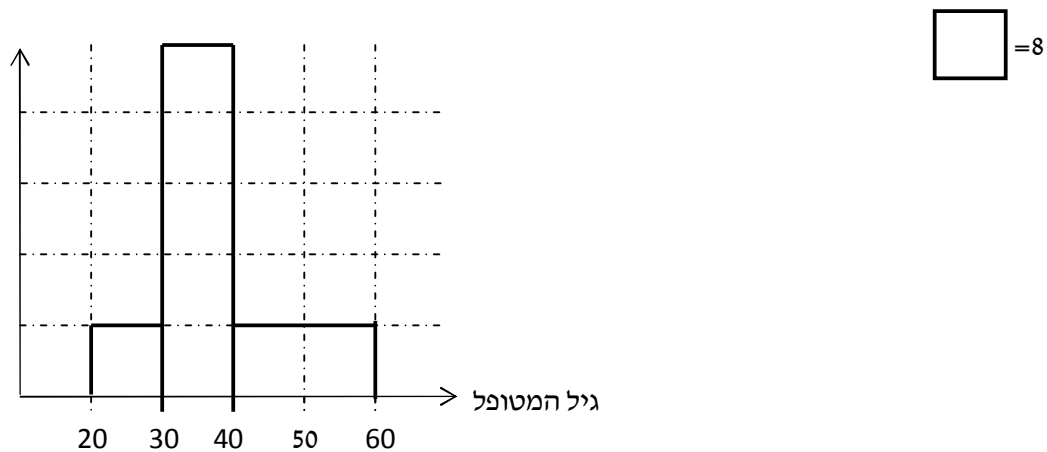


בטבלת שכיחויות במחלקות: אמצע המחלקה עם הצפיפות הגבוהה ביותר. התפלגות הציונים בכיתה.

$f(x)$	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.

להלן גיל המטופלים של ד"ר שוורץ בשנים:



כללי: יתכן שהתפלגות יותר משכיח אחד. השכיח הוא מדד הרלבנטי לכל סוגי המשתנים.

MIDRANGE – (טווח)

הממוצע בין התצפית הגבוהה ביותר לתצפית הנמוכה ביותר.

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

MEDIAN - החציון

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 יהיה הממוצע של האיבר ה- $\frac{n}{2}$ והאיבר ה- $\frac{n}{2} + 1$

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

ושיש 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.
בטבלת שכיחויות במחלקות:

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

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

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

$f(x_m)$ - 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

הנו מרכז הכובד של ההתפלגות.

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

בטבלת שכיחויות: $\bar{x} = \frac{\sum x \cdot f}{n}$

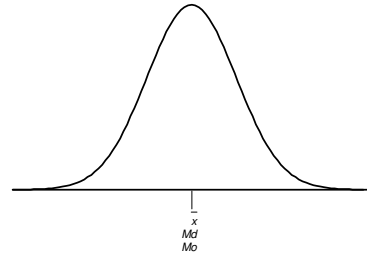
במחלקות: נשתמש באותה נוסחה רק נתייחס לאמצע המחלקה בתור ה- X . הממוצע הזה יהיה ממוצע מקורב.

כללי: הממוצע רלבנטי רק למשתנה כמותי.

מדדי המיקום המרכזי בהתפלגויות המיוחדות:

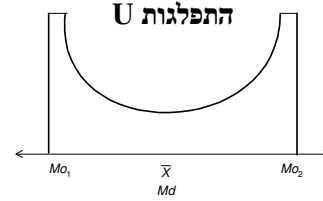
בהתפלגות סימטרית פעמונית כל מדדי המרכז שווים זה לזה:

התפלגות סימטרית



בהתפלגות סימטרית השכיח לא חייב להיות במרכז:

התפלגות U

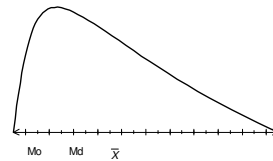


בהתפלגות אסימטרית

התפלגות
א-סימטרית
שמאלית או
שלילית



התפלגות א-סימטרית
ימנית או חיובית



תרגילים:

1. להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
7, 6, 8, 9, 10, 6, 4, 5, 8, 7, 6, 7, 6, 8, 9, 6, 7, 8, 5, 6
חשב 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: 4, 3, 4, 5.
א. כמה חדרים יש בדירה החמישית?
ב. מהו השכיח ומהו החציון?

3. להלן התפלגות מספר מקלטי הטלויזיה שנספרו עבור כל משפחה בישוב מסוים:

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

ג. חשב את הממוצע, החציון והשכיח של ההתפלגות.
ד. הסבר 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.

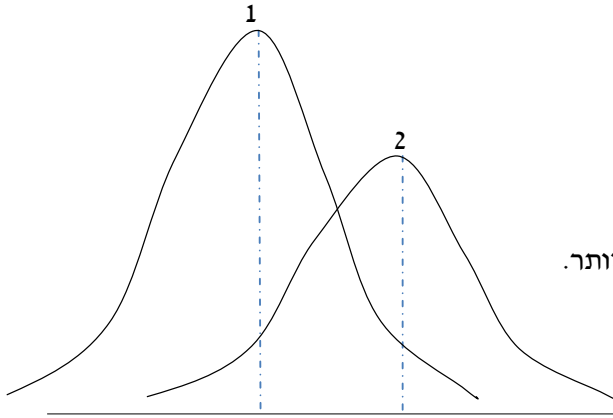
4. להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן"

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!

5. מורה לימד 2 כיתות, הוא תיאר באותה מערכת צירים את התפלגות הציונים בכל כיתה. בחר 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.

6. ביישוב מסוים בדקו לכל משפחה את מספר הטלויזיות שיש לה בבית. ביישוב גרות 200 משפחות. בממוצע יש למשפחה 1.5 טלויזיות.

מספר משפחות	מספר טלויזיות
28	0
62	1
	2
	3

א. השלימו את הטבלה.

ב. מהו השכיח, אמצע טווח והחציון.

ג. חלק מהמשפחות להן הייתה טלויזיה אחת בדיוק הוציאו את הטלויזיה מביתם, כיצד כל מדד ישתנה (יגדל, יקטן או לא ישתנה) הסבירו ללא חישוב.

7. להלן התפלגות המשקל של קבוצה מסוימת בק"ג:

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

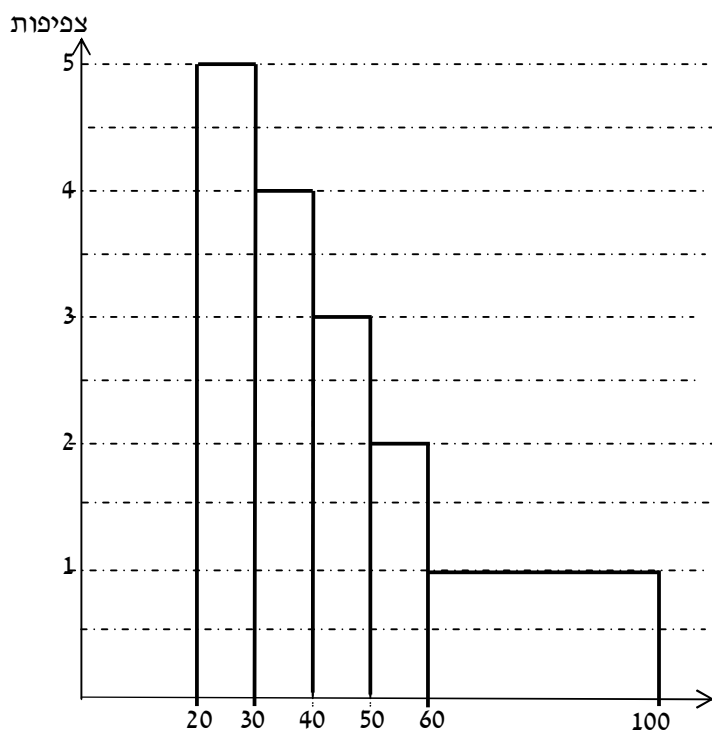
מה הממוצע והחציון של ההתפלגות?

8. להלן התפלגות הגבהים בס"מ בקבוצה מסוימת.

גובה בס"מ	שכיחות
150-160	30
160-170	40
170-175	60
175-180	70
180-190	40

חשב את הממוצע, החציון והשכיח של הגבהים בקבוצה זו.

9. בפקולטה מסוימת בדקו לסטודנטים העובדים בה את השכר לשעת עבודה. להלן התוצאות:



א. מצא את השכיח בהתפלגות.

ב. מצא את החציון בהתפלגות.

ג. הסבירו ללא חישוב האם הממוצע גדול/קטן/שווה לחציון.

ד. הסתבר שיש להוציא מספר תלמידים במחלקה בין 20-30 שקלים כיצד הדבר ישפיע על הממוצע, החציון והשכיח? הסבירו ללא חישוב.

פתרונות:**שאלה 1:**

החציון: 7

השכיח: 6

הממוצע: 6.9

שאלה 2:

א. 3

ב. שכיח: 3,4 חציון: 4

שאלה 3:

א. הממוצע: 1.7

החציון: 1.5

השכיח: 1

ב. הממוצע יגדל ויתר המדדים לא ישתנו.

שאלה 4:

א. 630

ב. 34.13%

ג. שכיח וחציון: 3

ממוצע: 2.952

שאלה 5:

תשובה ב:

שאלה 6:

ב חציון: 2 שכיח: 2 אמצע טווח: 1.5

שאלה 7:

חציון וממוצע: 55

פרק 6 - סטטיסטיקה תיאורית - מדדי פיזור : הטווח, השונות וסטיית התקן

רקע:

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

הטווחותחום RANGE:

ההפרש בין התצפית הגבוהה ביותר לנמוכה ביותר : $R = X_{\max} - X_{\min}$

שונות וסטיית תקן:

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

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2 \quad \text{עבור סדרת נתונים :}$$

דוגמה : נחשב את השונות של סדרת המספרים הבאה : 5,4,9

$$s_x^2 = \frac{\sum (x - \bar{x})^2 f}{n} = \frac{\sum x^2 \cdot f}{n} - \bar{x}^2 \quad \text{עבור טבלת שכיחויות :}$$

להלן התפלגות הציונים בכיתה מסוימת בה ממוצע הציונים הוא 7.44

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

$$s_x^2 = \frac{\sum x^2 f(x)}{n} - \bar{x}^2 = \frac{1430}{25} - 7.44^2 = 1.8464$$

$$s = \sqrt{s_x^2} = \sqrt{1.8464} = 1.3588$$

כשיש מחלקות נעזר באמצע המחלקה כדי לחשב את השונות.

תרגילים:

1. להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:

7, 6, 8, 9, 10, 6, 4, 5, 8, 7, 6, 7, 6, 8, 9, 6, 7, 8, 5, 6
 חשבו את השונות, סטיית התקן והטווח של הציונים.

2. להלן התפלגות מספר המכוניות למשפחה ב"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. חשב 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.

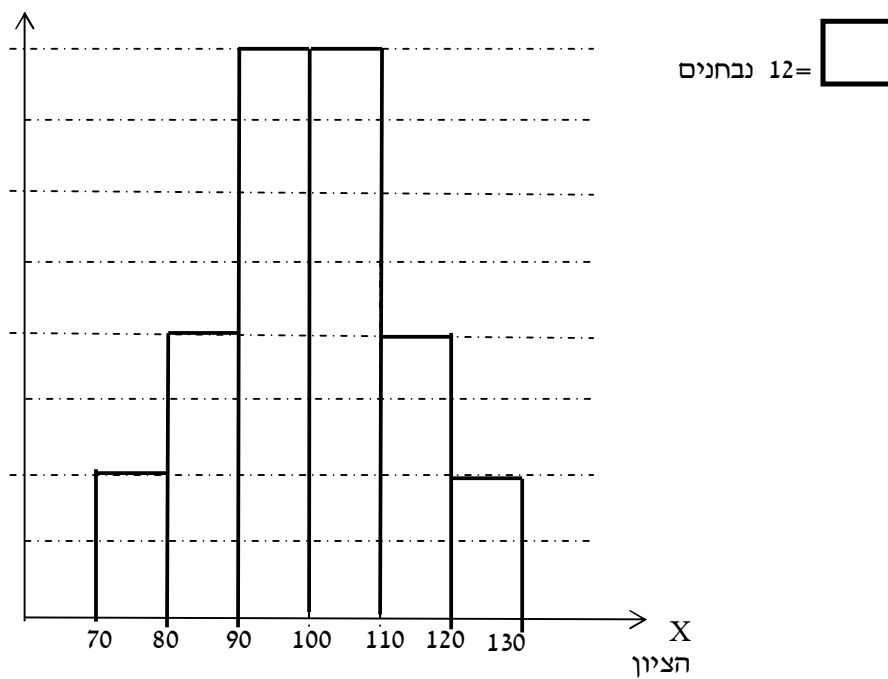
ג. חלק מבעלי הדירות בנות 2 החדרים הפכו את דירתם לדירת חדר. כיצד הדבר ישפיע (יקטין, יגדל, לא ישנה) כל מדד שחישבתם בסעיפים הקודמים.

6. להלן התפלגות המשקל של קבוצה מסוימת בק"ג:

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

מהי סטיית התקן של התפלגות המשקל?

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?

פתרונות :**שאלה 1:**

השונות : 2.19

סטיית תקן : 1.48

טווח : 6

שאלה 2:

א. סטיית תקן : 1.106

ב. טווח 4

שאלה 3:

א. ממוצע לא ישתנה, סטיית התקן תקטן.

ב. ממוצע לא ישתנה, סטיית התקן תגדל.

שאלה 4:

10.8

שאלה 5:

א. 3.05

ב. 1.16

שאלה 6:

7.73

שאלה 7:

א. 100

ב. 12.96

פרק 7 - סטטיסטיקה תיאורית - מדדי פיזור - טווח בין-רבעוני

רקע:

הטווח הבין-רבעוני נותן את הטווח בין הרבעונים בו נמצאים 50% מהתצפיות המרכזיות.

שלב א: במציאת טווח בין-רבעוני במחלקות:

F	f מספר עובדים (שכירות)	$L_1 - L_0$ רוחב	מספר שנות ותק
56	56	4	0.5 – 4.5
106	50	5	4.5 – 9.5
154	48	2	9.5 – 11.5
190	36	3	11.5 – 14.5
200	10	5	14.5 – 19.5

שלב א: נימצא את הרבעון התחתון (האחוזון ה-25) והרבעון העליון (האחוזון ה-75).

מיקום הרבעון התחתון יהיה: $\frac{n}{4}$

מיקום הרבעון העליון יהיה: $\frac{3n}{4}$

נוסחאות הרבעונים יהיו:

$$Q_1 = L_0 + \frac{\frac{n}{4} - F(x_{m-1})}{f(x_m)} \cdot (L_1 - L_0) \quad ; \quad Q_3 = L_0 + \frac{\frac{3n}{4} - F(x_{m-1})}{f(x_m)} \cdot (L_1 - L_0)$$

שלב ב: נחסר את הרבעונים:

$$IQR = Q_3 - Q_1$$

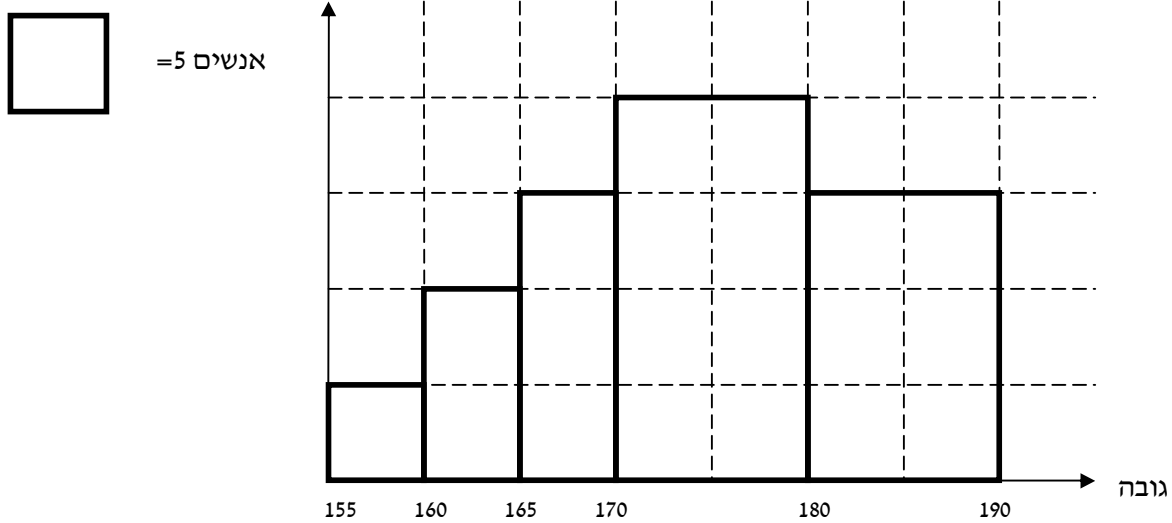
תרגילים:

1. להלן התפלגות המשקל של קבוצה מסוימת בק"ג:

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

מצא את הטווח הבין-רבעוני.

2. להלן היסטוגרמה המתארת את התפלגות הגבהים בס"מ של קבוצה מסוימת:



מצא את הטווח הבין-רבעוני.

פתרונות:

שאלה 1:

13.75

שאלה 2:

13.33

פרק 8 - סטטיסטיקה תיאורית - מדדי מיקום יחסי - ציון תקן

רקע:

המטרה למדוד איך תצפית ממוקמות יחסית לשאר התצפיות בהתפלגות.

ציון תקן:

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

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 8 אלפי ₪ עם סטית תקן של 2 אלפי ₪ 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 ?

תרגילים

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 ו70 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 האינפורמציה שנאספה על שני המשתנים:

מספר פועלים	תפוקה	
15	48	ממוצע
2	10	סטיית תקן

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?
בחר בתשובה הנכונה.

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 10 סנטימטר. המשקל הממוצע 66 ק"ג עם סטיית תקן 8 ק"ג. ערך התגייס, גובהו 180 ס"מ ומשקלו 59 ק"ג.
א. במה ערך חריג יותר ביחס לשאר המתגייסים- גובהו או משקלו?
ב. כמה ערך אמור לשקול כדי שמשקלו יהיה שקול לגובהו?

פתרונות:

שאלה 1:

א. לשון

ב. 72

שאלה 2:

תשובה ב

שאלה 3:

א. משקל

ב. 70

פרק 9 - סטטיסטיקה תיאורית - מדדי מיקום יחסי - אחוזונים
במחלקות

רקע:

האחוזון (המאון) ה- p הוא הערך בנתונים המחלק את הנתונים בצורה כזאת שעד אליו יש $p\%$ מהנתונים. מסמנים את האחוזון ה- p ב- X_p .
 למשל, המאון ה-25 הוא האחוזון ה-25 או הרבעון התחתון: ערך ש-רבע מהתצפיות קטנות ממנו והשאר גבוהות ממנו. מסומן: $X_{0.25}$
מציאת מאון במחלקות:

שלב א: נימצא את המחלקה הרלבנטית שמיקומה יהיה $\frac{np}{100}$.

$$x_p = L_0 + \frac{\frac{n \cdot p}{100} - F(x_{m-1})}{f(x_m)} \cdot (L_1 - L_0) \quad \text{שלב ב: נציב בנוסחה הבאה:}$$

$F(x_{m-1})$ - שכיחות מצטברת של מחלקה אחת לפני המחלקה הרלבנטית.

$f(x_m)$ - השכיחות של המחלקה הרלבנטית.

L_0 - גבול התחתון של המחלקה.

L_1 - גבול העליון של המחלקה.

אם רוצים לחלץ את אחוז התצפיות שמתחת לערך מסוים נשתמש בנוסחה הבאה:

$$P_x = \left[\frac{(x - L_0)}{(L_1 - L_0)} \cdot f(x_m) + F(x_{m-1}) \right] \cdot \frac{100}{n}$$

דוגמה : (פתרון בהקלטה)

להלן התפלגות השכר של עובדים בחברה מסוימת :

שכר בש"ח	f(x)
4000-6000	140
6000-10000	128
10000-15000	60
15000-20000	54
20000-40000	18

א. מצאו את המאון ה-40.

ב. מהו אחוז העובדים שמשתכרים מתחת ל5,000 ₪?

תרגילים:

1. להלן התפלגות השכר (באלפי שקלים) בחברה:

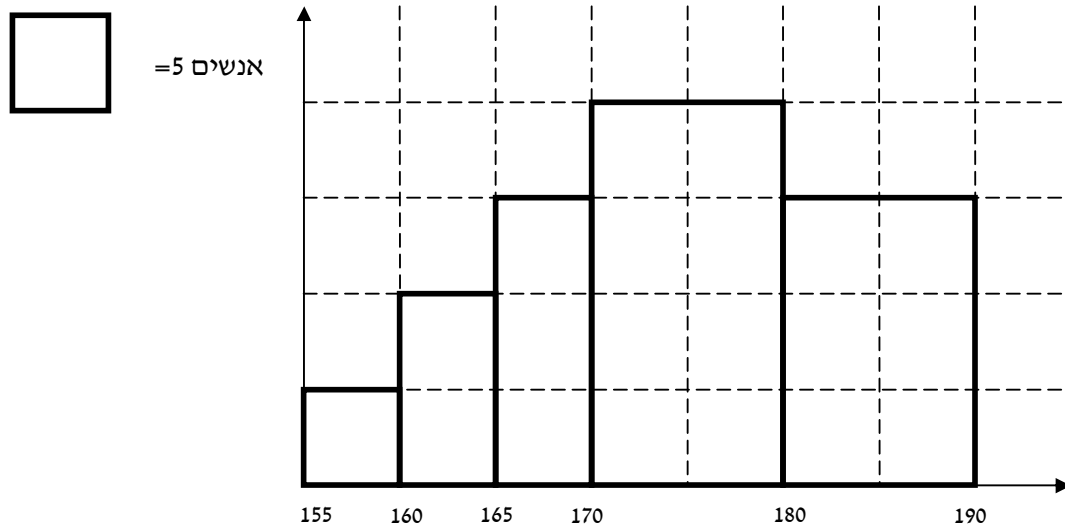
שכר X	שכיחות מצטברת
6-10	48
10-15	100
15-20	120
20-30	132
30-60	136

- א. חשבו את המאון ה-60.
- ב. מהו העשירון העליון?
- ג. 20% מהמשכורות הגבוהות ביותר הן משכורות של הבכירים, מהי המשכורת המינימאלית לבכיר?
- ד. מה אחוז האנשים שמשתכרים מתחת ל-7000 ₪?
- ה. איזה אחוז מהעובדים משתכרים מעל ל-25,000 ₪?
- ו. איזה אחוז מהעובדים משתכרים בין 7000 ל-25,000 ₪?

2. למבחן ניגשו 400 נבחנים. נתון שהעשירון התחתון הוא הציון 60. הרבעון העליון הוא הציון 80. כמו כן ההתפלגות של הציונים היא סימטרית. מלאו את השכיחות החסרות.

ציון - X	$f(X)$
50-60	
60-70	
70-80	
80-90	
90-100	

3. להלן היסטוגרמה המתארת את התפלגות הגבהים בס"מ של קבוצה מסוימת:



א. העשירון התחתון.

ב. האחוזון ה-30.

ג. הגובה ש-20% מהתצפית גדולות ממנו.

ד. את אחוז התצפיות מתחת לגובה 158 ס"מ.

ה. את אחוז התצפיות מעל לגובה 185 ס"מ.

ו. את אחוז התצפיות בין גובה 170 ס"מ ל-185 ס"מ.

פתרונות:**שאלה 1 :**

- א. 13.23
- ב. 22
- ג. 17.2
- ד. 8.82%
- ה. 7.36%
- ו. 83.82%

שאלה 3:

- א. 162.5
- ב. 170
- ג. 183.33
- ד. 3%
- ה. 15%
- ו. 55%

פרק 10 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

רקע:

מצב שבו מבצעים שינוי מסוג הוספה של קבוע (או החסרה) והכפלה של קבוע (או חילוק) לכל

$$y = a \cdot x + b \quad \text{התצפיות:}$$

וכך יושפעו המדדים השונים:

מדדי המרכז:

$$MR_y = a \cdot MR_x + b$$

$$Mo_y = a \cdot Mo_x + b$$

$$\bar{y} = a \cdot \bar{x} + b$$

$$Md_y = a \cdot Md_x + b$$

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 הנו 9000 ₪ וטווח 6000 ₪ חשבו את המדדים הללו לאחר שהעלו את כל המשכורות ב-10% ואחר כך קנסו אותם ב100 ₪.

תרגילים:

1. עבור סדרת נתונים התקבל :

$$\bar{X} = 80$$

$$S = 15$$

$$MO = 70$$

הוחלט להכפיל את כל התצפיות פי-4 ולהחסיר מהתוצאה 5. חשב 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 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. קבלו שבממוצע יש 3 גפרורים פגומים בקופסא. עם סטיית תקן של 1.5 גפרורים. מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקנים בקופסא?

5. חברת בזק הציעה את החבילה הבאה :
שלושים שקלים דמי מנוי חודשיים קבועים. ובנוסף 10 אגורות לכל דקה של שיחה יוצאת, אדם בדק במשך שנה את דקות השיחות היוצאות שלו, וקיבל שבממוצע בחודש יש לו 600 דקות שיחות יוצאות עם שונות 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.

6. הוכח שאם כל התצפיות בהתפלגות עברו טרנספורמציה לינארית :

$$Y_i = a \cdot X_i + b$$

אזי הממוצע והשונות של כלל התצפיות לאחר הטרנספורמציה יהיו בהתאמה :

$$\bar{y} = a \cdot \bar{x} + b$$

$$s_y^2 = a^2 s_x^2$$

פתרונות :**שאלה 2 :**

הממוצע : 46

השונות : 30.25

שאלה 1 :

הממוצע : 315

סטיית התקן : 60

השכיח : 275

שאלה 4 :

ממוצע : 37

סטיית תקן : 1.5

שאלה 3 :

טווח : 40

חציון : 77

עשירון עליון : 91

שאלה 5 :

ממוצע : 90

שונות : 25

ציון תקן : 2

פרק 11 - סטטיסטיקה תיאורית - מקדם ההשתנות

רקע:

כאשר מחשבים סטיית תקן למספר קבוצות בעלי ממוצע שונה, השוואת מידת פיזור הנתונים אינה מתייחסת לערך מרכז הנתונים (לממוצע למשל). על מנת לתת מדד פיזור המתחשב בממוצע הנתונים נחשב את **מקדם ההשתנות - Coefficient of Variation** :

$$CV = \frac{S(X)}{\bar{X}}$$

ככל שמקדם ההשתנות נמוך יותר המשתנה מרוכז יותר סביב הממוצע וככל שמקדם ההשתנות גבוהה יותר, מידת הפיזור סביב הממוצע גבוהה יותר.

תרגילים:

1. להלן נתונים לגבי ציונים במבחן באנגלית ב-3 כיתות מתוך שכבה י' בתיכון:

כיתה	ממוצע	מס' תלמידים	סטיית תקן
1	76	40	12
2	68	20	15
3	82	30	10

א. חשבו את מקדם ההשתנות בכל כיתה.

ב. מהי הכיתה הכי הטרוגנית?

2. נתונות שתי קבוצות:

הממוצע בקבוצה א 100 והשונות 100.

הממוצע בקבוצה ב 500 והשונות 400.

באיזו קבוצה מידת הפיזור יחסית קטן יותר?

3. במפעל לייצור מצברים לרכב בדקו במשך 40 ימים את התפוקה היומית (מספר מצברים

במאות) ואת מספר הפועלים שעבדו באותו היום.

להלן טבלה המסכמת את האינפורמציה שנאספה על שני המשתנים:

ממוצע	תפוקה	מספר פועלים
48	15	
10	2	

לפי קריטריון CV. בחר את התשובה הנכונה.

א. הפיזור באופן יחסי שווה בין התפוקה היומית לכמות הפועלים העובדים ביום.

ב. הפיזור יחסית יותר גדול עבור התפוקה היומית מאשר עבור מספר הפועלים ביום.

ג. הפיזור יחסית יותר גדול עבור מספר הפועלים ביום מאשר עבור התפוקה היומית.

ד. אין מספיק נתונים כדי לחשב את CV.

פתרונות:

שאלה 2:

קבוצה 2

שאלה 1:

ב. כתה ב

שאלה 3:

תשובה ב

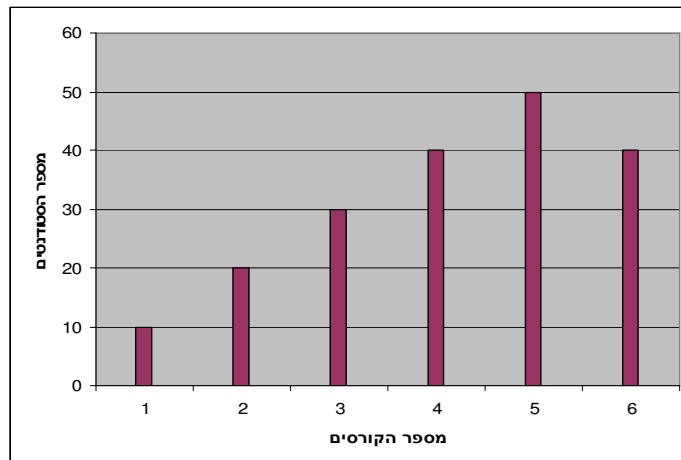
פרק 12 - סטטיסטיקה תיאורית - שאלות מסכמות

1. בדקו עבור 5 תלמידים את המשקל שלהם :

מספר תלמיד	משקל בק"ג
1	58
2	62
3	48
4	34
5	58

- א. מהו המשתנה הנחקר? האם הוא בדיד או רציף?
 ב. מהו המשקל החציוני, הממוצע והשכיח?
 ג. מה הטווח וסטיית התקן של המשקל?
 ד. לאותם תלמידים חישובו גם את הגובה בס"מ וקיבלו גובה ממוצע של 168 וסטיית תקן 6. במה תלמיד מספר 3 שגובהו 162 יותר חריג במשקל או בגובה?
 ה. הוסיפו עוד תלמיד השוקל 52 ק"ג בדיוק. הסבירו ללא חישוב כיצד הדבר ישפיע על הממוצע וסטיית התקן? (יגדיל יקטין או לא ישנה)

2. בפקולטה להנדסה אספה מזכירות הסטודנטים נתונים לגבי מס' הקורסים שכל סטודנט סיים בשנה הראשונה ללימודיו בשנת 2008.
 להלן התוצאות שהתקבלו :



- א. מה המשתנה הנחקר? האם הוא בדיד או רציף?
 ב. מהי צורת ההתפלגות?
 ג. תאר את הנתונים בטבלת שכיחויות.
 ד. חשב את השכיח, החציון והטווח.

3. להלן התפלגות הציונים בבחינה בלשון שנעשתה עבור תלמידי כיתות ד'. השתתפו במחקר 150 תלמידים.

$$\text{ממוצע הציונים שהתקבל: } \bar{X} = 7\frac{1}{15}$$

מספר התלמידים	ציון
12	4
16	5
	6
38	7
	8
14	9
10	10

- א. השלם את השכיחויות החסרות בטבלה.
 ב. חשב את הציון החציוני, השכיח.
 ג. חשב שונות וסטיית תקן להתפלגות הציונים.

4. חברה סלולארית דגמה 200 אנשים. עבור כל אדם נבדקה מידת שביעות הרצון של הלקוח מהחברה (1 – שביעות רצון נמוכה ועד 5 שביעות רצון גבוהה) להלן ההתפלגות שהתקבלה:

מספר האנשים	שביעות רצון
40	1
60	2
50	3
30	4
20	5

- א. מה אחוז האנשים עם רמת שביעות רצון נמוכה?
 ב. מה המשתנה הנחקר ומאיזה סוג הוא?
 ג. מהי הדרך הגרפית המתאימה ביותר לתיאור הנתונים?
 i. היסטוגרמה.
 ii. דיאגרמת מקלות.
 iii. דיאגרמת עוגה.
 ד. חשבו את המדדים הבאים:
 1. טווח
 2. שכיח
 3. חציון

5. להלן התפלגות מספר שעות העבודה לשבוע של העובדים בחברת "סטאר".
בחברה 200 עובדים.

שכיחות	שכיחות יחסית (פרופורציה)	מספר שעות עבודה
	15%	10-20
	20%	20-30
	30%	30-40
	20%	40-50
		50-60

- א. השלם את הטבלה.
- ב. חשב את החציון, השכיח, והממוצע של התפלגות מס' שעות העבודה בחברה.
- ג. מהי סטיית התקן של מס' שעות העבודה?
- ד. מה העשירון העליון של ההתפלגות?
- ה. איזה אחוז מהעובדים עובדים מעל 45 שעות בשבוע?
- ו. מה ציון התקן של רינה שעובדת 30 שעות בשבוע?
- ז. כיצד ישתנה החציון, הממוצע וסטיית התקן אם מספר שעות העבודה המינימאלי אינו 10 אלא 15? הסבר.

6. חברה סלולארית דגמה 200 אנשים. עבור כל אדם נבדק מס' המסרונים ששלח במשך חודש. להלן ההתפלגות שהתקבלה:

מספר האנשים	מספר המסרונים
40	0-50
60	50-100
50	100-150
30	150-250
20	250-ומעלה

- א. מה אחוז האנשים ששלחו פחות מ-80 מסרונים בחודש?
 ב. מה אחוז האנשים ששלחו בין 50 ל-120 מסרונים?
 ג. הוחלט להעניק מתנה עבור $\frac{1}{4}$ מהלקוחות שמשלמים במספר הרב ביותר של מסרונים בחודש. החל מאיזה כמות של מסרונים תחולק המתנה?
 ד. ציינו איזה מדד ניתן לחשב ואיזה לא ניתן. אם ניתן חשב:
1. ממוצע
 2. שכיח
 3. חציון
 4. שונות

7. נתנו לקבוצת ילדים לבצע משימה מסוימת ובדקו את התפלגות זמן ביצוע המשימה בדקות. להלן ההתפלגות שהתקבלה:

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

- א. שרטט היסטוגרמה לתיאור התפלגות זמן ביצוע המשימה.
 ב. מתוך ההיסטוגרמה שבנית בסעיף א מהי צורת ההתפלגות?
 ג. חשב את השכיח והחציון של ההתפלגות.
 ד. הסבר, ללא חישובים, האם הזמן הממוצע לביצוע המשימה, קטן או גדול או שווה ביחס לשכיח ולחציון.

8. התפלגות ציוני מבחן אינטיליגנציה היא סימטרית .

מספר הנבחנים	הציון
	50-70
	70-90
	90-100
	100-110
	110-130
	130-150

נתון שהעשירון העליון הוא 130 והרבעון התחתון הוא 90.

נתון שלמבחן נגשו 500 מועמדים.

א. השלימו את הטבלה.

ב. מהו הממוצע והחציון של ההתפלגות?

ג. מהו הציון ש 40% מהתלמידים קיבלו מעליו? באיזה אחוזון מדובר?

ד. אם יוחלט להעלות את כל הציונים ב-10 נקודות . כיצד הדבר ישפיע על הממוצע וסטיית התקן של הציונים?

9. להלן מספר טענות , עבור כל טענה ציין אם היא נכונה או לא נכונה ונמקו .

א. בסדרה שבה כל התצפיות שוות זו לזו השונות הינה 0.

ב. ציון התקן של החציון תמיד יהיה 0.

ג. ציון התקן של האחוזון ה-70 בהתפלגות אסימטרית ימנית (חיובית) תמיד יהיה חיובי.

ד. אם נוסיף תצפיות לסדרה של תצפיות, הדבר בהכרח יגדיל את הממוצע של הסדרה.

ה. בסדרה החציון הינו 80. הוספו שתי תצפיות אחת 79 ואחת 100 לכן החציון יגדל.

ו. אם נוסיף את הערך 4 לכל התצפיות אז סטיית התקן לא תשתנה.

ז. אם נחלק את כל התצפיות בהתפלגות ב-2 אז השונות תקטן פי 2.

ח. אם נגדיל את ממוצע המשכורות של עובדים בחברה אז גם השונות תגדל.

פתרונות:**שאלה 1:**

- א. המשתנה הנחקר כאן הוא משקל תלמיד בק"ג והוא משתנה כמותי רציף.
ב.

$$\bar{X} = 52$$

$$Md = X_{\frac{n+1}{2}} = X_3 = 58$$

השכיח הוא 58

$$R = 28 \quad \text{ג.}$$

$$s = 10.12$$

- ד. הוא חריג יותר בגובה כי שם ציון התקן בערך מוחלט יותר גבוה.
ה. הממוצע לא ישתנה אך סטיית התקן תקטן .

שאלה 2:

- א. מספר הקורסים. בדיד.
ב. התפלגות אסימטרית
שמאלית
ד. השכיח: 5
הטווח: 5

שאלה 3:

- א. 20 תלמידים קיבלו ציון 6 ו-40
תלמידים קיבלו ציון 8.
החציון: 7
השכיח: 8
ג. השונות: 2.533
סטיית התקן: 1.592

שאלה 4:

- א. 20%
ב. שביעות רצון (סדר)
ג. 2
ד. טווח: 4 שכיח: 2 חציון: 2.5
ה. חציון: 4

שאלה 5:

ב. החציון : 35

השכיח: 35

הממוצע: 35

ג. סטיית תקן : 12.65

ד. 53.333

ה. 25%

ו. -0.395

ז. חציון לא ישתנה, ממוצע יגדל סטיית התקן תקטן.

שאלה 6:

א. 38%

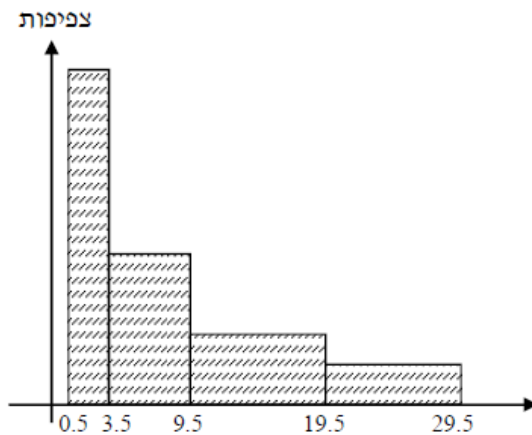
ב. 40%

ג. 150

ד. החציון : 100

שאלה 7:

א.



ב. ההתפלגות היא א-סימטרית ימנית .

ג. שכיח: 2 חציון: 6.83

ד. בהתפלגות א-סימטרית ימנית מתקיים $Mo < Md < \bar{X} < MR$

שאלה 8:

א.

מספר הנבחנים	ציון
50	50-70
75	70-90
125	90-100
125	100-110
75	110-130
50	130-150

ב. 100

ג. 104

ד. הממוצע יעלה ב-10 נקודות אך סטיית התקן לא תשתנה .

שאלה 9:

א. נכון

ב. לא נכון

ג. לא נכון

ד. לא נכון

ה. לא נכון

ו. נכון

ז. לא נכון

ח. לא נכון

פרק 13 - מדדי קשר - מדד הקשר של קרמר

רקע:

מתי משתמשים במדד הזה?
 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 האם הם בעד/נגד/נמנע/נשוא. להלן טבלת השכיחויות המשותפת שהתקבלה.

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

בהקשר של קרמר הטבלה נקראת טבלת O (observed)

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 $E_i = (f_{(x)} * f_{(y)}) / n$

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

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.

תרגילים:

1. להלן תוצאות מחקר שבדק את הקשר בין מין להשכלה. לגבי כל נחקר נבדק המין שלו והשכלתו. להלן התוצאות:

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

האם קיים קשר בין מין להשכלה? נמק!

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. בנה טבלת שכיחות משותפת לנתונים שהוצגו בשאלה.
- ב. האם קיים קשר בין פעילות גופנית למצב בריאותי? חשב לפי מדד הקשר של קרמר.

פתרונות :

שאלה 1

0.595

שאלה 2

ב. 0.19

פרק 14 - מדדי קשר - מדד הקשר של ספירמן

רקע:

מתי נשתמש במדד ספירמן ?

כאשר אחד המשתנים מסולם סדר והשני מסולם סדר ומעלה.

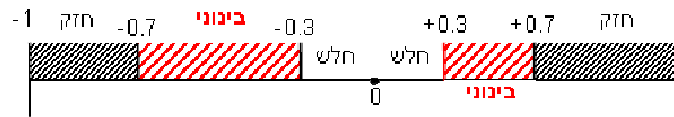
הקשר שהמדד בודק הוא קשר דירוגי.

מדד הקשר בודק :

1 כיוון של הקשר.

2 בודק את עצמת הקשר.

המדד מקבל ערכים בסקלה מ -1 ועד 1.



אם מדד הקשר של ספירמן יוצא 1 המשמעות היא שיש קשר דירוגי חיובי מלא : ככל המשתנה אחד עולה השני עולה ללא יוצא מן הכלל.

קשר דירוגי חיובי חלקי (שמקדם המתאם בין 0 ל-1) אומר שככל שמשתנה אחד עולה לשני יש נטייה לעלות אך לא באופן מוחלט.

אם מדד הקשר של ספירמן יוצא -1 המשמעות היא שיש קשר דירוגי שלילי מלא : ככל שהמשתנה אחד עולה השני יורד ללא יוצא מן הכלל.

קשר דירוגי שלילי חלקי (שמקדם המתאם הוא בין 0 ל- -1) אומר שככל שמשתנה אחד עולה לשני יש נטייה לרדת אך לא באופן מוחלט.

על מנת לחשב את הקשר יש לבצע פעולת דירוג (RANK) נלמד את פעולת הדירוג דרך הדוגמה הבאה (פתרון בהקלטה)

שם התלמיד	הערכה	דירוג R
ערן	בינוני	
מיכל	מצוין	
עודד	חלש	
רוני	טוב	
יעל	טוב	

כאשר מדרגים אם יש כמה תצפיות שתופסות את אותו הערך אז הדירוג שלהם הוא הממוצע של המקומות שהן תופסות.

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

הנוסחה של מדד הקשר :

דוגמה : (פתרון בהקלטה)

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

מהי מידת ההתאמה בין ציוני השופטים?

X- ציון שופט א (סולם סדר)

Y- ציון שופט ב (סולם סדר)

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			

תרגילים:

1. בתחרות יופי חילקו שני שופטים ציונים למועמדות:

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

האם קיים קשר בין שתי הערכות השופטים? נמק והסבר!

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? חשב באמצעות מדד הקשר המתאים והסבר.

3. אם $r_s = 1$ הדבר אומר שערכי X תמיד שווים לערכי Y . האם הטענה נכונה? הסבר.

פתרונות:**שאלה 1:**

0.973

שאלה 2:

-0.85

שאלה 3:

לא נכון

פרק 15 - מדדי קשר - מדד הקשר הלינארי (פירסון)

רקע:

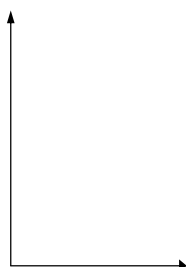
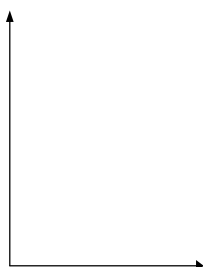
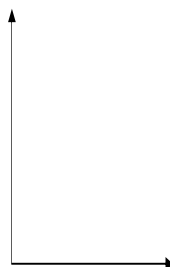
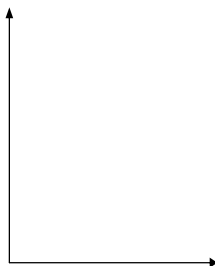
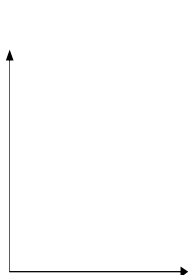
המטרה היא לבדוק האם קיים קשר (קורלציה, מתאם) של קו ישר בין שני משתנים כמותיים. מבחינת סולמות המדידה קשר בין סולמות רווחים ומנה. בדרך כלל, X הוא המשתנה המסביר (הבלתי תלוי) ו Y הוא המשתנה המוסבר (התלוי). למשל, נרצה להסביר כיצד השכלה של אדם הנמדדת בשנות לימוד X מסבירה את ההכנסה שלו Y . במקרה זה שנות ההשכלה 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. בשלב הראשון, נהוג לשרטט דיאגרמת פיזור. זו דיאגרמה שנותנת אינדיקציה ויזואלית על טיב הקשר בין שני המשתנים. למשל, 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 הללו דיאגרמת פיזור:



נתבונן בכמה מקרים של דיאגרמות פיזור וננתח אותן :



בשלב השני, מחשבים את מקדם המתאם (מדד הקשר) שבדוק עד כמה קיים קשר לינארי בין שני המשתנים. הממדד (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

$$y = bx + a$$

מתאם חיובי מלא (מקדם מתאם 1) אומר שקיים קשר לינארי מלא בו השיפוע b יהיה חיובי ואילו

מתאם שלילי מלא אומר שקיים קשר לינארי מלא בו השיפוע b שלילי (מקדם מתאם -1).

מתאם חיובי חלקי אומר שככל שמשנתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת

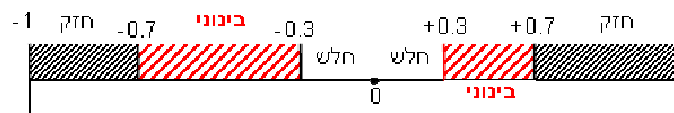
נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט ואילו מתאם שלילי חלקי אומר שככל

שמשנתנה 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.

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

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

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

$$r_{xy} = \frac{COV(x, y)}{s_x \cdot s_y} : \text{מקדם המתאם הלינארי}$$

תרגילים:

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	ציון
2	80
1	90
0	90
2	70
3	70
4	50

- א. שרטט דיאגרמת פיזור לנתונים. מה ניתן להסיק מהדיאגרמה על טיב הקשר בין מספר החיסורים של תלמיד לציונו? מיהו המשתנה הבלתי תלוי ומיהו המשתנה התלוי?
 ב. חשב את מדד הקשר של פירסון. האם התוצאה מתיישבת עם תשובתך לסעיף א?
 ג. הסבר 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:

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

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

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

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

ב. מה המשמעות של התוצאה שקיבלת בסעיף א?

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

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

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

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

חשב 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 קבע נכון/לא נכון ונמק!

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 אמריקאיות:

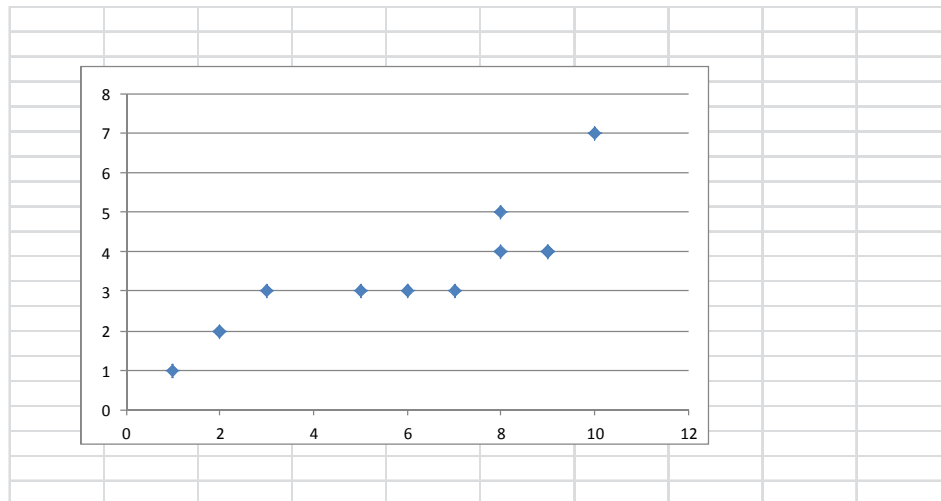
7. נמצא שקיים מקדם מתאם שלילי בין הציון בעברית לציון בחשבון בבחינה לכן :

- 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.

8. נלקחו 20 מוצרים וניבדק 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 דיאגרמת פיזור :



מה יהיה מקדם המתאם בין שני המשתנים?

- א. 1
- ב. 0.85
- ג. 0.15
- ד. 0

פתרונות:**שאלה 1:**

א. בהקלטה

ב. -0.9325 **שאלה 2:**א. $\bar{y} = 16$ ב. $\bar{x} = 15.4$ ב. $r_{xy} = 0.96$ **שאלה 3:**

א : 0.8

שאלה 4:

0.8

שאלה 5:

1

שאלה 6:

א. נכון

ב. לא נכון

ג. נכון

שאלה 7:

התשובה : ג

שאלה 8:

התשובה : א

שאלה 9:

התשובה : ב

פרק 16 - 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 - בעצם נותן את ערך Y כאשר X הנו אפס על גבי קו הניבויים. הוא ניקרא 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:

$$\tilde{Y} = bX + a$$

$$b = r \frac{S_y}{S_x}$$

$$a = \bar{Y} - b\bar{X}$$

אם נרצה לבנות קו ניבויים לניבוי X על סמך Y נצטרך לעדכן את הנוסחאות בהתאם.

תרגילים:

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

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

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

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

- א. חשב את מדד הקשר הלינארי בין X ל- Y . מיהו המשתנה התלוי?
 ב. מצא את קו הרגרסיה לניבוי ההוצאה של משפחה על סמך הכנסה שלה. הסבר 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:

$$S_y = 5 \quad S_x = 2$$

$$\bar{Y} = 8 \quad \bar{X} = 14$$

$$COV(X, Y) = 7.5$$

- א. חשב 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. על פי משוואת הרגרסיה תלמיד שייגש למבחן ללא שעות הכנה כלל יקבל ציון?
 ג. מהו קו הרגרסייה לניבוי הציון לפי שעות ההכנה?
4. נתונים 2 משתנים Y, X . כמו כן נתון: X ממוצע = 1.5, שונות X = שונות Y = 4, וכן שקו הרגרסיה של Y על בסיס X הינו $Y = -0.2X + 0.5$. חשב מהו מקדם המתאם בין X ל- Y ?

פתרונות:**שאלה 1:**

א. 0.8

ב. $\tilde{Y} = 0.8X + 0.4$

ג. 12.4

שאלה 2:

א. 0.75

ב. 4.25 אלפי ש"ח

ג. 14.6 שנים

שאלה 3:

א. 1.2

ב. 29

ג. $y = 1.2x + 29$

שאלה 4:

-0.2

פרק 17 - בעיות בסיסיות בהסתברות

רקע :

ניסוי מקרי : תהליך לו כמה תוצאות אפשריות. התוצאה המתקבלת נודעת רק לאחר ביצוע התהליך.
למשל : תוצאה בהטלת קובייה, מזג האוויר בעוד שבועיים.

מרחב מדגם : כלל התוצאות האפשריות בניסוי המקרי :

בהטלת קובייה : $\{1,2,3,4,5,6\}$.
מזג האוויר בעוד שבועיים : { נאה, שרבי, מושלג, גשום, מעונן חלקית, אביך }

מאורע : תת קבוצה מתוך מרחב במדגם. מסומן באותיות : A, B, C, \dots

בהטלת קובייה, למשל, לקבל לפחות 5 : $A = \{5, 6\}$
לקבל תוצאה זוגית : $B = \{2, 4, 6\}$

גודל מרחב המדגם : מספר התוצאות האפשריות במרחב המדגם :

בהטלת הקובייה : $|\Omega| = 6$

גודל המאורע : מספר התוצאות האפשריות במאורע עצמו.

בהטלת הקובייה : $|A| = 2$ $|B| = 3$

מאורע משלים : מאורע המכיל את כל התוצאות האפשריות במרחב המדגם פרט לתוצאות במאורע אותו הוא משלים :

בהטלת הקובייה : $\bar{A} = \{1, 2, 3, 4\}$ $\bar{B} = \{1, 3, 5\}$

מרחב מדגם אחיד (סימטרי) : מרחב מדגם בו לכל התוצאות במרחב המדגם יש את אותה עדיפות, אותה סבירות למשל, קובייה הוגנת, אך לא כמו מזג האוויר בשבוע הבא.

הסתברות במרחב מדגם אחיד :

במרחב מדגם אחיד הסיכוי למאורע יהיה: $p(A) = \frac{|A|}{|\Omega|}$

למשל, מה הסיכוי בהטלת קובייה לקבל לפחות 5? $p(A) = \frac{|A|}{|\Omega|} = \frac{2}{6}$

מה הסיכוי בהטלת קובייה לקבל תוצאה זוגית? $p(B) = \frac{|B|}{|\Omega|} = \frac{3}{6}$

הסתברות במרחב לא אחיד :

יחושב לפי השכיחות היחסית: $\frac{f}{n}$

להלן התפלגות הציונים בכיתה מסוימת :

הציון X-	מספר התלמידים – השכיחות-f
5	2
6	4
7	8
8	5
9	4
10	2

א. מה ההסתברות שתלמיד אקראי שניבחר בכיתה קיבל את הציון 8? $\frac{f}{n} = \frac{5}{25} = 0.2$

ב. מה ההסתברות שתלמיד אקראי שניבחר בכיתה יכשל?

$$\frac{f}{n} = \frac{2}{25} = 0.08$$

הסתברות למאורע משלים :

$$p(\bar{A}) = 1 - P(A)$$

למשל, בדוגמה הקודמת הסיכוי לעבור את הבחינה יכול להיות מחושב לפי הסיכוי להיכשל :

$$p(A) = 1 - \frac{2}{25} = \frac{23}{25}$$

תרגילים:

1. מהאותיות E, F ו-G יוצרים מילה בת 2 אותיות לא בהכרח בת משמעות.
- הרכב את כל המילים האפשריות.
 - רשום את המקרים למאורע:
 - A במילה נמצאת האות E.
 - B במילה האותיות שונות.
 - ג. רשום את המקרים למאורע \bar{A} .
2. מטיילים זוג קוביות.
- רשום את מרחב המדגם של הניסוי. האם המרחב מדגם הוא אחיד?
 - רשום את כל האפשרויות למאורעות הבאים:
 - A סכום התוצאות 7.
 - C מכפלת התוצאות 12.
 - ג. חשב את הסיכויים למאורעות שהוגדרו בסעיף ב.
3. בוחרים באקראי ספרה מבין הספרות 0-9.
- מה ההסתברות שהספרה שנבחרה גדולה מ-5?
 - מה ההסתברות שהספרה שנבחרה היא לכל היותר 3?
 - ג. מה ההסתברות שהספרה שנבחרה היא אי זוגית?
4. 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

- נבחרה משפחה באקראי מהישוב.
- מה ההסתברות שאין מקלטים למשפחה?
 - ג. מה ההסתברות שיש 3 מקלטים למשפחה?

5. להלן התפלגות מספר המכוניות למשפחה ביישוב "עדן":

מספר מכוניות	מספר משפחות
0	20
1	40
2	100
3	30
4	10

נבחרה משפחה אקראית מן הישוב.

א. מה ההסתברות שאין לה מכוניות?

ב. מה ההסתברות שבבעלות המשפחה לפחות 3 מכוניות?

ג. מה הסיכוי שבבעלותה פחות מ-3 מכוניות?

6. מטילים מטבע רגיל 3 פעמים. בצד אחד של המטבע מוטבע עץ ובצד השני פלי.

א. רשום את מרחב המדגם של הניסוי. האם המרחב מדגם הוא אחיד?

ב. רשום את כל האפשרויות למאורעות הבאים:

A- התקבל פעם אחת עץ.

D- התקבל לפחות פלי אחד.

ג. מהו המאורע המשלים ל-D.

ד. חשבו את הסיכויים למאורעות שהוגדרו בסעיפים ב- ג.

פתרונות:**שאלה 2**

ג. הסיכוי ל-A: $\frac{1}{6}$

הסיכוי ל-B: $\frac{1}{9}$

שאלה 3

א. 0.4

ב. 0.4

ג. 0.5

שאלה 4

א. 0.22

ב. 0.78

ג. 0.32

פרק 18 - פעולות בין מאורעות (חיתוך ואיחוד), מאורעות זרים ומכילים

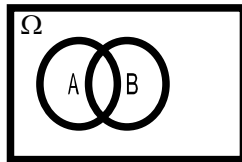
רקע:

פעולת חיתוך:

נותנת את המשותף בין המאורעות הנחתכים, חיתוך בין המאורע A למאורע B יסומן כך:

$$A \cap B$$

מדובר בתוצאות שנמצאות ב-A וגם ב-B.



$A = \{5, 6\}$: בהטלת קובייה, למשל, לקבל לפחות 5

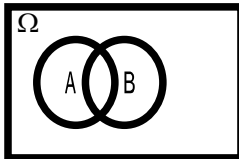
$B = \{2, 4, 6\}$: לקבל תוצאה זוגית

$$A \cap B = \{6\}$$

פעולת איחוד:

נותנת את כל האפשרויות שנמצאות לפחות באחת מהמאורעות. הסימון הוא: $A \cup B$ נותנת את

אשר נימצא ב-A או ב-B. כלומר, לפחות אחד מהמאורעות קורה.



$A = \{5, 6\}$: בהטלת קובייה, למשל, לקבל לפחות 5

$B = \{2, 4, 6\}$: לקבל תוצאה זוגית

$$A \cup B = \{2, 4, 5, 6\}$$

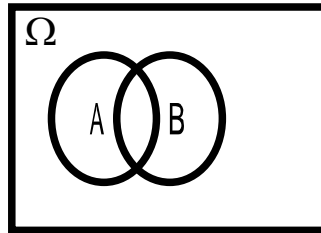
דוגמה (הפתרון נמצא בהקלטה)

סטודנט ניגש בסמסטר לשני מבחנים. מבחן בסטטיסטיקה ומבחן בכלכלה. ההסתברות שלו לעבור את המבחן בסטטיסטיקה הוא 0.9. ההסתברות שלו לעבור את המבחן בכלכלה הוא 0.8. ההסתברות לעבור את המבחן בסטטיסטיקה ובכלכלה היא 0.75.

- א. מה ההסתברות שלו לעבור את המבחן בסטטיסטיקה בלבד?
 ב. מה ההסתברות שלו להיכשל בשני המבחנים?
 ג. מה ההסתברות לעבור לפחות מבחן אחד?

נוסחת החיבור לשני מאורעות :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

חוקי דה מורגן לשני מאורעות:

$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

$$P(A \cap B) = 1 - P(\bar{A} \cup \bar{B})$$

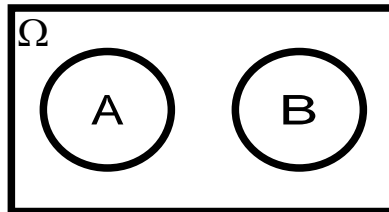
$$P(A \cup B) = 1 - P(\bar{A} \cap \bar{B})$$

שיטת ריבוע הקסם:

השיטה רלבנטית רק אם יש שני מאורעות במקביל בדומה לתרגיל הקודם:

	\bar{A}	A	
B	$P(\bar{A} \cap B)$	$P(A \cap B)$	$P(B)$
\bar{B}	$P(\bar{A} \cap \bar{B})$	$P(A \cap \bar{B})$	$P(\bar{B})$
	$P(\bar{A})$	$P(A)$	1

מאורעות זרים: מאורעות שאין להם מהמשותף: לא יכולים להתרחש בו זמנית.



$$A \cap B = \{ \}$$

$$P(A \cap B) = 0$$

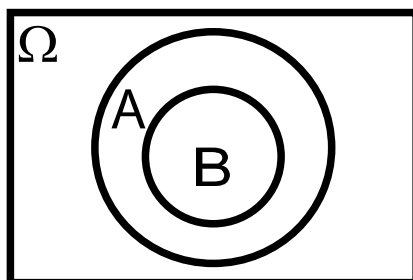
$$P(A \cup B) = P(A) + P(B)$$

למשל, בהטלת קובייה

$$A = \{5, 6\} \quad : \text{לקבל לפחות 5}$$

$$B = \{3\} \quad : \text{לקבל 3}$$

$$A \cap B = \{ \}$$

מאורעות מכילים :

מאורע A מכיל את מאורע B כל התוצאות שנמצאות ב-B מוכלות בתוך המאורע-A.

קשר זה מסומן באופן הבא: $B \subset A$

$$A \cap B = B$$

$$A \cup B = A$$

$$P(A \cap B) = P(B)$$

$$P(A \cup B) = P(A)$$

למשל:

$$A = \{2, 4, 6\}$$

$$B = \{2, 4\}$$

תרגילים:

1. מהאותיות E, F ו-G יוצרים מילה בת 2 אותיות לא בהכרח בת משמעות. נגדיר את המאורעות הבאים :
- E - במילה נמצאת האות E.
 F - במילה אותיות שונות.
 א. רשום את כל האפשרויות לחיתוך A עם B.
 ב. רשום את כל האפשרויות לאיחוד של A עם B.
2. תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ומבחן בסטטיסטיקה. נגדיר את המאורעות הבאים :
- A - לעבור את המבחן בסטטיסטיקה.
 B - לעבור את המבחן בכלכלה.
 העזר בפעולות חיתוך, איחוד ומשלים בלבד כדי להגדיר את המאורעות הבאים וסמן בדיאגרמת וון את השטח המתאים :
- א. התלמיד עבר רק את המבחן בכלכלה.
 ב. התלמיד עבר רק את המבחן בסטטיסטיקה.
 ג. התלמיד עבר את שני המבחנים.
 ד. התלמיד עבר לפחות מבחן אחד.
 ה. התלמיד נכשל בשני המבחנים.
 ו. התלמיד נכשל בכלכלה.
3. נתבקשתם לבחור ספרה באקראי. נגדיר את A להיות הספרה שנבחרה היא זוגית. נגדיר את B להיות הספרה שנבחרה קטנה מ-5.
- א. רשמו את כל התוצאות למאורעות הבאים :
- $A =$
 $B =$
 $\bar{B} =$
 $A \cap B =$
 $A \cup B =$
- ב. חשבו את ההסתברויות לכל המאורעות מהסעיף הקודם.

4. נסמן ב- Ω את מרחב המדגם וב- ϕ קבוצה ריקה.

נתון כי A הינו מאורע בתוך מרחב המדגם.

להלן מוגדרים מאורעות שפתרונם הוא Ω או ϕ או A .

קבע עבור כל מאורע מה הפתרון שלו.

$$\begin{aligned}
 &= \\
 &A \\
 &A \cap \phi \\
 &A \cup \phi \\
 &A \cap \Omega \\
 &A \cup \Omega \\
 &A \cap \bar{A} \\
 &\bar{\phi} \\
 &A \cup \bar{A}
 \end{aligned}$$

5. הוגדרו המאורעות הבאים:

A = אדם שגובהו מעל 1.7 מטר

B = אדם גובהו מתחת ל-1.8 מטר

קבע את גובהם של האנשים הבאים:

א. $A \cap B$

ב. $A \cup B$

ג. $\bar{A} \cap B$

ד. $\bar{A} \cup \bar{B}$

ה. \bar{A}

6. נגדיר את המאורעות הבאים :
- A - אדם דובר עברית.
 B - אדם דובר ערבית.
 C - אדם דובר אנגלית.
- השתמש בפעולות איחוד, חיתוך והשלמה לתיאור המאורעות הבאים :
- א. אדם דובר את כל שלוש השפות.
 ב. אדם דובר רק עברית.
 ג. אדם דובר לפחות שפה אחת מתוך השפות הללו.
 ד. אדם אינו דובר אנגלית.
 ה. קבוצת התלמידים דוברי 2 שפות בדיוק (מהשפות הנ"ל).
7. שתי מפלגות רצות לכנסת הבאה. מפלגת "גדר" תעבור את אחוז החסימה בהסתברות של 0.08. מפלגת עתיד תעבור את אחוז החסימה בהסתברות של 0.20. בהסתברות של 76% שתי המפלגות לא תעבורנה את אחוז החסימה.
- א. מה ההסתברות שלפחות אחת מהמפלגות תעבור את אחוז החסימה?
 ב. מה ההסתברות ששתי המפלגות תעבורנה את אחוז החסימה?
 ג. מה ההסתברות שרק מפלגות "עתיד" תעבור את אחוז החסימה?
8. במקום עבודה מסוים 40% מהעובדים הם גברים. כמו כן 20% מהעובדים הם אקדמאים. 10% מהעובדים הינן נשים אקדמאיות.
- א. איזה אחוז מהעובדים הם גברים אקדמאיים?
 ב. איזה אחוז מהעובדים הם גברים או אקדמאיים?
 ג. איזה אחוז מהעובדים הם נשים לא אקדמאיות?
9. הסיכוי של מניה A לעלות הנו 0.5 ביום מסוים והסיכוי של מניה B לעלות ביום מסוים הנו 0.4. בסיכוי של 0.7 לפחות אחת מהמניות תעלה ביום מסוים. חשב את ההסתברויות הבאות לגבי שתי המניות הללו ביום מסוים :
- א. ששתי המניות תעלנה.
 ב. שאף אחת מהמניות לא תעלנה.
 ג. שמניה A בלבד תעלה.

10. מטילים זוג קוביות אדומה ושחורה. נגדיר את המאורעות הבאים:

A - בקובייה האדומה התקבלה התוצאה 4 ובשחורה 2.

B - סכום התוצאות משתי הקוביות 6.

C - מכפלת התוצאות בשתי הקוביות 10.

א. האם A ו-B מאורעות זרים?

ב. האם המאורע B מכיל את המאורע A?

ג. האם A ו-C מאורעות זרים?

ד. האם A ו-C מאורעות משלימים?

11. עבור המאורעות A ו-B ידועות ההסתברויות הבאות:

$$p(\bar{A} \cap \bar{B}) = 0.1 \quad p(B) = 0.3 \quad p(A) = 0.6$$

א. האם A ו-B מאורעות זרים?

ב. חשב את $p(\bar{A} \cap B)$

12. מטבע הוטל פעמיים. נגדיר את המאורעות הבאים:

A - קיבלנו עץ בהטלה הראשונה.

B - קיבלנו לפחות עץ אחד בשתי ההטלות.

איזו טענה נכונה?

א. A ו-B מאורעות זרים.

ב. A ו-B מאורעות משלימים.

ג. B מכיל את A.

ד. A מכיל את B.

13. בהגרלה חולקו 100 כרטיסים על 3 מהם רשום חופשה ועל 2 מהם רשום מחשב שאר

הכרטיסים ריקים. אדם קיבל כרטיס אקראי.

א. מה הסיכוי לזכות בחופשה או במחשב? האם המאורעות הללו זרים?

ב. מה ההסתברות לא לזכות בפרס?

$$P(A) = 0.3$$

$$P(B) = 0.25$$

$$P(A \cup B) = 0.49$$

א. חשב את הסיכוי ל- $P(A \cap B)$.

ב. האם A ו- B מאורעות זרים?

ג. מה ההסתברות שרק A יקרה או רק B יקרה?

15. A ו- B מאורעות זרים. נתון ש: $2 \cdot P(B \cap \bar{A}) = P(A \cap \bar{B}) = P(\bar{A} \cap \bar{B})$.

מה הסיכוי למאורע A ומה ההסתברות למאורע B?

פתרונות:**שאלה 7**

א. 0.24

ב. 0.04

ג. 0.16

שאלה 8

א. 10%

ב. 50%

ג. 50%

שאלה 9

א. 0.2

ב. 0.3

ג. 0.3

שאלה 10

א. לא.

ב. כן.

ג. כן.

ד. לא.

שאלה 11

א. כן

ב. 0.3

שאלה 12

התשובה הנכונה ג

שאלה 13

א. 0.05

ב. 0.95

שאלה 14

א. 0.06

ב. לא זרים

ג. 0.43

פרק 19 - קומבינטוריקה - כלל המכפלה

רקע:

כלל המכפלה:

כלל המכפלה הוא כלל שבאמצעותו אפשר לחשב את גודל המאורע או גודלו של מרחב המדגם. אם לתהליך יש k שלבים : n_1 אפשרויות לשלב הראשון, n_2 אפשרויות לשלב השני ... n_k אפשרויות לשלב k :

מספר האפשרויות לתהליך כולו יהיה : $n_1 \cdot n_2 \cdot n_3 \cdots n_k$

למשל, כמה אפשרויות יש למשחק בו מטיילים קובייה וגם מטבע? (הסבר בהקלטה)

למשל, כמה לוחיות רישוי בני 5 תווים ניתן ליצור כאשר התו הראשון הוא אות אנגלי והיתר ספרות? (הסבר בהקלטה)

תרגילים:

1. חשבו את מספר האפשרויות לתהליכים הבאים:
 - א. הטלת קובייה פעמים.
 - ב. מספר תלת ספרתי.
 - ג. בחירת בן ובת מכתה שיש בה שבעה בנים ועשר בנות.
 - ד. חלוקת שני פרסים שונים לעשרה אנשים שונים כאשר אדם לא יכול לקבל יותר מפרס אחד.

2. במסעדה מציעים ארוחה עסקית. בארוחה עסקית יש לבחור מנה ראשונה, מנה עיקרית ושתייה. האופציות למנה ראשונה הן: סלט ירקות, סלט אנטיפסטי ומרק היום. האופציות למנה עיקרית הן: סטייק אנטרקוט, חזה עוף בגריל, לזניה בשרית ולזניה צמחונית. האופציות לשתייה הן: קפה, תה ולימונדה.
 - א. כמה ארוחות שונות ניתן להרכיב בעזרת התפריט הזה?
 - ב. אדם מזמין ארוחה אקראית. חשב את ההסתברויות הבאות:
 1. בארוחה סלט ירקות, לזניה בשרית ולימונדה.
 2. בארוחה סלט, לזניה ותה.

3. בוחרים באקראי מספר בין חמש ספרות. חשבו את ההסתברויות הבאות:
 - א. המספר הוא זוגי.
 - ב. במספר כל הספרות שונות.
 - ג. במספר כל הספרות זהות.
 - ד. במספר לפחות שתי ספרות שונות.
 - ה. במספר לפחות שתי ספרות זהות.
 - ו. המספר הוא פלינדרום (מספר הנקרא מימין ומשמאל באותה צורה).

4. חמישה אנשים אקראיים נכנסו למעלית בבנין בן 8 קומות. חשבו את ההסתברויות הבאות:
 - א. כולם ירדו בקומה החמישית?
 - ב. כולם ירדו באותה קומה?
 - ג. כולם ירדו בקומה אחרת?
 - ד. ערך ודני ירדו בקומה השישית והיתר בשאר הקומות?

5. במפלגה חמישה עשר חברי כנסת. יש לבחור שלושה חברי כנסת לשלושה תפקידים שונים. בכמה דרכים ניתן לחלק את התפקידים אם:
- חבר כנסת יכול למלא יותר מתפקיד אחד.
 - חבר כנסת לא יכול למלא יותר מתפקיד אחד.
6. מטילים קובייה 4 פעמים.
- מה ההסתברות שכל התוצאות תהינה זהות?
 - מה ההסתברות של התוצאות תהינה שונות?
 - מה ההסתברות שלפחות שתי תוצאות תהינה זהות?
 - מה ההסתברות שלפחות שתי תוצאות תהינה שונות?
7. יש ליצור מילה בת חמש אותיות לא בהכרח עם משמעות מאותיות ה-ABC (26 אותיות) בת 5 אותיות.
- מה ההסתברות שבמילה שנוצרה אין האותיות L, A, D?
 - מה ההסתברות שבמילה שנוצרה כל האותיות זהות?
 - מה ההסתברות שבמילה שנוצרה לפחות שתי אותיות שונות זו מזו?
 - מה ההסתברות שהמילה היא פלינדרום (מילה אשר משמאל לימין, ומימין לשמאל נקראת אותו הדבר).
8. יוצרים קוד עם a ספרות (מותר לחזור על אותה ספרה בקוד). חשבו את ההסתברויות הבאות: (בטאו את תשובתיכם באמצעות a)
- בקוד אין את הספרה 5.
 - בקוד מופיעה הספרה 3.
 - בקוד לא מופיעות ספרות אי זוגיות.
9. במשחק מזל יש למלא טופס בו n משבצות. כל משבצת מסומנת בסימון V או בסימון X. בכמה דרכים שונות ניתן למלא את טופס משחק המזל?

פתרונות :**שאלה 2**

- א. 36
 ב. $1/36$
 ג. $1/9$

שאלה 4

- א. 0.00003
 ב. 0.00024
 ג. 0.20508
 ד. 0.01047

שאלה 6

- א. $1/216$
 ב. $5/18$
 ג. $13/18$
 ד. $215/216$

שאלה 9

2"

שאלה 1

- א. 36
 ב. 900
 ג. 70
 ד. 90

שאלה 3

- א. 0.5
 ב. 0.3024
 ג. 0.0001
 ד. 0.9999
 ה. 0.6976
 ו. 0.01

שאלה 5

- א. 3,375
 ב. 2,730

שאלה 7

- א. 0.5417
 ב. $\frac{1}{26^4}$
 ד. 0.0015

פרק 20 - קומבינטוריקה - תמורה - סידור עצמים בשורה

רקע:

תמורה:

מספר האפשרויות לסדר n עצמים שונים בשורה:

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$$

$$0! = 1$$

למשל, בכמה דרכים שונות ניתן לסדר את האותיות a, b, c, d ? (הפתרון בהקלטה)

למשל, בכמה דרכים שונות ניתן לסדר את האותיות a, b, c, d , כך שהאותיות a, b יהיו ברצף?

(הפתרון בהקלטה)

למשל, בכמה דרכים שונות ניתן לסדר את האותיות a, b, c, d , כך שהאותיות a, b יופיעו בתור

הרצף ba ? (הפתרון בהקלטה)

תרגילים:

1. חשבו בכמה אופנים :
 - א. אפשר לסדר 4 ספרים שונים על מדף?
 - ב. אפשר לסדר חמישה חיילים בטור?

2. סידרו באקראי 10 דיסקים שונים על מדף שמתוכם שניים בשפה העברית.
 - א. מה ההסתברות שהדיסקים בעברית יהיו צמודים זה לזה?
 - ב. מה ההסתברות שהדיסקים בעברית לא יהיו צמודים זה לזה?
 - ג. מה ההסתברות ששני הדיסקים בעברית יהיו כל אחד בקצה השני של המדף?

3. בוחנים 5 בנים ו-4 בנות בכיתה ומדרגים אותם לפי הציון שלהם בבחינה. נניח שאין תלמידים להם אותו ציון.
 - א. מהו מספר הדירוגים האפשריים?
 - ב. מהו מספר הדירוגים האפשריים, אם מדרגים בנים ובנות בנפרד?

4. מסדרים 10 ספרים שונים על מדף.
 - א. בכמה אופנים ניתן לסדר את הספרים על המדף?

- שני ספרים מתוך ה-10 הם ספרים בסטטיסטיקה.
 - ב. מה ההסתברות שאם נסדר את הספרים באקראי, הספרים בסטטיסטיקה יהיו צמודים זה לזה?
 - ג. מה ההסתברות שהספרים בסטטיסטיקה לא יהיו צמודים זה לזה?
 - ד. מה ההסתברות שהספרים בסטטיסטיקה יהיו בקצות המדף (כל ספר בקצה אחר)?

5. אדם יצר בנגן שלו פלייליסט (רשימת השמעה) של 12 שירים שונים. 4 בשפה העברית, 5 באנגלית ו-3 בצרפתית. האדם הריץ את הפלייליסט באקראי.
 - א. מה ההסתברות שכל השירים באנגלית יופיעו כשירים הראשונים כמקשה אחת?
 - ב. מה ההסתברות שכל השירים באנגלית יופיעו ברצף (לא חובה ראשונים)?
 - ג. מה ההסתברות ששירים באותה השפה יופיעו ברצף (כלומר כל השירים באנגלית ברצף, כל השירים בעברית ברצף וכך גם השירים בצרפתית)?

6. 4 בנינים ו-4 בנות התיישבו באקראי בשורת קולנוע בכיסאות 1-8.
- א. מה ההסתברות שיוסי ומיכל לא ישבו זה לצד זה?
- ב. מה ההסתברות שהבנים יתיישבו במקומות האי-זוגיים?
- ג. מה ההסתברות שכל הבנים ישבו זה לצד זה?
- ד. מה ההסתברות שהבנים ישבו זה לצד זה והבנות תשבנה זו לצד זו?

פתרונות :שאלה 1

א. 24

ב. 120

שאלה 2

א. 0.2

ב. 0.8

ג. 0.022

שאלה 3

א. 362,880

ב. 2,880

שאלה 4

א. 3,628,800

ב. 0.2

ג. 0.8

ד. $\frac{1}{45}$ שאלה 5א. $\frac{1}{792}$ ב. $\frac{1}{99}$ ג. $\frac{1}{4620}$ שאלה 6

א. 0.75

ב. 0.014

ג. $\frac{1}{14}$ ד. $\frac{1}{35}$

פרק 21 - קומבינטוריקה - דגימה סידורית ללא החזרה ועם החזרה

רקע:

מדגם סדור בדגימה עם החזרה

מספר האפשרויות בדגימת k עצמים מתוך n עצמים שונים כאשר הדגימה היא עם החזרה והמדגם סדור הוא: n^k .

למשל,
בחרים שלושה תלמידים מתוך עשרה לייצג ועד בו תפקידים שונים, תלמיד יכול למלא יותר מתפקיד אחד.
כמה ועדים שונים ניתן להרכיב?

$$n = 10$$

$$k = 3$$

$$10^3 = 1,000$$

מדגם סדור ללא החזרה

מספר האפשרויות בדגימת k עצמים שונים מתוך n עצמים שונים ($n \geq k$) כאשר המדגם סדור ואין החזרה של עצמים נדגמים הינו:

$$(n)_k = n(n-1)(n-2)\cdots(n-(k-1)) = \frac{n!}{(n-k)!}$$

למשל,
שלושה תלמידים נבחרים מתוך 10 לייצג וועד בו תפקידים שונים. תלמיד לא יכול למלא יותר מתפקיד אחד.

$$\frac{10!}{7!} = 10 \cdot 9 \cdot 8 = 720$$

תרגילים:

1. במפלגה 20 חברי כנסת, מעוניינים לבחור שלושה חברי כנסת לשלושה תפקידים שונים.
 - א. חבר כנסת יכול למלא יותר מתפקיד אחד. כמה קומבינציות ישנן לחלוקת התפקידים?
 - ב. חבר כנסת לא יכול למלא יותר מתפקיד אחד. כמה קומבינציות יש לחלוקת התפקידים?

2. במשחק מזל יש 4 משבצות ממוספרות מ A-D (A עד D). בכל משבצת יש למלא סיפרה (0-9).
 - א. מה ההסתברות לזכות במשחק?
 - ב. מה ההסתברות שבאף משבצת לא תהיה את הספרה 3 במספר הזוכה?
 - ג. מה ההסתברות שהתוצאה 4 תופיע לפחות פעם אחת במספר הזוכה?

3. קבוצה מונה 22 אנשים, מה ההסתברות שלפחות לשניים מהם יהיה יום הולדת באותו התאריך?

4. שלושה אנשים קבעו להיפגש במלון הילטון בסינגפור. הבעיה היא שבסינגפור ישנם 5 מלונות הילטון.
 - א. מה ההסתברות שכל השלושה ייפגשו?
 - ב. מה ההסתברות שכל אחד יגיע לבית מלון אחר?

5. בכיתה 40 תלמידים. מעוניינים לבחור חמישה מהם לוועד כיתה. בכמה דרכים ניתן להרכיב את הוועד אם:
 - א. בוועד 5 תפקידים שונים ותלמיד יכול למלא יותר מתפקיד אחד.
 - ב. בוועד 5 תפקידים שונים ותלמיד לא יכול למלא יותר מתפקיד אחד.

פתרונות :**שאלה 1 :**

א. 8000

ב. 6840

שאלה 2 :

א. 0.0001

ב. 0.6561

ג. 0.3439

שאלה 3 :

0.476

שאלה 4 :

א. 0.04

ב. 0.48

פרק 22 - קומבינטוריקה - דגימה ללא סדר וללא החזרה

רקע:

מדגם לא סדור בדגימה ללא החזרה

מספר האפשרויות לדגום k עצמים שונים מתוך n עצמים שונים כאשר אין משמעות לסדר העצמים הנדגמים ואין החזרה:

$$\frac{n!}{(n-k)!k!} = \binom{n}{k} = \frac{(n)_k}{k!}$$

דוגמה

מתוך 10 תלמידים יש לבחור שלושה נציגים לוועד ללא תפקידים מוגדרים:

$$\binom{10}{3} = \frac{10!}{7!3!} = 120$$

הערות

$$\binom{n}{k} = \binom{n}{n-k} \quad .1$$

$$\binom{n}{n-1} = \binom{n}{1} = n \quad .2$$

$$\binom{n}{n} = \binom{n}{0} = 1 \quad .3$$

תרגילים :

1. בכיתה 15 בנות ו-10 בנים. יש לבחור 5 תלמידים שונים מהכיתה לנציגות הכיתה. בכמה דרכים אפשר להרכיב את הנציגות אם-
 - א. אין שום הגבלה לבחירה.
 - ב. מעוניינים ש-3 בנות ו-2 בנים ירכיבו את המשלחת.
 - ג. לא יהיו בנים במשלחת.

2. סטודנט מעוניין לבחור 5 קורסי בחירה בסמסטר זה. לפניו רשימה של 10 קורסים לבחירה:
 - 5 במקצועות מדעי הרוח.
 - 3 במקצועות מדעי החברה.
 - 2 מתחום המתמטיקה.
 - א. כמה בחירות שונות הוא יכול ליצור לעצמו?
 - ב. כמה בחירות יש לו בהן 3 קורסים הם ממדעי הרוח?
 - ג. כמה בחירות יש לו אם 2 מהן לא ממדעי הרוח?
 - ד. כמה בחירות יש לו אם 2 ממדעי הרוח, 2 ממדעי החברה ו-1 ממתמטיקה?

3. בכיתה 30 תלמידים מתוכם 12 תלמידים ו-18 תלמידות. יש לבחור למשלחת 4 תלמידים מהכיתה. התלמידים נבחרים באקראי.
 - א. מה ההסתברות שהמשלחת תורכב רק מבנות?
 - ב. מה ההסתברות שבמשלחת תהיה רק בת אחת?
 - ג. מה ההסתברות שבמשלחת תהיה לפחות בת אחת?

4. במשחק הלוטו יש לבחור 5 מספרים מתוך 45. המספרים הם 1-45.
 - א. מה ההסתברות שבמשחק הזוכה כל המספרים הם זוגיים?
 - ב. מה ההסתברות שבמספר הזוכה יש לכל היותר מספר זוגי אחד?
 - ג. מה ההסתברות שבמספר הזוכה לפחות פעם אחת יש מספר זוגי?
 - ד. מה ההסתברות שבמספר הזוכה כל המספרים גדולים מ-30?

5. בחפיסת קלפים ישנם 52 קלפים: 13 בצבע שחור בצורת עלה, 13 בצבע אדום בצורת לב, 13 בצבע אדום בצורת יהלום ו- 13 בצבע שחור בצורת תלתן. מכל צורה (מתוך ה-4) יש 9 קלפים שמספרם 2-10, שאר הקלפים הם; נסיך, מלכה, מלך ואס (בעצם מדובר בקופסת קלפים רגילה ללא ג'וקר). שני אנשים משחקים פוקר. כל אחד מקבל באקראי 5 קלפים (ללא החזרה).

- א. מה ההסתברות שעודד יקבל את כל המלכים וערן את כל המלכות?
- ב. מה ההסתברות שאחד השחקנים יקבל את הקלף אס-לב?
- ג. מה ההסתברות שערן יקבל קלפים שחורים בלבד ועודד יקבל שני קלפים שחורים בדיוק?
- ד. מה ההסתברות שערן יקבל לפחות 3 קלפים שהם מספר (אס אינו מספר)?

6. במכללה 4 מסלולי לימוד. בכל מסלול לימוד 5 מזכירות. יש ליצור וועד של 5 מזכירות מתוך כלל המזכירות במכללה. יוצרים וועד באופן אקראי. חשבו את ההסתברויות הבאות:

- א. כל המזכירות בוועד יהיו ממסלול "מדעי ההתנהגות".
- ב. כל המזכירות בוועד יהיו מאותו המסלול.
- ג. מכל מסלול תבחר לפחות מזכירה אחת.

פתרונות:**שאלה 2**

א. 252

ב. 100

ג. 100

ד. 60

שאלה 1

א. 53,130

ב. 20,475

ג. 3003

שאלה 4

א. 0.02

ב. 0.187

ג. 0.972

ד. 0.00246

שאלה 3

א. 0.1117

ב. 0.1445

ג. 0.9819

שאלה 6א. $6.45 \cdot 10^{-5}$ ב. $2.58 \cdot 10^{-4}$

ג. 0.3225

פרק 23 - הסתברות מותנית - במרחב מדגם אחיד

רקע:

לעיתים אנו נדרשים לחשב הסתברות למאורע כלשהו כאשר ברשותנו אינפורמציה לגבי מאורע אחר. הסתברות מותנית הינה סיכוי להתרחשות מאורע כלשהו אשר ידוע שמאורע אחר התרחש/ לא התרחש.

ההסתברות של A בהינתן ש B כבר קרה:

$$P(A|B)$$

$$P(A|B) = \frac{|A \cap B|}{|B|} \quad \text{כשמרחב המדגם אחיד:}$$

למשל, (פתרון בהקלטה)

מטילים קובייה.

נגדיר :

A – התוצאה זוגית.

B – התוצאה גדולה מ-3 .

נרצה לחשב את :

$$P(A|B)$$

תרגילים:

1. נבחרה ספרה זוגית באקראי. מה הסיכוי שהספרה גדולה מ-6?
2. יוסי הטיל קובייה. מה הסיכוי שקיבל את התוצאה 4 אם ידוע שהתוצאה שהתקבלה זוגית?
3. מטילים צמד קוביות.
נגדיר:
A – סכום התוצאות בשתי ההטלות הינו 7
B – מכפלת התוצאות 12
חשבו את $P(A|B)$
4. הוטל מטבע פעמיים. ידוע שהתקבל לכל היותר ראש אחד, מה הסיכוי שהתקבלו שני ראשים?
5. אדם הטיל זוג קוביות והתקבל שהתוצאות זהות. מה הסיכוי שלפחות אחת התוצאות 5?
6. אדם הטיל זוג קוביות והתקבל לפחות פעם אחת. מה הסיכוי שאחת התוצאות 5?
7. נבחרה משפחה בת שני ילדים. ידוע שאחד הילדים בן. מה ההסתברות שבמשפחה שני בנים בקרב הילדים?
8. נבחרה משפחה בת שלושה ילדים. נתון שהילד האמצעי בן. מה הסיכוי שיש בנות בקרב הילדים?

השאלות הבאות משלבות קומבינטוריקה:

9. בכיתה 6 בנים ו-7 בנות. נבחרו ארבעה ילדים מהכיתה.
אם ידוע שנבחרו 2 בנים ושתי בנות, מה הסיכוי שאלעד לא נבחר?
10. חמישה חברים יצאו לבית קולנוע והתיישבו זה ליד זה באקראי בכיסאות מספר 5 עד 9.
אם ידוע שערך ודין התיישבו זה ליד זה. מה ההסתברות שהם יושבים בכיסאות מספר 6 ו 7?

פתרונות:**שאלה 1**

0.2

שאלה 2 $1/3$ **שאלה 3**

0.5

שאלה 4

0

שאלה 5 $1/6$ **שאלה 6** $2/11$ **שאלה 7** $1/3$ **שאלה 8** $3/4$ **שאלה 9** $2/3$ **שאלה 10** $1/4$

פרק 24 - הסתברות מותנית - מרחב לא אחיד

רקע:

הסיכוי שמאורע A יתרחש בהינתן ש – מאורע B כבר קרה :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

במונה : הסיכוי לחיתוך של שני המאורעות זה הנשאל וזה הנתון שהתרחש.

במכנה : הסיכוי למאורע שנתון שהתרחש :

למשל,

נלקחו משפחות שיש להם שתי מכוניות. ל- 30% מהמשפחות הללו המכונית הישנה יותר היא מתוצרת אירופה ואצל 60% מהמשפחות הללו המכונית החדשה יותר מתוצרת אירופה. כמו כן 15% מהמשפחות הללו שתי המכוניות הן מתוצרת אירופאית.
אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית? (פתרון בהקלטה)

תרגילים:

1. תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ומבחן בסטטיסטיקה :
 נגדיר את המאורעות הבאים : A- לעבור את המבחן בסטטיסטיקה. B- לעבור את המבחן בכלכלה.
 כמו כן נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8 והסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9. הסיכוי לעבור את שני המבחנים הנו 0.75. חשבו את הסיכויים למאורעות הבאים :
- התלמיד עבר בסטטיסטיקה, מה ההסתברות שהוא עבר בכלכלה?
 - התלמיד עבר בכלכלה, מה ההסתברות שהוא עבר בסטטיסטיקה?
 - התלמיד עבר בכלכלה, מה ההסתברות שהוא נכשל בסטטיסטיקה?
 - התלמיד נכשל בסטטיסטיקה מה ההסתברות שהוא נכשל בכלכלה?
 - התלמיד עבר לפחות מבחן אחד מה ההסתברות שהוא יעבור את שני המבחנים?
2. במדינה שתי חברות טלפון סלולארי "סופט" ו"בל". 30% מהתושבים הבוגרים רשומים אצל חברת "בל". 60% מהתושבים הבוגרים רשומים אצל חברת "סופט".
 ל-15% מהתושבים הבוגרים אין טלפון סלולארי בכלל.
- איזה אחוז מהתושבים הבוגרים רשומים אצל שתי החברות?
 - נבחר אדם שרשום אצל חברת "סופט", מה ההסתברות שהוא רשום גם אצל חברת "בל"?
 - אם אדם לא רשום אצל חברת "בל", מה ההסתברות שהוא כן רשום בחברת "סופט"?
 - אם אדם רשום אצל חברה אחת בלבד, מה ההסתברות שהוא רשום בחברת "סופט"?
3. במכללה שני חניונים : חניון קטן וחניון גדול. בשעה 08:00 יש סיכוי של 60% שבחניון הגדול יש מקום, סיכוי של 30% שבחניון הקטן יש מקום וסיכוי של 20% שבשני החניונים יש מקום.
- מה ההסתברות שיש מקום בשעה 08:00 רק בחניון הגדול של המכללה?
 - ידוע שבחניון הקטן יש מקום בשעה 08:00, מה הסיכוי שבחניון הגדול יש מקום?
 - אם בשעה 08:00 בחניון הגדול אין מקום, מה ההסתברות שבחניון הקטן יהיה מקום?
 - נתון שלפחות באחד מהחניונים יש מקום בשעה 08:00, מה ההסתברות שבחניון הגדול יש מקום?

פתרונות:**שאלה 1**

- א. 0.833
- ב. 0.9375
- ג. 0.0625
- ד. 0.5
- ה. 0.789

שאלה 2

- א. 5%
- ב. 0.0833
- ג. 0.786
- ד. 0.6875

שאלה 3

- א. 0.4
- ב. $\frac{2}{3}$
- ג. 0.25
- ד. $\frac{6}{7}$

פרק 25 - תלות ואי תלות בין מאורעות

רקע:

אם מתקיים ש: $P(B|A) = p(B)$ נגיד שמאורע B בלתי תלוי ב-A.

הדבר גורר גם ההפך: $P(A|B) = p(A)$ כלומר A אינו תלוי גם ב-B.

כשהמאורעות בלתי תלויים מתקיים ש: $P(A \cap B) = P(A) \cdot P(B)$.

הוכחה לכך:

$$P(A|B) = P(A) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A) \cdot P(B)$$

נשתמש בנוסחאות של מאורעות בלתי תלויים רק אם נאמר במפורש שהמאורעות בלתי תלויים

בתרגיל או שמהקשר אפשר להבין ללא צל של ספק שהמאורעות בלתי תלויים.

למשל,

חוקר מבצע שני ניסויים בלתי תלויים הסיכוי להצליח בניסוי הראשון הנו 0.7 והסיכוי להצליח

בניסוי השני הוא 0.4.

א. מה הסיכוי להצליח בשני הניסויים יחדו?

כיוון שהמאורעות הללו בלתי תלויים:

$$p(A \cap B) = P(A) \cdot P(B) = 0.7 \cdot 0.4 = 0.28$$

ב. מה הסיכוי להיכשל בשני הניסויים?

$$p(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B}) = (1-0.7)(1-0.4) = 0.18 \quad \text{באופן דומה:}$$

הרחבה: אי תלות בין n מאורעות

n מאורעות A_1, \dots, A_n הם בלתי תלויים אם ורק אם:

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i)$$

תרגילים:

1. נתון:

$$p(A) = 0.2$$

$$P(B) = 0.5$$

$$P(A \cup B) = 0.6$$

האם המאורעות הללו בלתי תלויים?

2. תלמיד ניגש לשני מבחנים שהצלחתם לא תלויה זו בזו. הסיכוי שלו להצליח במבחן הראשון

הוא 0.7 והשני 0.4 .

א. מה הסיכוי להצליח בשני המבחנים יחדו?

ב. מה הסיכוי שניכשל בשני המבחנים ?

3. במדינה מסוימת 8% אבטלה, נבחרו באקראי שני אנשים מהמדינה.

א. מה ההסתברות ששניהם מובטלים?

ב. מה ההסתברות שלפחות אחד מהם מובטל?

4. מוצר צריך לעבור בהצלחה ארבע בדיקות בלתי תלויות לפני שיווקו, אחרת הוא נפסל ולא

יוצא לשוק. הסיכוי לעבור בהצלחה כל אחת מהבדיקות הוא 0.8. בכל מקרה מבוצעות כל 4

הבדיקות.

א. מה הסיכוי שהמוצר יפסל?

ב. מה ההסתברות שהמוצר יעבור בהצלחה לפחות בדיקה אחת?

5. מדינה מסוימת 8% אבטלה, נבחרו באקראי חמישה אנשים מהמדינה.

א. מה ההסתברות שכולם מובטלים במדגם?

ב. מה ההסתברות שלפחות אחד מהם מובטל?

6. עבור שני מאורעות A ו-B המוגדרים על אותו מרחב מדגם נתון ש: $P(A \cup B) = 0.9$,
 $P(A|B) = 0.6$, $P(A \cap \bar{B}) = 0.3$. האם A ו-B מאורעות בלתי תלויים?

פתרונות :

שאלה 1

כן

שאלה 2

א. 0.28

ב. 0.18

שאלה 3

א. 0.0064

ב. 0.1536

שאלה 4

א. 0.5904

ב. 0.9984

פרק 26 - התפלגויות בדידות מיוחדות - התפלגות בינומית

רקע:

נגדיר את המושג ניסוי ברנולי: ניסוי ברנולי הנו ניסוי שיש לו שתי תוצאות אפשריות: "הצלחה" ו"כישלון" כמו: מוצר פגום או תקין אדם עובד או מובטל עץ או פלי בהטלת מטבע וכדומה.

בהתפלגות בינומית חוזרים על אותו ניסוי ברנולי n פעמים באופן בלתי תלוי זה בזה. מגדירים את X להיות מספר ההצלחות שהתקבלו בסך הכול. נסמן ב p את הסיכוי להצלחה בניסוי בודד וב q את הסיכוי לכישלון בניסוי בודד.

ואז נגיד ש: $X \sim B(n, p)$.

פונקציית ההסתברות של X :

$$k = 0, 1, 2, \dots, n; P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{לכל}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}; \quad n! = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 1; \quad 0! = 1; \quad \text{כאשר}$$

לגודל $\binom{n}{k}$: ניתן לחשב באמצעות המחשבונית.

$$E(X) = np \quad \text{תוחלת}$$

$$V(X) = npq \quad \text{שונות}$$

שימו לב כדי לזהות שמדובר בהתפלגות בינומית צריכים להתקיים כל התנאים הבאים:

- (1) חוזרים על אותו ניסוי ברנולי באופן בלתי תלוי זה בזה.
- (2) חוזרים על הניסוי n פעמים.
- (3) X – מוגדר כמספר ההצלחות המתקבלות בסך הכול.

דוגמה: (פתרון בהקלטה)

במדינה מסוימת ל-80% מהתושבים יש רישיון נהיגה. נבחרו 10 תושבים אקראיים מהמדינה.

א. מהי ההסתברות שבדיוק ל-9 מהם יש רישיון נהיגה?

ב. מה ההסתברות שלפחות ל-9 מהם יש רישיון נהיגה?

ג. מהי התוחלת ומהי סטיית התקן של מספר התושבים שנדגמו ושיש להם רישיון נהיגה?

תרגילים:

1. במדינה 10% מהאוכלוסייה מובטלת. נבחרו 5 אנשים באקראי מאותה אוכלוסייה. נגדיר את X להיות מספר המובטלים שהתקבלו במדגם.
- מהי ההתפלגות של X ?
 - מה ההסתברות שיהיה בדיוק מובטל אחד?
 - מה ההסתברות שכולם יעבדו במדגם?
 - מה ההסתברות ששלושה יעבדו במדגם?
 - מה ההסתברות שלפחות אחד יהיה מובטל?
 - מה תוחלת ומהי השונות של מספר המובטלים במדגם?
2. על פי נתוני משרד התקשורת ל-70% מהאוכלוסייה יש סמארט-פון. נבחרו 10 אנשים באקראי. נגדיר את X כמספר האנשים שנדגמו עם סמארט-פון.
- מהי ההתפלגות של X ? הסבירו.
 - מה ההסתברות שבמדגם ל-8 אנשים יש סמארט-פון?
 - מה ההסתברות שבמדגם לפחות ל-9 יהיו סמארט-פון?
 - מה התוחלת ומה סטיית התקן של מספר האנשים שנדגמו ולהם סמארט-פון?
3. במדינה מסוימת התפלגות ההשכלה בקרב האוכלוסייה מעל גיל 30 היא כזו:

השכלה	נמוכה	תיכונית	תואר I	תואר II ומעלה
פרופורציה	0.1	0.6	0.2	0.1

- נבחרו 20 אנשים אקראיים מעל גיל 30 מהמדינה הנ"ל.
- מה ההסתברות ש-5 מהם אקדמאים?
 - מה התוחלת של מס' בעלי השכלה הנמוכה?

4. 5% מקו היצור פגום. המוצרים נארזים בתוך קופסת קרטון. בכל קופסא 10 מוצרים שונים. הקופסאות נארזות בתוך מכולה. בכל מכולה 20 קופסאות.
- א. מה ההסתברות שבקופסא אקראית לפחות מוצר פגום אחד?
- ב. מה התוחלת ומהי סטיית התקן של מספר הקופסאות במכולה בהן לפחות מוצר פגום אחד?

פתרונות :**שאלה 4 :**

- א. 0.401
ב. תוחלת : 8.025
סטיית תקן : 2.193

שאלה 2 :

- ב. 0.2335
ג. 0.1493
ד. תוחלת : 7
סטיית תקן : 1.449

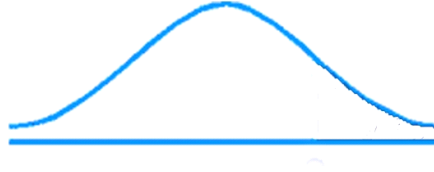
שאלה 3 :

- א. 0.1789
ב. 2

פרק 27 - התפלגויות רציפות מיוחדות - התפלגות נורמלית

רקע:

התפלגות נורמלית הינה התפלגות של משתנה רציף. ישנם משתנים רציפים מסוימים שנהוג להתייחס אליהם כנורמליים כמו: זמן ייצור, משקל תינוק ביום היוולדו ועוד. פונקציית הצפיפות של ההתפלגות הנורמלית נראית כמו פעמון:



לעקומה זו קוראים גם עקומת גאוס ועקומה אחת נבדלת מהשנייה באמצעות הממוצע וסטיית התקן שלה. אלה הם הפרמטרים שמאפיינים את ההתפלגות.

$$X \sim N(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} : \text{נוסחת פונקציית הצפיפות}$$

כדי לחשב הסתברויות בהתפלגות נורמלית יש לחשב את השטחים הרלבנטיים שמתחת לעקומה. כדי לחשב שטחים אלה נמיר כל התפלגות נורמלית להתפלגות נורמלית סטנדרטית על ידי תהליך הנקרא תקנון.

התפלגות נורמלית סטנדרטית היא התפלגות נורמלית שהממוצע שלה הוא אפס וסטיית התקן היא אחת והיא תסומן באות Z .

$$Z \sim N(0, 1^2)$$

תהליך התקנון מבוצע על ידי הנוסחה הבאה:

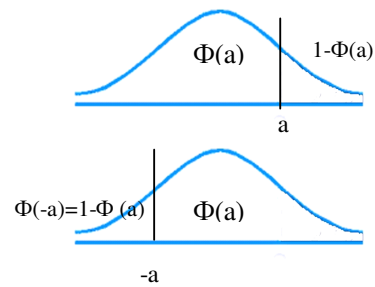
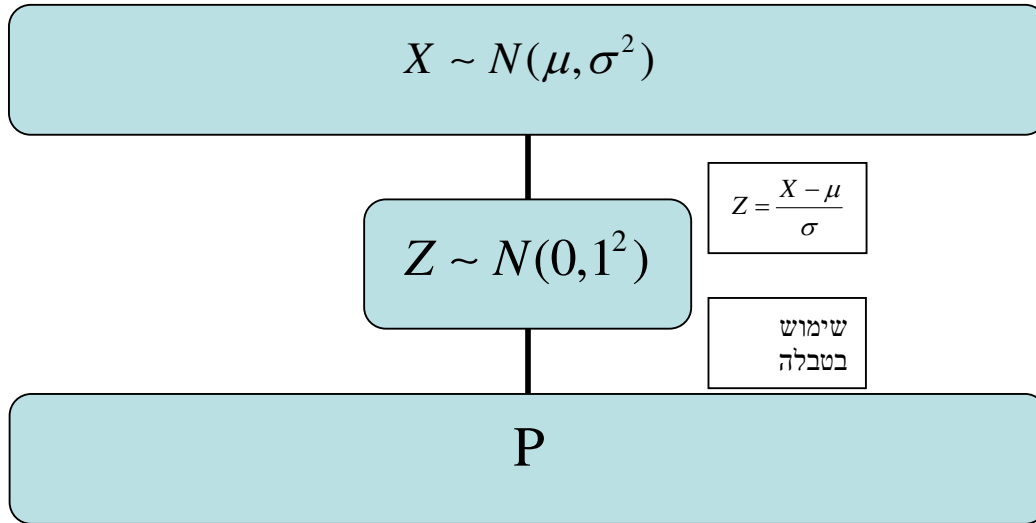
$$Z = \frac{X - \mu}{\sigma}$$

אחרי תקנון מקבלים ערך הנקרא ציון תקן.

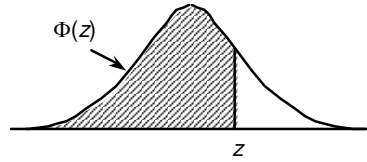
ציון התקן משמעו בכמה סטיות תקן הערך סוטה מהממוצע.

לאחר חישוב ציון התקן של ערך מסוים נעזרים בטבלה של ההתפלגות הנורמלית הסטנדרטית לחישוב השטח הרצוי.

ובאופן כללי נתאר את הסכמה הבאה :



טבלת ההתפלגות המצטברת הנורמלית סטנדרטית – ערכי $\Phi(z)$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

z	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.891	4.417
$\Phi(z)$	0.90	0.95	0.975	0.99	0.995	0.999	0.9995	0.99995	0.999995

דוגמה: (הפתרון בהקלטה)

משקל חפיסות שוקולד המיוצרות בחברה מתפלג נורמלית עם ממוצע 100 גרם בסטיית תקן של 8 גרם.

- א. מה אחוז חפיסות השוקולד ששוקלות מתחת ל- 110 גרם?
- ב. מה אחוז חפיסות השוקולד השוקלות מעל 110 גרם?
- ג. מה אחוז חפיסות השוקולד השוקלות מתחת ל 92 גרם?
- ד. מהו המשקל ש90% מהחפיסות בקו הייצור שוקלים פחות מהם?

תרגילים:

1. הגובה של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 170 ס"מ וסטית תקן של 10 ס"מ.
 - א. מה אחוז האנשים שגובהם מתחת ל- 182.4 ס"מ?
 - ב. מה אחוז האנשים שגובהם מעל 190 ס"מ?
 - ג. מה אחוז האנשים שגובהם בדיוק 173.6 ס"מ?
 - ד. מה אחוז האנשים שגובהם מתחת ל- 170 ס"מ?
 - ה. מה אחוז האנשים שגובהם לכל היותר 170 ס"מ?

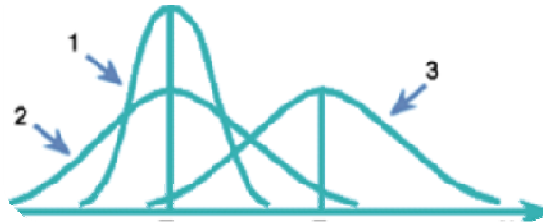
2. נתון שהזמן שלוקח לתרופה מסוימת להשפיע מתפלג נורמלית עם ממוצע של 30 דקות ושונות של 9 דקות רבועות.
 - א. מהי פרופורציית המקרים בהן התרופה תעזור אחרי יותר משעה?
 - ב. מה אחוז מהמקרים שבהן התרופה תעזור בין 35 ל-37 דקות?
 - ג. מה הסיכוי שהתרופה תעזור בדיוק תוך 36 דקות?
 - ד. מה שיעור המקרים שבהן ההשפעה של התרופה תסטה מ-30 דקות בפחות מ-3 דקות?

3. המשקל של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 60 ק"ג וסטיית תקן של 8 ק"ג.
 - א. מה אחוז האנשים שמשקלם נמוך מ- 55 ק"ג?
 - ב. מהי פרופורציית האנשים באוכלוסייה שמשקלם לפחות 50 ק"ג?
 - ג. מהי השכיחות היחסית של האנשים באוכלוסייה שמשקלם בין 60 ל- 70 ק"ג?
 - ד. לאיזה חלק מהאוכלוסייה משקל הסוטה מהמשקל הממוצע בלא יותר מ- 4 ק"ג?
 - ה. מה הסיכוי שאדם אקראי ישקול מתחת ל- 140 ק"ג?

4. משקל תינוקות ביום היוולדם מתפלג נורמלית עם ממוצע של 3300 גרם וסטיית תקן 400 גרם.
 - א. מצאו את העשירון העליון.
 - ב. מצאו את האחוזון ה-95.
 - ג. מצאו את העשירון התחתון.

5. ציוני מבחן אינטליגנציה מתפלג נורמלית עם ממוצע 100 ושונוות 225 .
- מה העשירון העליון של הציונים במבחן האינטליגנציה?
 - מה העשירון התחתון של ההתפלגות?
 - מהו הציון ש- 20% מהנבחנים מקבלים מעליו?
 - מהו האחוזון ה- 20?
 - מהו הציון ש- 5% מהנבחנים מקבלים מתחתיו?
6. נפח משקה בבקבוק מתפלג נורמלית עם סטיית תקן של 20 מ"ל, נתון ש-33% מהבקבוקים הם עם נפח שעולה על 508.8 מ"ל.
- מה ממוצע נפח משקה בבקבוק ?
 - 5% מהבקבוקים המיוצרים עם הנפח הגבוה ביותר נשלחים לבדיקה, החל מאיזה נפח שולחים בקבוק לבדיקה?
 - 1% מהבקבוקים עם הנפח הקטן ביותר נתרמים לצדקה, מהו הנפח המקסימלי לצדקה?
7. אורך חיים של מכשיר מתפלג נורמלית . ידוע שמחצית מהמכשירים חיים פחות מ- 500 שעות, כמו כן ידוע ש- 67% מהמכשירים חיים פחות מ- 544 שעות.
- מהו ממוצע אורך חיי מכשיר?
 - מהי סטיית בתקן של אורך חיי מכשיר?
 - מה הסיכוי שמכשיר אקראי יחיה פחות מ- 460 שעות?
 - מהו המאון העליון של אורח חיי מכשיר?
 - 1% מהמכשירים בעלי אורך החיים הקצר ביותר נשלח למעבדה לבדיקה מעמיקה. מהו אורך החיים המקסימלי לשליחת מכשיר למעבדה?

8. להלן שלוש התפלגויות נורמליות של שלוש קבוצות שונות ששורטטו באותה מערכת צירים. ההתפלגויות מוספרו כדי להבדיל ביניהן.



א. לאיזו התפלגות הממוצע הגבוה ביותר?
 ב. במה מבין המדדים הבאים התפלגות 1 ו 2 זהות?

א. בעשירון העליון.

ב. בממוצע.

ג. בשונות.

ג. לאיזו התפלגות סטיית התקן הקטנה ביותר?

א. 1

ב. 2

ג. 3

ד. אין לדעת.

9. הזמן שלוקח לאדם להגיע לעבודתו מתפלג נורמלית עם ממוצע של 40 דקות וסטית תקן של 5 דקות.

א. מה ההסתברות שמשך הנסיעה של האדם לעבודתו יהיה לפחות שלושת רבעי השעה?

ב. אדם יצא לעבודתו בשעה 08:10 מביתו. הוא צריך להגיע לעבודתו בשעה 09:00. מה הסיכוי שיאחר לעבודתו?

ג. אם ידוע שזמן נסיעתו לעבודה היה יותר משלושת רבעי השעה. מה ההסתברות שזמן הנסיעה הכולל יהיה פחות מ- 50 דקות?

ד. מה הסיכוי שבשבוע (חמישה ימי עבודה) בדיוק פעם אחת יהיה זמן הנסיעה לפחות שלושת רבעי השעה?

פתרונות :

<u>שאלה 3</u>		<u>שאלה 1</u>	
א.	26.43%	א.	89.25%
ב.	89.44%	ב.	2.28%
ג.	39.44%	ג.	0
ד.	0.383	ד.	50%
ה.	100%		

<u>שאלה 7</u>		<u>שאלה 5</u>	
א.	500	א.	119.2
ב.	100	ב.	80.8
ג.	0.3446	ג.	112.6
ד.	733	ד.	87.4
ה.	267		

<u>שאלה 9</u>		<u>שאלה 8</u>	
א.	0.1587	א.	3
ב.	0.0228	ב.	בממוצע.
ג.	0.8563	ג.	1
ד.	0.3975		

<u>שאלה 10</u>	
א.	1925
ב.	0.2266
ג.	0.1587

פרק 28 - התפלגות הדגימה

א. התפלגות ממוצע המדגם ומשפט הגבול המרכזי

1. משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.
 - א. מה ההסתברות שתינוק אקראי בעת הלידה ישקול פחות מ-3800 גרם?
 - ב. נתון שביום מסוים נולדו 4 תינוקות.
 - ג. מה ההסתברות שהמשקל הממוצע שלהם יעלה על 4 ק"ג?
 - ד. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-2.5 ק"ג?
 - ה. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת בלא יותר מ-50 גרם?

2. הזמן הממוצע שלוקח לאדם להגיע לעבודתו 30 דקות עם שונות של 16 דקות רבועיות. האדם נוסע לעבודה במשך שבוע 5 פעמים. לצורך פתרון הניחו שזמן הנסיעה לעבודה מתפלג נורמאלית.
 - א. מה ההסתברות שבמשך שבוע משך הנסיעה הממוצע יהיה מעל 33 דקות?
 - ב. מהו הזמן שבהסתברות של 90% ממוצע משך הנסיעה השבועי יהיה פחות ממנו?
 - ג. מה ההסתברות שממוצע משך הנסיעה השבועי יהיה מרוחק מ-30 דקות בלפחות 2 דקות?

3. נפח היין בבקבוק מתפלג נורמאלית עם תוחלת של 750 סמ"ק וסטיית תקן של 10 סמ"ק.
 - א. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה בדיוק 755 סמ"ק?
 - ב. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה יותר מ-755 סמ"ק?
 - ג. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה לפחות 755 סמ"ק?
 - ד. בקבוקי היין שבארגז נמוזגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?

4. משתנה מתפלג נורמאלית עם תוחלת 80 וסטיית תקן 4 .
- א. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?
- ב. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה שגודל המדגם הוא 16?
- ג. הסבר את ההבדל בתשובות של שני הסעיפים.
5. לפי הערכות הלשכה המרכזית לסטטיסטיקה השכר הממוצע במשק הוא 8000 ₪ עם סטיית תקן של 3000 ₪. מה ההסתברות שבמדגם מקרי של 100 עובדים השכר הממוצע יהיה יותר מ- 8500 ₪?
6. אורך צינור שמפעל מייצר הינו עם ממוצע של 70 ס"מ וסטיית תקן של 10 ס"מ
- א. נלקחו באקראי 100 מוטות, מה ההסתברות שממוצע אורך המוטות יהיה בין 68 ל 78 ס"מ?
- ב. יש לחבר 2 בניינים באמצעות מוטות. המרחק בין שני הבניינים הינו 7200 ס"מ. מה ההסתברות ש 100 המוטות יספיקו למלאכה?
- ג. מה צריך להיות גודל המדגם המינימאלי, כדי שבהסתברות של 5% ממוצע המדגם יהיה קטן מ-69 ס"מ. העזר במשפט הגבול המרכזי.

התפלגות מספר ההצלחות במדגם ופרופורציות ההצלחות במדגם וקירוב נורמאלי
להתפלגות הבינומית

1. נתון ש-20% מאוכלוסייה מסוימת אקדמאית. נבחרו באקראי 10 אנשים באותה אוכלוסייה.
 - א. מה ההסתברות שלושה מהם אקדמאים?
 - ב. מה ההסתברות שלכל היותר אחד מהם אקדמאי?
 - ג. מה התוחלת ומהי סטיית התקן של מספר האקדמאים במדגם?

2. הסיכוי שמבוגר יחלה בשפעת בחורף מוערך ב-0.4. נדגמו 30 מבוגרים אקראיים.
 - א. חשבו את הסיכוי שבדיוק 15 יחלו בשפעת (1) לפי ההתפלגות הבינומית ; (2) לפי הקירוב הנורמאלי.
 - ב. חשבו את הסיכוי לכל היותר 14 יחלו בשפעת.
 - ג. חשבו את הסיכוי שלפחות 11 יחלו בשפעת.

3. במפעל 10% מהמוצרים פגומים. נלקחו 100 מוצרים באקראי מקו הייצור.
 - א. מה ההסתברות שנדגמו לכל היותר 6 מוצרים פגומים?
 - ב. מה ההסתברות שמספר המוצרים הפגומים יהיה לכל היותר 11 במדגם?
 - ג. מה ההסתברות שיהיו בדיוק 11 פגומים?

4. ציוני פסיכומטרי בקרב הנרשמים למוסד מסוים מתפלגים נורמאלית עם ממוצע 500 וסטיית תקן 100. למוסד מסוים הוחלט לקבל אך ורק סטודנטים שקיבלנו מעל 600 בפסיכומטרי. 100 סטודנטים אקראיים נרשמו למוסד. מה ההסתברות שלפחות 20 יתקבלו?

5. במדינה יש 10% של אבטלה. נדגמו באקראי 140 אנשים מהמדינה.
 - א. מה התוחלת ומהי השונות של פרופורציות המובטלים שנדגמו?
 - ב. מה ההסתברות שבמדגם לפחות 10% יהיו מובטלים?
 - ג. מה ההסתברות שלכל היותר 9% מהמדגם יהיו מובטלים?

6. נניח ש-30% מהאוכלוסייה תומכת בהצעת חוק מסוימת. אם נדגום מהאוכלוסייה 200 איש. חשבו את ההסתברויות הבאות :
 - א. לפחות 35% יתמכו בהצעת החוק במדגם.
 - ב. לכל היותר 25% יתמכו בהצעת החוק במדגם.
 - ג. יותר מ – 27% יתמכו בהצעת החוק במדגם.

פתרונות:**פרק א' - התפלגות ממוצע מדגם ומשפט הגבול המרכזי**

<u>שאלה 2</u>	<u>שאלה 1</u>
א. 0.0465	א. 0.8413
ב. 32.29	ב. 0.0013
ג. 0.2628	ג. 0
	ד. 0.1974
<u>שאלה 4</u>	<u>שאלה 3</u>
א. 0.5468	א. 0
ב. 0.6826	ב. 0.1587
	ג. 0.1587
	ד. 0.5

שאלה 5

0.0475

שאלה 6

א. 0.9772

ב. 0.0228

ג. 271

פרק ב' - התפלגות מספר ההצלחות במדגם ופרופורציות ההצלחות במדגם וקירוב נורמאלי להתפלגות הבינומית

<u>שאלה 1</u>	<u>שאלה 2</u>
א. 0.201	א. 0.0783 (1) ; 0.0794 (2)
ב. 0.3758	ב. 0.8238
ג. התוחלת: 2	ג. 0.7123
סטיית התקן: 1.2649	
<u>שאלה 3</u>	<u>שאלה 4</u>
א. 0.121	0.1611
ב. 0.6915	
ג. 0.124	
<u>שאלה 5</u>	<u>שאלה 6</u>
א. התוחלת: 0.1	א. 0.0618
השונות: 0.00064	ב. 0.0618
ב. 0.5	ג. 0.8238
ג. 0.3446	