

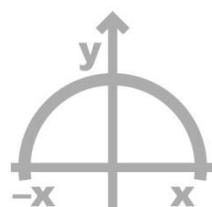
# ביостטיסטיקה לרפואת שיניים



$$\sqrt{2}$$
A gray square containing the number 1 in the top-left corner and the number 1 in the bottom-right corner, with a diagonal line through it.



$$\{\sqrt{x}\}^2$$
A large white bracketed expression where the inner term is  $\sqrt{x}$  and the outer term is squared, resulting in  $\{\sqrt{x}\}^2$ .



## תוכן העניינים

1.	סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה	1
5.	סטטיסטיקה תיאורית- הצגה של נתונים	
16.	סטטיסטיקה תיאורית- סכימה.	
(לא ספר)	סטטיסטיקה תאורית- מדדי מרכז	
(לא ספר)	סטטיסטיקה תיאורית- מדדי פיזור	
20.	סטטיסטיקה תיאורית - מדדי פיזור - טווח בין רבוני	
24.	סטטיסטיקה תיאורית- מדדי מיקום יחס-ציוון תקן	
26.	סטטיסטיקה תיאורית-אחווזונים בטליה בדידה	
28.	סטטיסטיקה תיאורית-תרשים קופסה	
30.	סטטיסטיקה תיאורית- ניתוח פלטים	
32.	סטטיסטיקה תיאורית - טרנספורמציה לינארית	
35.	סטטיסטיקה תיאורית שאלות אמריקאיות	
42.	מקדם המתאים של פירסון	
58.	יסודות ההסתברות	
62.	פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכללים	
70.	הסתברות מותנית-במרחב מודגם אחד	
73.	הסתברות מותנית - מרחב לא אחד	
76.	דיגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה	
81.	תלות ואי תלות בין מאורעות	
83.	הערכת כלים אבחנתיים	
87.	שאלות מסכימות בהסתברות	
90.	המשתנה המקרי הבדיקה - פונקציית ההסתברות	
93.	המשתנה המקרי הבדיקה - תוחלת - שונות וסטיית תקן	

## תוכן העניינים

24. המשטנה המקרי הבודד - טרנספורמציה לינארית.....	96
25. התפלגיות בדיםות מיוחדות - התפלגות בינומית.....	99
26. התפלגיות רציפות מיוחדות - התפלגות נורמלית.....	103
27. הסקה סטטיסטית - הקדמה.....	111
28. התפלגות הדגימה ומשפט הגבול המרכזי.....	114
29. מושגי יסוד באמידה.....	130
30. רוח סמך לתוחלת (ממוצע).....	135
31. רוח סמך לפרופורציה.....	145
32. רוח סמך להפרש תוחלות (ממוצעים) במדגים בלתי תלויים.....	151
33. רוח סמך לתוחלת (ממוצע) ההפרשים במדגים מזוגים.....	155
34. שאלות מסכימות על רוחי סמך.....	157
35. מבוא לבדיקה השערות על פרמטרים.....	159
36. בדיקת השערות על תוחלת (ממוצע).....	165
37. בדיקת השערות על פרופורציה.....	197
38. בדיקת השערות על הפרש תוחלות במדגים בלתי תלויים.....	210
39. בדיקת השערות לתוחלת ההפרש במדגים מזוגים.....	225
40. הקשר בין רוח סמך לבדיקה השערות להפרש תוחלות.....	235
41. מבחני חি בריבוע.....	238
42. ניתוח שונות חד כיוונית.....	256
43. שאלות מסכימות בבדיקה השערות.....	265

# bijostatystyka dla medycyny

פרק 1 - סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה

תוכן העניינים

1. כללי .....

## סטטיסטיקה תיאורית – סיווג משתנים וסולמות מדידה:

### רקע:

סטטיסטיקה תיאורית הוא ענף בו לומדים כיצד לאסוף נתונים, להציג אותם ולנתה אותם. בסטטיסטיקה תיאורית אנו פונים לקבוצה מסוימת, ובאותה קבוצה אנו אוספים נתונים על היחסות בהולה קבוצה. משתנה – תוכנה שיכולה לקבל מספר ערכים: דעה פוליטית, מקום מגורים, גובה של אדם וכדומה. חלוקה אחת של המשתנים הנמדדים היא לפי סולמות מדידה:

### מיון משתנים לפי סולמות המדידה:

1. **סולם שמי (גומינלי)** – משתנה של ערכיו יש משמעות רק מבחינת הזהות ואין עניין של יותר או פחות.  
לדוגמה: מצב משפחתי (רווק/ נשוי/ אלמן/ גירוש), אזרח מגורים. משתנה דיקוטומי (הינו מסולם שמי) אותו משתנים שיש להם רק שני ערכים אפשריות זכר/ נקבה. מעש/לא מעש.
2. **סולם סדר (אורדינלי)** – כאשר לערכים של המשתנה בנוסף לשם ישנה גם משמעות לסדר אבל אין משמעות לגודל ההפרש. למשל, דרגה בצבא.
3. **סולם רוחים (אינטראולי)** – משתנה של ערכים שלו בנוסף לשם ולסדר בניהם יש משמעות לרוחים בין הערכים אבל אין משמעות לייחס בין הערכים. למשל, קומה בבניין. סולם לא כל כך פופולרי.

### טולט מנה/יחס:

משתנה של ערכיו בנוסף לשם, לסדר ולרוח יש משמעות גם לייחס בין הערכים. למשל, מספר מכוניות למשפחה, משקל אדם בק"ג. הדרך הקלה ביותר כדי לזהות עם הטולט הוא סולם מנה היא על ידי מבחן האפס. בסולם מנה האפס הוא מוחלט, אבסולוטי, ומיצג אין.

## סוגי משתנים:

נבע סיווג של המשתנים :

### משתנה איקומי

משתנה של ערכיו אין משמעות של יותר או פחות, אין עניין כמותי לערכים המתקבלים. כמו: מקומות מגורים של אדם (רעננה, תל אביב, אשדוד...),מין האדם (זכר, נקבה), מצב משפחתי (רווק, נשוי, גירוש, אלמן).

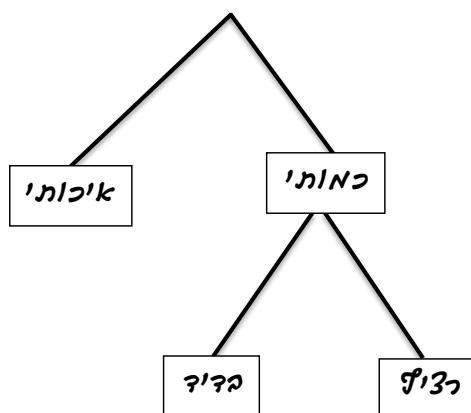
### משתנה כמותי

משתנה שערךיו הם מספריים להם יש משמעות כמותית כמו: גובה אדם בס"מ, ציון בבחינה ובדומה. את המשתנה הכמותי נסוג לשני סוגים :

משתנה בדיד : משתנה שערךיו מתקבלים מתוך סידרה של ערכים אפשריים. כמו: מספר ילדים למשפחה (1,2,3...), ציון בבחינה (מ-0 ועד 100 בקפיצות של 1).

משתנה רציף : משתנה שערךיו מתקבלים מתוך אינסוף ערכים בתחום מסוים, הערכים מתקבלים ברצף – ללא קפיצות של ערכים.

דוגמאות: גובה בס"מ – אם הגובה הנמוך ביותר הוא 150 ס"מ ועד 190 ס"מ – הגבהים בקבוצה הם ברצף. גם בין 160 ל-161 ס"מ יש רצף אינסופי של ערכים אפשריים (כמו 160.233 ס"מ, למשל).



**שאלות:**

**1)** באיזה סולם מדידה המשתנים הבאים נחקרים (שמי/סדר/רווחים/מנה) :

- א. גובה (בס"מ).
- ב. מספר ילדים למשפחה.
- ג. מידת החרדה לפני מבחון.
- ד. שביעות רצון משירות לקוחות בסקלה מ-1 עד 7 (1 - כלל לא מרוצה עד 7 - מרוצה מאוד)
- ה. השכלה.
- ו. מספר אוטובוס.
- ז. מקום מגוריים.
- ח.מין (1=גבר ; 2=אישה).
- ט. מידת נעלים.

**2)** להלן התפלגות מספר האיחורים לעובדה בחודש של העובדים בחברת "סטארר" :

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

בחברה 200 עובדים.

- א. מהו המשתנה הנחקר כאן?
- ב. האם מדובר במשתנה איקוטי או כמותי?  
אם הוא כמותי האם הוא בדיד או רציף?  
באיזה סולם מדידה המשתנה?

**3)** להלן רשימה של משתנים כמותיים. ציינו האם הוא משתנה רציף/בדיד :

- א. שכר ב-₪.
- ב. ציון בחינות בגרות.
- ג. תוצאה של הטלת קובייה.
- ד. מהירות ריצה בתחרויות.
- ה. שיעור התמיכה במשלה.

**תשובות סופיות:**

- |                                                       |             |          |
|-------------------------------------------------------|-------------|----------|
| ג. סדר.                                               | ב. מנה.     | א. מנה.  |
| ו. שמי.                                               | ה. מנה/סדר. | ד. סדר.  |
| ט. סדר.                                               | ח. שמי.     | ז. שמי.  |
| (2) א. מספר האichenרים.      ב. כמותי בדיד בסולם מנה. |             |          |
| ג. בדיד.                                              | ב. בדיד.    | א. רציף. |
|                                                       | ג. רציף.    | ד. רציף. |

# bijostatystyka dla medycyny

פרק 2 - סטטיסטיקה תיאורית- הצגה של נתונים

תוכן העניינים

5 ..... 1. כללי .....

## סטטיסטיקה תיאורית – הצגה של נתונים:

---

**רקע:**

דרכים להציג נתונים שנאספו :

**רישימה של תצפיות:**

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

**טבלת שכיחיות בדידה:**

שכיחותיחסית ב אחוזים	שכיחות – $f(x)$	שם המשתנה – $X$
$\frac{f_1}{N} \cdot 100$	$f_1$	$X_1$
$\frac{f_2}{N} \cdot 100$	$f_2$	$X_2$
$\frac{f_3}{N} \cdot 100$	$f_3$	$X_3$
⋮	⋮	⋮
$\frac{f_x}{N} \cdot 100$	$f_k$	$X_k$
100%	$N = \sum_{i=1}^k f_i$	<b>סה"כ</b>

רושמים את התצפיות בטבלה שבה עמודה אחת מבטא את ערכי המשתנה והשנייה את השכיחות. עיל עבור משתנה איקומי וכמותי בדיד וככיש מספר רב של תצפיות. לא עיל למשתנה כמותי רציף.

**דוגמה:**

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 השכיחויות.

השכיחויות  $F_i$  – השכיחות המצטברת נותנת כמה תצפויות קטנות או שותת לערך.

שכיחות יחסית (פרופורציה) – השכיחות מחולקת לכמויות התצפויות הכללי:

$$\frac{f_i}{n} \text{ -- איזה חלק מהתצפויות בקבוצה שותת לערך.}$$
**טבלת שכיחיות בחלוקת:**

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

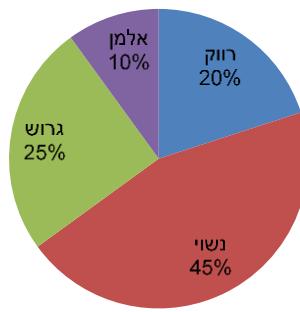
**דוגמה:**

נתנו לקבוצת ילדים לבצע משימה, בדקו את התפלגות זמן הביצוע, בדיקות.  
להלן החתפלגות שהתקבלה :

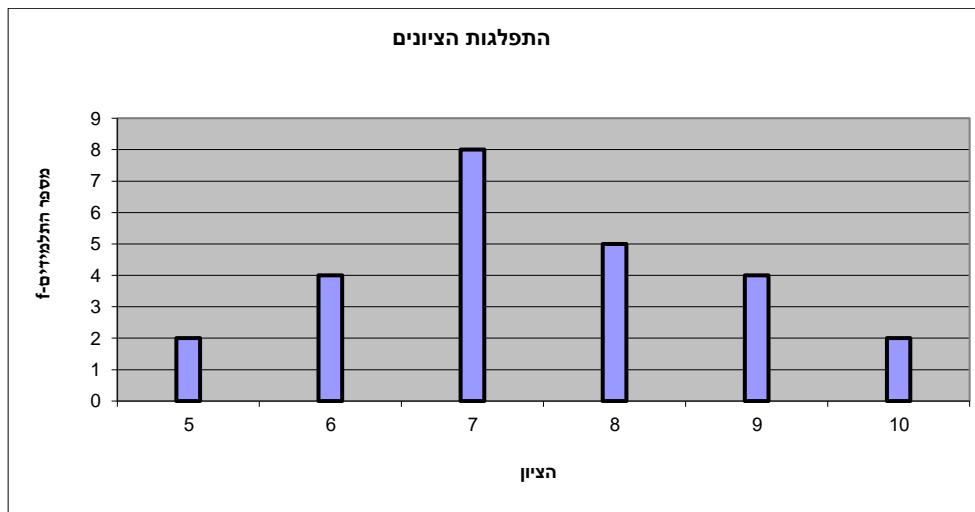
זמן בדיקות	מספר הילדים
20	0.5-3.5
18	3.5-9.5
14	9.5-19.5
8	19.5-29.5

**דיאגרמת עוגה:**

זהו התיאור הגרפי של משתנה איקומי. בדיאגרמת עוגה כל ערך במשתנה מקבל "נתח", שהוא פרופורציוני לשכיחות היחסית של ערך המשתנה בתנאים.

**התפלגות המצב המשפחתי****דיאגרמת מקלות:**

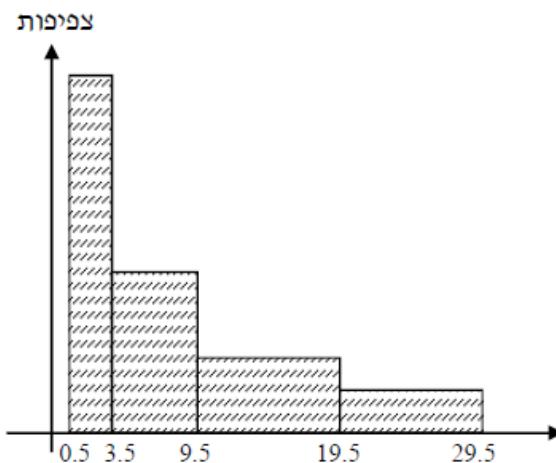
הציר האופקי הוא הציר של המשתנה והציר האנכי של השכיחות, כך שהגובה של המקל מעיד על השכיחות. לבנתי למשתנה כמותי בלבד. לא נהוג להשתמש בתיאור למשתנה איקומי וכמו כן לא למשתנה כמוותי רציף, וכן בסולמות מדידה עבור משתנה מסולם סדר.



**ההיסטוגרמה:**

ההיסטוגרמה היא הדרך הגרפי כדי לתאר טבלת שכיחיות בחלוקת, והיא רלוונטי למשתנה כמותי רציף. בההיסטוגרמה הציר האופקי הוא הציר של המשטנה והציר האנכי הוא הציר של הצפיפות. הצפיפות מחושבת בכל מחלוקת על ידי חלוקת השכיחות ברוחב של כל המחלוקת, והוא נותנת את מספר התצפיות הממוצע בכל מחלוקת ייחודה. אם המחלוקות הן שוות ברוחב, ניתן לשרטט את הההיסטוגרמה לפי השכיחות ואין צורך בцеיפויות.

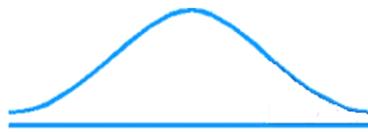
cefipot	cefipot	מצטברת	שכיחות	ממוצע	רוחב	X
6.6667	20	20	2	3	0.5 - 3.5	
3	38	18	6.5	6	3.5 - 9.5	
1.4	52	14	14.5	10	9.5 - 19.5	
0.8	60	8	24.5	10	19.5 - 29.5	

**פוליגון – מצולעון:**

אם נחבר את אמצע קצה כל מלבן בקווים ישרים. ניתן לראות חזותי לצורה של התפלגות המשטנה.

**צורות התפלגות נפוצות:****התפלגות סימטרית פעמונית**

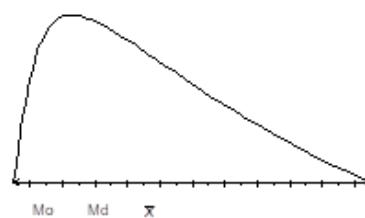
רוב התצפויות במרכז, וככל שנתרחק מהמרכז יהיה פחות תצפויות באופן סימטרי. לדוגמה, ציוני IQ.



ישנן התפלגויות סימטריות שאינן פעמוניות, כגון :

**התפלגות אסימטרית ימנית ( חיובית )**

רוב התצפויות מתקבלות ערכים נמוכים ויש מיעוט הולך וקטן של תצפויות שמקבלות ערכים גבוהים קיצוניים. לדוגמה, שכר במשק.

**התפלגות א-סימטרית  
ימנית או חיובית**
**התפלגות אסימטרית שמאלית (שלילית)**

רוב התצפויות מתקבלות ערכים גבוהים ויש מיעוט הולך וקטן של תצפויות שמקבלות ערכים נמוכים קיצוניים. לדוגמה, אורך חיים.



**שאלות:**

**1)** בסקר צפיה בטלוייזיה התקבלו התוצאות הבאות : 25 צפו בערוץ הראשון, 25 צפו בערוץ 10, 75 צפו בערוץ השני, 50 צפו באחד מערוצי הcabלים ו-25 לא צפו בטלוייזיה 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":

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

- א. מהו המשתנה הנחקר?
- ב. מהי פרופורצית התלמידים שمعدיפים תנ"ך?

**3)** להלן התפלגות ההשכלה במקום העבודה מסוימת :

השכלה	מספר העובדים
ນמוכה	60
תיכונית	120
אקדמאית	20

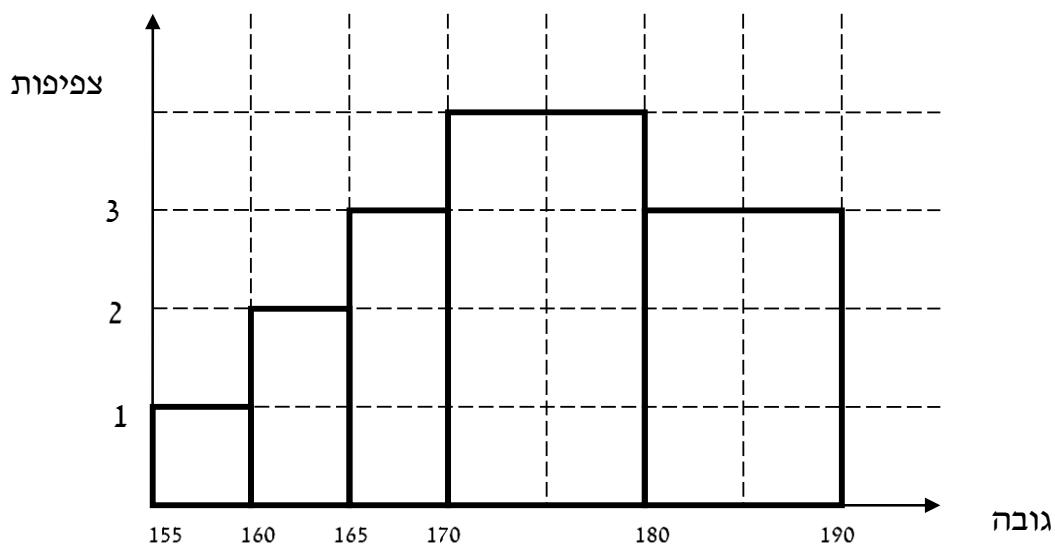
- א. מהו המשתנה הנחקר?
- מماיזה סולם הוא?
- ב. תארו את הנתונים באופן גרפי.

**4)** להלן רשימת הציונים של 20 תלמידים שנבחנו ב מבחון הבנת הנקרא :

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

- א. מהו המשתנה? האם הוא בדיד או רציף?
- ב. תארו את הרשימה בטבלת שכיחויות.
- ג. הוסיפו שכיחיות יחסית לטבלה.
- ד. תארו את הנתונים באופן גרפי.

5) להלן היסטוגרמה המתארת את התפלגות הגבאים בס"מ של קבוצה מסוימת:



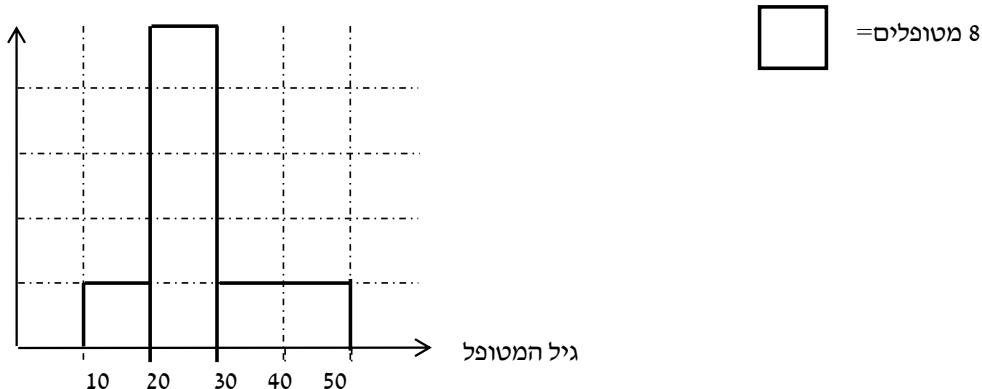
- מהו המשתנה הנחקר? האם הוא בדיד או רציף?
- תארו את הנתונים בטבלת שכיחיות בחלוקת.
- הוסיפו שכיחות יחסית לטבלה.
- הוסיפו את הצפיפות של כל מחלוקת לטבלה.
- מהי צורת ההתפלגות של הגבאים?

6) 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

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

7) להלן גיל המטופלים של ד"ר שורץ בשנים :  
 קנה מידת :



- א. מה המשטנה הנחקר? האם הוא בדיד או רציף?
- ב. מהי הקבוצה הנחקרת?
- ג. תרגמו את ההיסטוגרמה לטבלת שכיחות.
- ד. מהי הפרופורציה של המטופלים של ד"ר שורץ בגילאים 20-30?

**תשובות סופיות:**

ב. עיין גרף מלא בסרטון הוידאו.

(1) א. להלן טבלה :

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

ב. 19.6%.

(2) א. מקצוע מועדף.

ב. עיין גרף מלא בסרטון הוידאו.

(3) א. משתנה נחקר : השכלה, סוג : סדר.

ב+ג. להלן טבלה:

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

- (4) א. המשתנה: ציון, משתנה בדיד.  
 ד. עיין גרף מלא בסרטון הויידאו.

ה. אסימטריה:

(5) א. גובה בס"מ, רציף.

$d$	%	$\frac{f(x)}{n}$	$f(x)$	$x$
1	5%	$\frac{5}{100}$	5	<b>155-160</b>
2	10%	$\frac{10}{100}$	10	<b>160-165</b>
3	15%	$\frac{15}{100}$	15	<b>165-170</b>
4	40%	$\frac{40}{100}$	40	<b>170-180</b>
3	30%	$\frac{30}{100}$	30	<b>180-190</b>

- ב. סימטרית.  
 ב. המטופלים של ד"ר שורץ.  
 ה. 62.5%.

- א. עין גוף מלא בסרטון הוידאו.  
 א. המשתנה : גיל בשנים, משתנה רציף.  
 ד. להלן טבלה:

$f(x)$	$x$
8	<b>10-20</b>
40	<b>20-30</b>
16	<b>30-50</b>

# bijostatystyka dla medycyny

## פרק 3 - סטטיסטיקה תיאורית- סכימה

תוכן העניינים

1. כללי .....

16 .....

## סטטיסטיקה תיאורית – סכימה:

**רקע:**

בסטטיסטיקה ישנה צורת רישום מקובלת לסקום 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

(הסביר מלא מופיע בסרטוניים באתר).

**שאלות:**

- 1) במבנה 5 דירות. לכל דירה רשמו את מספר החדרים שיש בדירה ( $X$ ), ומספר הנפשות החיים בדירה ( $Y$ )如下：

<b><math>Y</math></b>	<b><math>X</math></b>	<b>מספר דירה</b>
1	2	1
1	3	2
2	2	3
3	4	4
2	3	5

.  $\sum_{i=1}^3 X_i$  . א.

.  $\sum_{i=1}^5 Y_i$  . ב.

.  $\sum_{i=1}^4 X_i$  . ג.

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

.  $\sum X_i$  . ה.

.  $\sum X_i Y_i$  . ו.

.  $\sum (X_i) \sum (Y_i)$  . ז.

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

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

$$\cdot \sum_{i=1}^4 y_i . \text{א}$$

$$\cdot \sum_{i=1}^6 a . \text{ב}$$

$$\cdot \sum_{i=1}^6 x_i y_i . \text{ג}$$

$$\cdot \sum_{i=1}^6 (x_i + y_i) . \text{ד}$$

$$\cdot \sum_{i=1}^6 x_i + a . \text{ה}$$

**3)** קבעו לכל זהות האם היא נכונה:

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

$$\cdot \sum_{i=1}^n a = a \cdot n . \text{ב}$$

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

**תשובות סופיות:**

- |          |              |           |           |     |
|----------|--------------|-----------|-----------|-----|
| .121. ז. | .11.ג        | .9.ב      | .7.א.     | (1) |
| .126. ז. | .27.ו        | .14.ה.    | .14.ה.    |     |
| .12.ד    | .7.ג         | .12.ב     | .3.א.     | (2) |
|          | ג. לא נכוна. | ב. נכוна. | .14.ה.    |     |
|          |              |           | א. נכוна. | (3) |

# bijustisjka le-refoat shinim

## פרק 4 - סטטיסטיקה תאורית- מדדי מרכז

תוכן העניינים

1. כללי .....  
(ללא ספר) .....

# bijustisjka le-refoat shinim

## פרק 5 - סטטיסטיקה תיאורית- מדדי פיזור

תוכן העניינים

1. כללי .....

(ללא ספר) .....

# bijostatystyka dla medycyny

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

תוכן העניינים

20 ..... 1. טווח בין רבוני

## סטטיסטיקה תיאורית – מדדי פיזור – טווח בין רביעוני:

**רקע:**

הטווח הבינו-רביעוני (יש הקוראים לו התחום הבינו-רביעוני) נותן את הטווח בין הרבעונים בו נמצאים 50% מהתציפות המרכזית. הרעיון ליצור מדד פיזורי שלא ניתן לתוצאות חריגות ביותר. כדי לחשב את הטווח הבינו-רביעוני יש למצוא את הרבעון התיכון והעליו של התפלגות התציפות.

רביעון תיכון – ערך שמחולק את ההתפלגות לשניים.  
25% מהמקרים נמצאים ממנה או שווים לו ו-75% מהמקרים גבוהים או שווים לו.  
סימון:  $Q_2$ .

רביעון עליון – ערך שמחולק את ההתפלגות לשניים.  
75% מהמקרים נמצאים ממנה או שווים לו ו-25% מהמקרים גבוהים או שווים לו.  
סימון:  $Q_3$ .

הטווח הבינו-רביעוני הוא הפער בין שני הרבעונים:  $IQR = Q_3 - Q_1$ .

**שלבים במציאת טווח בין-רביעוני בטבלת שכיחיות:**

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

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

שלב ג: נמצא את הטווח הבינו-רביעוני: נחסר את הרבעונים:  $IQR = Q_3 - Q_1$ .

**דוגמה (פתרון בהקלטה):**

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

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

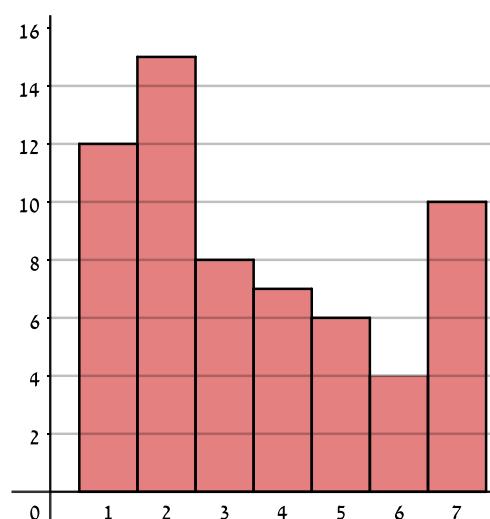
**שאלות:**

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

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

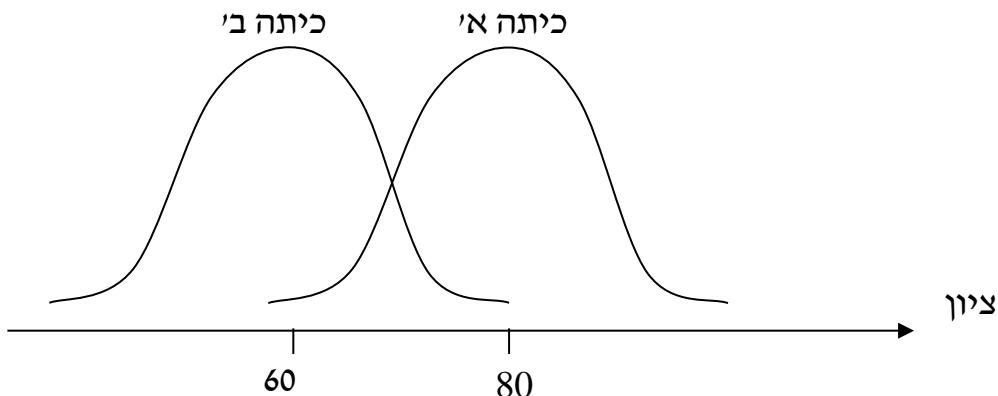
מהו הטווח הבין-רביעוני של מספר המכוניות למשפחה בישוב "הגורן"?

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



- א. מה מייצגים הערכים בציר האופקי?
- ב. מהו הטווח הבין-רביעוני של מספר ימי המחללה של המורים
- ג. אם נוסיף 25 מורים אשר הצהירו שמספר ימי המחללה השנתיים שלהם הוא 4 ימים, כיצד הדבר ישנה את הטווח הבין-רביעוני? הסבירו.
- ד. אם מסתבר שחלק מהמורים בסקר הצהירו שהם חולו 7 ימים בשנה אבל בפועל הם חולו 8 ימים, כיצד הדבר ישנה את הטווח הבין-רביעוני? הסבירו.

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



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

4) הוספה גודל קבוע לכל תצפיות סדרת נתונים :

- א. תגדיל את הטוחה הבין-רביעוני.
- ב. תקטין את הטוחה הבין-רביעוני.
- ג. לא תנסה הטוחה הבין-רביעוני.
- ד. לא ניתן לדעת מה יקרה לטוחה הבין-רביעוני.

5) חושב הטוחה הבין-רביעוני עבור התפלגות מסויימת והתקבלת התוצאה אפס. לכן :

- א. לפחות 50% מהתצפיות זהות.
- ב. סטיית התקן היא אפס.
- ג. ההתפלגות היא סימטרית.
- ד. מצב זה כלל לא יכול.

- 6) סניף מס' 543 של בנק "רואה" בדק ל-80 לקוחות את מספר הפעמים שככל  
לקוח נכנס לסניף הבנק במשך שבוע. התוצאות שהתקבלו הן:  
 50 אנשים נכנסו 0 פעמים לסניף.  
 20 אנשים נכנסו פעם אחת לסניף.  
 5 אנשים נכנסו פעמיים לסניף.  
 5 אנשים נכנסו יותר מפעםיים.  
 מהו הטווח הבין-רבוני?  
 א. 60.  
 ב. 2.  
 ג. 50.  
 ד. 1.
- 7) התפלגות הציונים ב מבחון ווקסלר היא סימטרית בכך:  
 א. טווח הציונים הוא אפס.  
 ב. הטווח הבין-רבוני של הציונים אפס.  
 ג. סעיפים א ו-ב הם נכונים.  
 ד. אף סעיף אינו נכון.

**תשובות סופיות:**

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

# bijostatystyka dla medycyny

פרק 7 - סטטיסטיקה תיאורית - מדריך מיקום יחסי-ציוון תקן

תוכן העניינים

24 ..... 1. כללי .....

## סטטיסטיקה תיאורית – מדדי מיקום יחסי – ציון תקן:

**רקע:**

המטרה למדוד איך תצפית ממוקמת ביחס לשאר התצפיות בהתפלגות.

**ציון תקן:**

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

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

- ציון תקן חיובי אומר שההתצפית מעל הממוצע.
- ציון תקן שלילי אומר שההתצפית מתחת לממוצע.
- ציון תקן אפס אומר שההתצפית בדיק בממוצע.

**דוגמה (פתרון בהקלטה):**

במקומות העבודה מסוימים, ממוצע המשכורות הוא 8 אלף ₪, עם סטיית התקן של אלפיים ₪. באותו מקום העבודה ההשכלה הממוצעת של העובדים הנה 14 שנים, עם סטיית התקן של 1.5 שנים. עורך מרוויח במקום העבודה זה 11 אלף ₪ והשכלהו 16 שנים.  
מה ערך יותר, באופן יחסי, משכיל או משתכר?

**שאלות:**

**1)** תלמידי כיתה ח' ניגשו לבחן בלשון ולבוחן במתמטיקה.  
להלן התוצאות שהתקבלו :

המבחן	סטטיסט Takon	ממוצע
לשון	74	12
מתמטיקה	80	16

עודד קיבל : 68 בלשון ו-70 במתמטיקה.

- א. באיזה מקצוע עודד טוב יותר באופן יחסית לשכבה שלו?  
ב. איזה ציון עודד צריך לקבל במתמטיקה כדי שייהה שקול לציונו בלשון?

**2)** במבצע ליצור מצלבים לרכב בדקו במשך 40 ימים את התפוקה היומית (מספר מצלבים במאוט) ואת מספר הפעלים שעבדו באותו היום.  
להלן טבלה המסכםת את המידע שנאסף על שני המשתנים :

סטטיסט Takon	ממוצע	תפוקה	מספר פעילים
10	48	15	15
16	10	2	2

באחד הימים מתוך כלל הימים שנבדקו התפוקה הייתה 50 מאות מצלבים ובאותו היום עבדו 13 פעילים.  
מה יותר חריג באותו היום, ייחסית לשאר הימים שנבדקו : נתוני התפוקה או  
כמות הפעלים?  
א. התפוקה.  
ב. כמות הפעלים.  
ג. חריגים באותה מידה.  
ד. חסרים נתונים כדי לדעת זאת.

**3)** הגובה הממוצע של המתגייסים לצבאות הוא 175 סנטימטר עם סטטיסט Takon של 10 סנטימטר. המשקל הממוצע הוא 66 ק"ג עם סטטיסט Takon של 8 ק"ג.  
ערן המתגייס כshaw 180 ס"מ ומשקלנו 59 ק"ג.  
א. כמה ערן חריג יותר ביחס לשאר המתגייסים, גובהו או משקלו?  
ב. כמה ערן אמר לשקלן כדי שמשקלו יהיה שקול לגובהו?

**תשובות סופיות:**

- 1)** א. לשון.      ב. 72.  
**2)** ב'.  
**3)** א. משקל.      ב. 70.

# bijostatystyka dla medycyny

## פרק 8 - סטטיסטיקה תיאורית-אחסונים בטבלה בדידה

תוכן העניינים

- 26 ..... 1. כללי .....

## **סטטיסטיקה תיאורית – מדדי מיקום יחסי – אחווזוניים בטבלה בדידה:**

**רקע:**

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

**чисוב האחווזון מתוך נתוניים בטבלה שכיחיות בדידה:**

האחווזון הוא הערך שבו בפעם הראשונה השכיחות היחסית המצטברת (באחווזים) גדולה או שווה ל-  $p\%$ .

**דוגמה (פתרון בהקלטה):**

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

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

א. מצאו את האחווזון ה-25.

ב. מצאו את הערך ש-20% מהמקרים מעליו.

**שאלות:**

**1)** להלן התפלגות של משתנה קלשחו:

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

מצאו להתפלגות את :

- א. האחוזון ה-60.
- ב. המאון ה-40.
- ג. העשרון העליון.
- ד. הטווח בין הרבעונים.

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

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

חשבו את :

- א. העשרון התחתון.
- ב. האחוזון ה-30.
- ג. הערך ש-20% מהתצפית גודלות ממנו.
- ד. רביעון עליון.

**תשובות סופיות:**

- |       |       |       |       |
|-------|-------|-------|-------|
| .1 .ד | .3 .ג | .1 .ב | .2 .א |
| .4 .ד | .4 .ג | .2 .ב | .1 .א |

# bijostatystyka dla medycyny

## פרק 9 - סטטיסטיקה תיאורית- תרשימים קופסא

תוכן העניינים

1. כללי .....

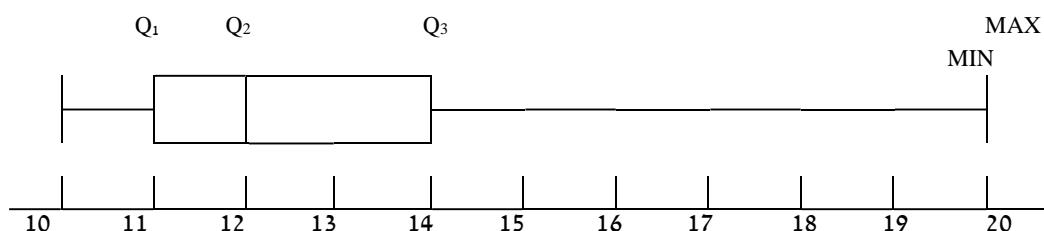
- 28 .....

## סטטיסטיקה תיאורית – תרשימים קופסה (Boxplot):

**רקע:**

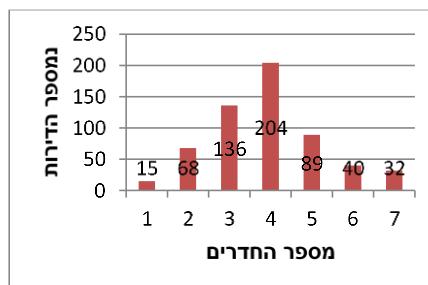
תרשימים קופסה הינו תרשימים שבuzzרתו ניתן לבחון :

- 1) את המרכז של ההתפלגות על ידי החציון ( $Q_2$ ).
- 2) את הפיזור של הנתונים (הטוחה והטוחה הבין רבוני).
- 3) את צורת ההתפלגות (סימטריה או אסימטריה שמאלית).



**שאלות:**

**1)** להלן התפלגות מספר החדרים לדירות שנבנו בשנת 2009 בעיר אשדוד :

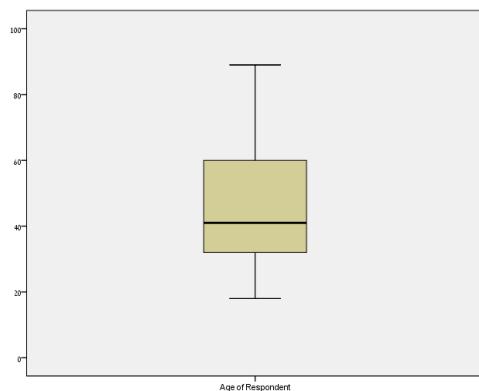


א. מצאו את החציון, הרבעון תחתון והרביעון עליון של ההתפלגות.

ב. שרטטו דיאגרמת קופסה להתפלגות.

ג. מה ניתן לומר על צורת ההתפלגות?

**2)** להלן דיאגרמת קופסה המתארת את התפלגות הגיל (בשנים) באוכלוסייה מסוימת :



א. מה הגיל החזionario?

ב. מה בערך טווח הגילאים?

ג. מה ניתן להגיד על צורת ההתפלגות?

**תשובות סופיות:**

- 1)** א. חציון : 4 , רבעון תחתון : 3 , רבעון עליון : 5 .  
 ב. ראה גרף מלא בסרטון וידאו. ג. כמעט סימטרית.  
**2)** א. חציון : 40 . ב. טווח : 70 . ג. התפלגות אסימטרית ימנית.

# bijostatystyka dla medycyny

## פרק 10 - סטטיסטיקה תיאורית- ניתוח פלטים

תוכן העניינים

- |         |               |
|---------|---------------|
| 30..... | 1. כללי ..... |
|---------|---------------|

## סטטיסטיקה תיאורית – ניתוח פלטים:

### שאלות:

**1)** להלן פلت על התפלגות הגילאים אוכלוסייה מסוימת :

	Statistic
Age of Respondent	Mean
	45.63
	Median
	41.00
	Variance
	317.140
	Std. Deviation
	a
	Minimum
	18
	Maximum
	b
	Range
	71
	Interquartile Range
	28

- א. מצאו את הערכים בטבלה המסומנים ב- a ו- b.  
 ב. נתנו שההתפלגות היא אסימטרית, האם היא נוטה ימינה או שמאלה?

**2)** להלן התפלגות ההשכלה של העובדים בחברת "מתאר" :

	Statistic
years of education	Mean
	?
	Median
	12.0000
	Variance
	?
	Std. Deviation
	2.54786
	Minimum
	?
	Maximum
	?
	Range
	?
	Interquartile Range
	?

מלאו את הערכים המסומנים בסימני שאלה.

**תשובות סופיות:**

- (1) א.  $a = 17.81$ ,  $b = 89$ .    ב. אסימטריה ימנית.
- (2) ממוצע: 11.909, שונות: 6.492, טווח: 10, טב"ר: 3.

# bijostatystyka dla medycyny

פרק 11 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

תוכן העניינים

32 ..... 1. כללי

## סטטיסטיקה תיאורית – טרנספורמציה לינארית:

---

**רקע:**

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

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

$$MO_y = a \cdot MO_x + b$$

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

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

**מדדי המרכז:**

$$R_y = |a| R_x$$

$$S_y = |a| S_x$$

$$S_y^2 = a^2 S_x^2$$

$$Y_p = a \cdot X_p + b$$

$$Z_Y = \frac{a}{|a|} Z_X$$

**שלבי העבודה:**

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

**דוגמה (פתרון בהקלטה):**

השכר הממוצע של עובדים הינו 9000 ש"ח וטוחה 6000 ש"ח. חשבו את המדדים הללו לאחר שהעלו את כל המשכורות ב-10% ולאחר כך קנסו אותן ב-100 ש"ח.

**שאלות:**

- 1) עברו סדרת נתוניים התקבל:  $70, S = 15, MO = 70, \bar{x} = 80$ .  
הוחלט להכפיל את כל התצפויות ב-4 ולהחסיר מההתוצאה 5.  
חשבו את המדדים הללו לאחר השינוי.
- 2) בחברה מסוימת השכר הממוצע הוא 40 ש' לשעה עם סטיית תקן של 5 ש' לשעה.  
הוחלט להעלות את כל המשכורות ב-10%, אך זה לא סיפק את העובדים ולכן  
הם קיבלו לאחר מכן תוספת של 2 ש' לשעה.  
מה הממוצע ומה הבדלים של השכר לשעה לאחר כל השינויים.
- 3) במחקר מסויים הציון החיצוני היה 73, טווח הציונים היה 40 נקודות והעשורן  
העליון היה הציון 87. כיוון שהציונים בבחינה היו נמוכים, המורה החליט לתת  
פקטור של 4 נקי לכל התלמידים.  
חשבו את המדדים לאחר הפקטור.
- 4) דגמו מקו ייוצר 50 קופסאות של גפרורים. בדקו בכל קופסה בה יש 40  
גפרורים את כמות הגפרורים הפגומים. התקבל שבממוצע יש 3 גפרורים  
פגומים בקופסה, עם סטיית תקן של 1.5 גפרורים.  
מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקינים בקופסה?
- 5) חברת בזק הציעה את ההצעה הבאה: שלושים שקלים דמי מנוי חודשיים  
קבועים וכן 10 אגרות לכל דקה של שיחת יוצאת. אדם בדק במשך שנה את  
דקות השיחות היוצאות שלו, וkiemל שבממוצע חודשי יש לו 600 דקות שיחות  
יוצאות עם שונות של 2500 דקות רבעות, כמו כן בחודש ינואר ציון התקן  
היה 2.  
חשבו את המדדים הללו עבור חשבון הטלפון החודשי של אותו אדם בשקלים  
אם היה משתמש בחיבור המוצעת לו על ידי בזק.
- 6) הוכחו שאם כל התצפויות בהתפלגות עברו טרנספורמציה לינארית:  
$$Y_i = a \cdot X_i + b$$
  
אזី הממוצע והבדלים של כל התצפויות לאחר הטרנספורמציה יהיו בהתאם:  
$$\bar{y} = a \cdot \bar{x} + b, S_y^2 = a^2 S_x^2$$

**תשובות סופיות:**

- (1) ממוצע: 315, סטיית תקן: 60, שכיח: .275.
- (2) ממוצע: 46, שונות: .30.25.
- (3) טוח: 40, חציון: 77, עשירון עליון: .91.
- (4) ממוצע: 37, סטיית תקן: 1.5.
- (5) ממוצע: 90, שונות: 25, חציון תקן: 2.
- (6)  $a^2 \cdot S_x^2$

# bijostatystyka dla medycyny

פרק 12 - סטטיסטיקה תיאורית שאלות אמריקאיות

תוכן העניינים

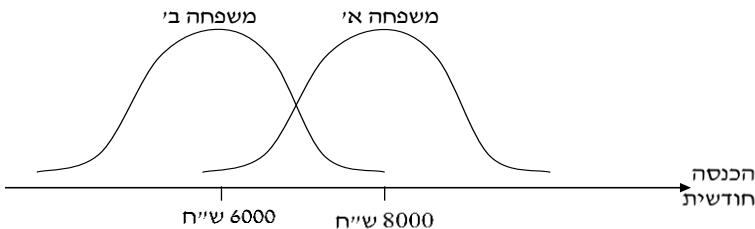
35 ..... 1. כללי .....

## סטטיסטיקה תיאורית – שאלות אמריקאיות:

**שאלות:**

**שאלות 3-1 מתייחסות לקטע הבא:**

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



**1)** לאיזו משפחה הכנסה שכיחה גבולה יותר?

- א. משפחה א'.
- ב. משפחה ב'.
- ג. לשתיهن אותה הכנסה שכיחה.
- ד. לא ניתן לדעת – אין מספיק נתונים.

**2)** באיזו משפחה הכנסה החזיונית שווה להכנסה הממוצעת?

- א. משפחה א'.
- ב. משפחה ב'.
- ג. לשתיهن הכנסה החזיונית שווה להכנסה הממוצעת.
- ד. לא ניתן לדעת – אין מספיק נתונים.

**3)** באיזו משפחה סטיית התקן של הכנסה החודשית גבולה יותר?

- א. משפחה א'.
- ב. משפחה ב'.
- ג. לשתיهن אותה סטיית התקן.
- ד. לא ניתן לדעת – אין מספיק נתונים.

**הנתונים הבאים מתייחסים לשאלות 6-4:**

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

$F(x)$	$x$
?	0
10	1
6	2
15	3
?	4
<b>50</b>	<b>סה"כ</b>

**4) השכיח של הנתונים הוא:**

- א. 0.
- ב. 1.5.
- ג. ישים שני שכיחים: 0 ו-3.
- ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של השכיח.

**5) חציון הנתונים הוא:**

- א. 2.
- ב. 1.5.
- ג. 25.5.
- ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של החציון.

**6) הטווח של הנתונים:**

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

**7) בהתפלגות אסימטרית ימנית של משתנה כמוותי רציף, הערך המתאים למאון ה-30, ציון התקן שלו הוא בהכרח:**

- א. שלילי.
- ב. חיובי.
- ג. אפס.
- ד. לא ניתן לדעת ללא הנתונים.

- 8)** סדרת נתונים סטטיסטיים מונה 10 תצפיות. נתון כי סדרת הנתונים סימטרית סביבה הממוצע. ממוצע הסדרה-40 ושונות הסדרה-100. בשלב מאוחר יותר נוסףו שתי תצפיות נוספות : 50 ו-30. השונות של 12 תצפיות :
- תקפן.
  - תגדל.
  - לא תשנה.
  - לא ניתן לחשב את השונות ללא ידיעת התצפיות.

**הנתונים הבאים מתיחסים לשאלות 10-9:**

בחברת "תיק" המשכורת הממוצעת היא 4,600 ש"ח וסטיית התקן של משכורת זו הינה 200 ש"ח. לאחר מוי"ם עם ועד עובדי הנהלה סוכם כי המשכורת תוכפל פי 1.5.

- 9)** מהי המשכורת הממוצעת החדשה (ב-ש"ח)?
- 2,300.
  - 6,900.
  - 4,650.
  - 4,600.
  - חסרים נתונים כדי לדעת.

- 10)** מהי סטיית התקן של המשכורת לאחר יישום המוי"ם לגבי השכר (ב-ש"ח)?
- .200.
  - .300.
  - .675.
  - לא ניתן לדעת.

- 11)** הוספה גודל קבוע לכל תצפיות סדרת נתונים :
- תגדיל את סטיית התקן.
  - תקטין את סטיית התקן.
  - לא תשנה את סטיית התקן.
  - לא ניתן לדעת.

**הנתונים הבאים מתייחסים לשאלות 13-12:**

להלן נתונים על ציוני תלמידים שנבחנו במועדים שוניםsstattistikah.com

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

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

- א. מייטב.
- ב. צבי.
- ג. לובה.
- ד. שרית.
- ה. סטף.

13) פנינה נבחנה עם סטף וציוון התקן שלה שווה לציוון התקן של שרית לכן ציונה הוא :

- א. 80.55
- ב. 65
- ג. 80
- ד. 81.66

**הנתונים הבאים מתייחסים לשאלות 17-14:**

בבדיקה פתע של משרד הבריאות במפעל שוקולד, נמצא ש :

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

14) מהו החציון של מספר הפגומים בקופסה :

- א. 1
- ב. 2
- ג. 4
- ד. לא ניתן לדעת.

15) מהו הרביעון התחתון של מספר הפגומים בקופסה?

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

**16)** מספר הפוגומים בקופסא הוא משתנה :

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

**17)** השכיח של מספר הפוגומים בקופסא :

- א. 63.
- ב. 1.
- ג. 200.
- ד. לא ניתן לדעת.

**18)** ביחס לציר המספרים, רוב הערכים בהתקלגות א-סימטרית ימנית נמצאים :

- א. בערכים הגבוהים.
- ב. בחלוקת זהה בין הערכים הגבוהים והנמוכים.
- ג. בערכים הנמוכים.
- ד. לא ניתן לדעת.
- ה. אף לא תשובה מהניל נכונה.

**19)** בוצע מחקר על מספר העובדים בחברות מזון לעומת חברות תקשורת.

החציון וה ממוצע בשתייהן שווה 8.

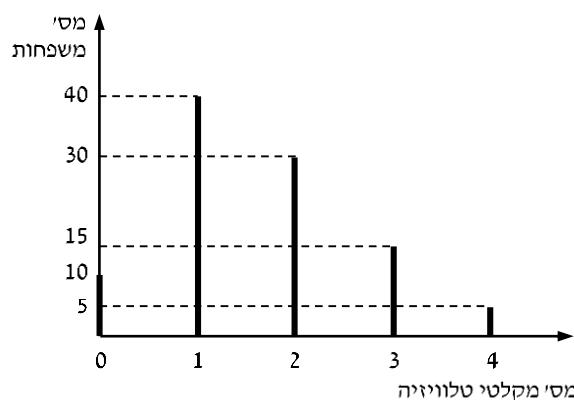
איזה מהטענות הבאות היא הנכונה והמלאת ביותר :

- א. השכיחות ב-2 החברות זהה אך שונה מ-8.
- ב. השכיח ב-2 החברות זהה אך לא ניתן לדעת מהו.
- ג. השכיח בשתי חברות הינו בהכרח 8.
- ד. שכיח בחברה אחת שונה מ-8 ובשנייה הוא 8.
- ה. אף תשובה אינה נכונה.

#### הנתונים הבאים מתיחסים לשאלות 20 עד 24:

נערך סקר על מספר מקלטי הטלוויזיה הנמצאים בבית.

תוצאות הסקר נתונות בדיאגרמת מקלות הבאה :



**(20)** המשנה הנחקר כאן הוא :

- א. משנה שמי.
- ב. משנה מסולם סדר.
- ג. משנה כמותי בדיד.
- ד. משנה כמותי רציף.

**(21)** הטווח של ההתפלגות הוא :

- .35
- .4
- .3
- .2

**(22)** ממוצע מספר מקלט טלוויזיה למשפחה הוא :

- .1.65
- .1.5
- .1
- .2

**(23)** השכיח של התפלגות זו היא :

- .40
- .1.5
- .1
- .2

**(24)** מסתבר שיש בין 2 ל-5 משפחות נוספות שאין להם מקלט טלוויזיה ויש לצרף את המשפחות הללו להתפלגות. כיצד הנתון זה ישפייע על סטיית התקן?

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

**תשובות סופיות:**

- |        |        |        |        |        |
|--------|--------|--------|--------|--------|
| 5) ב'  | 4) ג'  | 3) ג'  | 2) ג'  | 1) א'  |
| 10) ב' | 9) ג'  | 8) ג'  | 7) א'  | 6) ג'  |
| 15) א' | 14) ב' | 13) ד' | 12) ה' | 11) ג' |
| 20) ג' | 19) ח  | 18) ג' | 17) ב' | 16) ג' |
|        | 24) ב' | 23) ג' | 22) א' | 21) ב' |

# bijustisjka le-refoat shinim

## פרק 13 - מקדם המתאים של פירסון

### תוכן העניינים

42 .....	1. מקדם המתאים הלנארי (פירסון).....
53 .....	2. חישוב מקדם המתאים הלנארי (פירסון).....

## مقدم המתאים (מדד קשר) הلينארי ומובהקותו

### **מדד הקשר הلينארי (פירסון) – מבוא**

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

#### **דוגמה (פתרו בהקלטה) :**

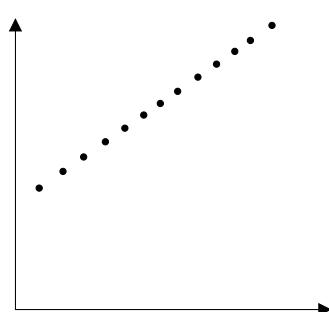
בבנייה 8 דירות בדקו לכל דירה את מספר החדרים שלה וכמו כן את מספר הנפותות הגורות בדירה. להלן התוצאות שהתקבלו :

מספר חדרים בדירה	מספר הנפותות בדירה
4	4
5	4

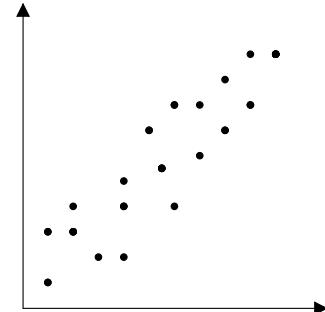
- 1) כמה תציפות ישן בדוגמה?
- 2) כמה משתנים ישנס בדוגמה, מי הם?
- 3) שרטטו לנواتים דיאגרמת פיזור.
- 4) מי המשתנה תלוי ומיהו המשתנה הבלתי תלוי?

### דיאגרמות פיזור לקשר בין משתנים וניתוחם

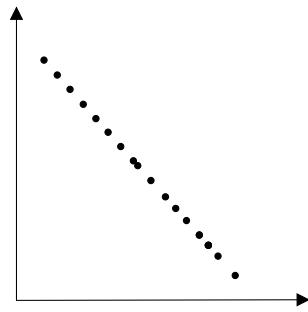
**קשר לנארី חיובי מלא**



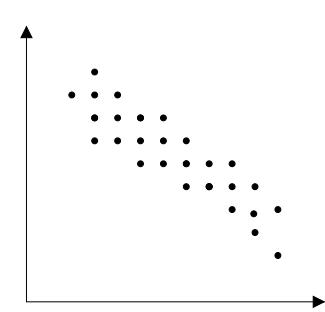
**קשר לנארី חיובי חלק**



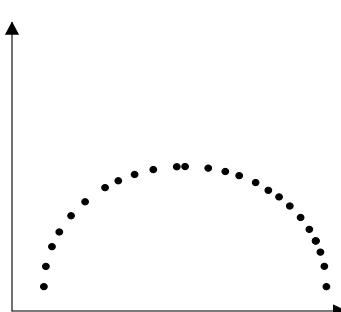
**קשר לינארី שלילי מלא**



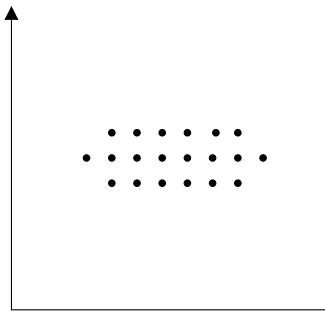
**קשר לינארី שלילי חלק**



**אין קשר לינארី**



**אין קשר**



#### משמעות מקדם המתאים:

כדי לבדוק עד כמה קיים קשר לינארី בין שני המשתנים ישנו מודד קשר שנקרא גם מקדם המתאים הליינארי הידוע גם בשם מקדם המתאים של פירסון. מקדם מתאים זה מקבל ערכים בין 1 ל-1.

-1

0

1

מقدم מותאם 1-או 1 אומר שקיים קשר לנארי מלא בין המשתנים שנייתן לבטא על ידי נוסחה של קו ישר:  $y = ax + b$ .

### **מתאים חיובי מלא (מقدم מותאם 1):**

קיים קשר לנארי מלא בו השיפוע  $a$  יהיה חיובי ואילו מותאם שלילי (מقدم מותאם-1) מלא אומר שקיים קשר לנארי מלא בו השיפוע  $a$  שלילי.

### **מתאים חיובי חלק:**

כל משתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחהلينארית שמקשרת את  $X$  ל- $Y$  באופן מוחלט ואילו מותאם שלילי חלקי אומר שככל משתנה אחד עולה לשני יש נטייה לרדת אבל לא קיימת נוסחה לינארית שמקשרת את  $X$  ל- $Y$  באופן מוחלט. ככל שמדובר המתאים הקרוב לאפס עצמת הקשר יותר חלשה וככל שהמדד רחוק יותר מהאפס העוצמה יותר חזקה. לsicום, מقدم המתאים בודק את עצמת הקשר הלינארי, ואת כיוון הקשר.

מقدم המתאים הלינארי אינו מושפע מייחדות המדידה. כל שינוי ביחסות המדידה של המשתנים, לא ישנה את מقدم המתאים.

מדד הקשר הלינארי באוכולוסייה, שנראה גם מقدم המתאים של פירסון או מדד הקשר של פירסון באוכולוסייה מסומן ב:  $r$  - פרמטר המאפיין את עצמת הקשר הלינארי באוכולוסייה וכיונו בין שני המשתנים הנחקרים. כאשר:

- מדד הקשר הלינארי במדגם שמהווה אומד לפרמטר  $r$ .

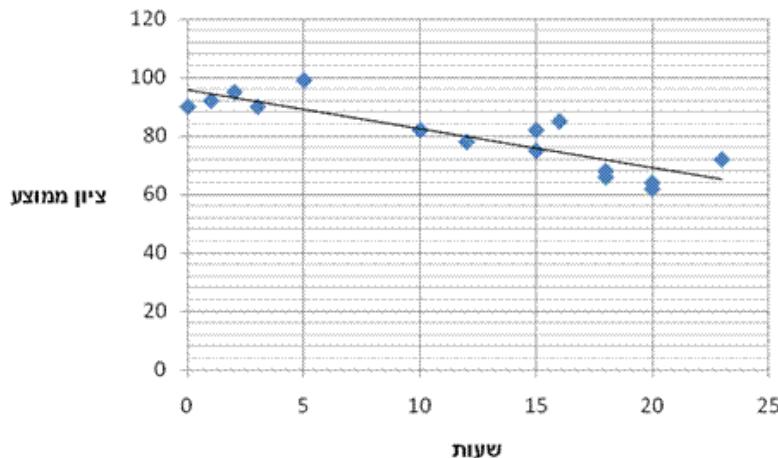
קיומו של מותאם בין שני משתנים אינו מצביע על סיבות בבחירה. למשל, אם נמצא מותאם חיובי בין כמות הסוכרזיות שאדם אוכל לבין משקל שלו אין זה אומר שהסיבה להשמנה היא הסוכרזית. מדד הקשר של פירסון הוא מדד קשר סימטרי,قولمر אם נחליף את  $X$  ב- $Y$  התוצאה תהיה זהה.

### **דוגמה (פתרון בהקלטה):**

- מה ניתן להגיד על מقدم המתאים של שני המשתנים על סמך דיאגרמת הפיזור שרטטנו?
- אם היינו משנים את הشرط כך שבציר האנכי היה המשתנה "מספר החדרים" ובציר האופקי היה "מספר הנפשות", האם הדבר היה משנה על מדד הקשר של פירסון?

## שאלות

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



- א. מיהו המשתנה הבלתי תלוי?
- ב. מה ניתן לומר על כיוון הקשר בין מספר שעות הבילוי השבועית לבין הציון הממוצע של הסמסטר? מה ניתן להגיד על עוצמת הקשר?

**2)** להלן טבלה המסכםת את מקדמי המתאים הליינארי בין ציוני מבחנים שונים שהתקבלו עבור תלמידים בכיתה מסוימת:

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

א. השלימו את מקדמי המתאים שמשמעותם בסימן שאלה בטבלה.

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

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

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

- א. מיהו המשתנה התלו依 ומיהו המשתנה הבלתי תלוי בדוגמה זו?
- ב. שרטטו דיאגרמת פיזור לנוטונים.
- ג. מה ניתן לומר על הקשר בין המשתנים במדגם?
- ד. מסתבר שבסופו של דבר נתנו פקטור של 5 נקודות לציון הסופי. כיצד הדבר היה משנה את מקדם המתאים של המדגם?

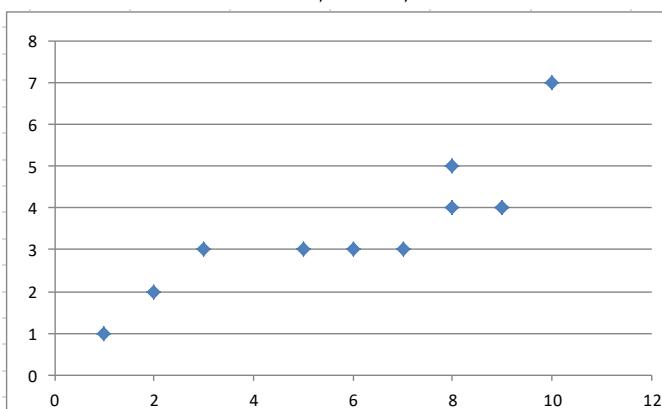
4) בتحقנה המטאורולוגית רצוי לבדוק את הקשר שבין הטמפרטורה במערכות כלזיות לכמות המשקעים במ"מ. הם אספו נתונים על 10 ימים במהלך חודש ינואר. המתאים שהתקבל היה 0.8.

א. השלימו את המשפט:

בחודש ינואר ככל שהטמפרטורה היומית נוטה לרדת, כך כמות המשקעים נוטה \_\_\_\_\_.

ב. הוחלט להעביר את הטמפרטורה למערכות פרנהייט על מנת שיוכלו להשוות אותה לנ נתונים מארה"ב. נוסחת המעבר היא  $F^0 = 32 + \frac{9}{5}C^0$ . כיצד הדבר ישפיע על מקדם המתאים בין הטמפרטורה במערכות פרנהייט לכמות המשקעים במ"מ?

5) להלן דיאגרמת פיזור המראה קשר בין שני משתנים:



א. השלימו: ניתן לראות קשר הוא לינארי \_\_\_\_\_ (מלאו חלקי) כיוון שהקשר הוא (חיובי ושלילי).

ב. השלימו: אם היינו מושפעים תצפית שערך ה- X שלה הוא 4 וערך ה- Y שלה הוא 7, מקדם המתאים של פירסון היה \_\_\_\_\_ (גדלו קטו לא משתנה).

שאלות רב ברירה (יש לבחור את התשובה הנכונה):

6) חוקר אקלים דגם כמה ימים בשנה ומדד את הטמפרטורה בטורונטו שבקנדזה ואת הטמפרטורה בסידני שבאוסטרליה באותו היום. הוא חישב ומצא מקדם מתאים שלילי בין הטמפרטורה היומית בטורונטו לבין הטמפרטורה היומית בסידני. משמעות מקדם המתאים השילי במדגם:

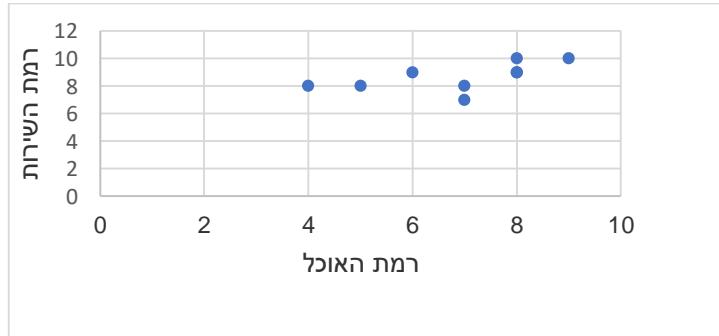
א. אין קשר בין הטמפרטורה בטורונטו לבין הטמפרטורה בסידני ביום שנדגמו.

ב. במדגם, רוב הטמפרטורות בטורונטו היו שליליות.

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

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

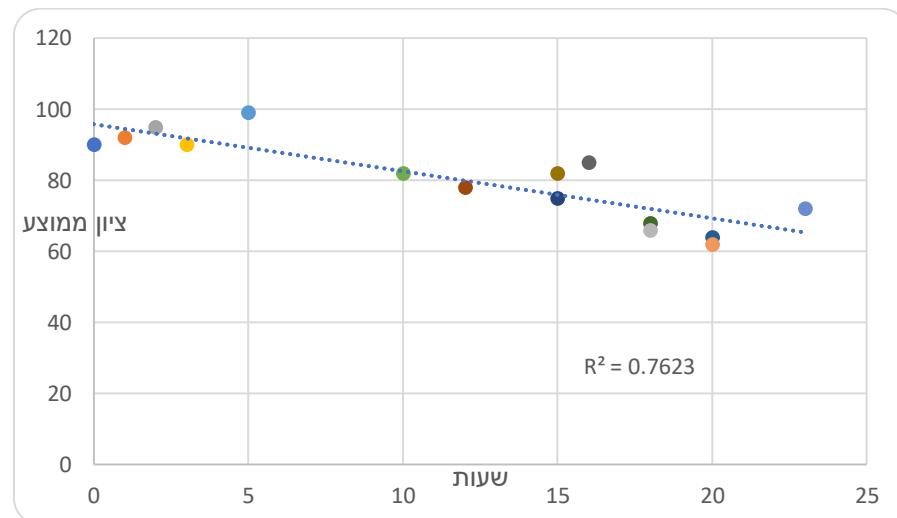
- 7) בסקר שביעות רצון שנערך בבית הקפה "fat لלחס" התבקו הלקוחות לדרג את מידת שביעות הרצון שלהם (בסולם 1-10) בשני נושאים: רמת האוכל ורמת השירות.



מה יהיה ערכו של מועד המתאים ( $r$ )?

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

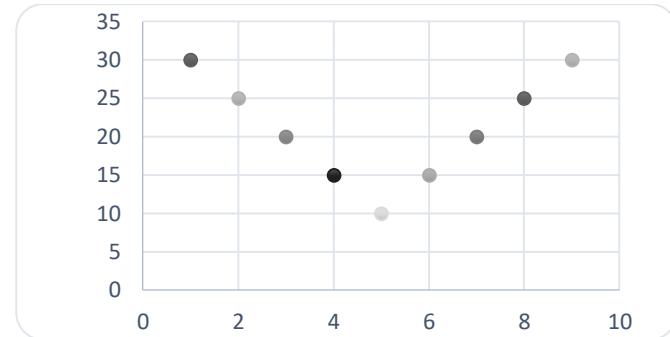
- 8) חוקר רצה לאפיין את הקשר בין מספר השעות בשבוע שסטודנט מקדיש לבילויים לבין הציון הממוצע שלו בסוף הסמסטר. לשם כך הוא אסף נתונים של 15 סטודנטים ויוצר דיאגרמת פיזור.



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

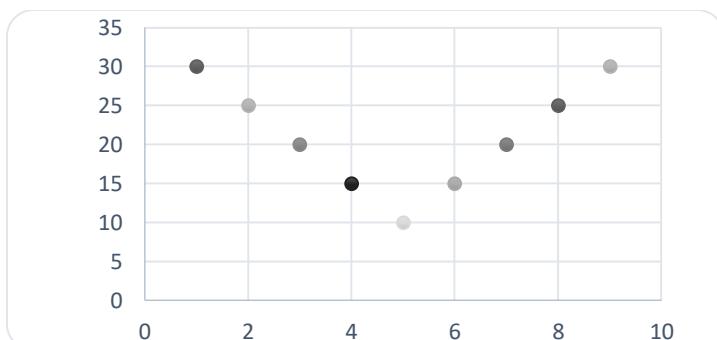
- א. ככל שմבלים יותר הציון נוטה לרדת.
- ב. אין קשר בין ساعات הבילוי לציון.
- ג. ככל שմבלים פחות הציון נוטה לרדת.
- ד. ככל שהציון נוטה לרדת הסטודנט מבלה פחות.

**9)** התרשימים הבא מתאר קשר בין שני משתנים, איזה מהמתאים הבאים הוא המתאים ביותר לתיאור הקשר בין שני המשתנים?



- א.  $r = 1$  היות ושני המשתנים יוצרים קוים ישרים.
- ב.  $r = 2$  היות ויש שני קוים בעלי קשר מושלם.
- ג.  $r = 0$  היות והקו יורד ולאחר מכן עולה באותו האופן.
- ד.  $r = \pm 1$  היות ויש קו עולה וגם קו יורד.

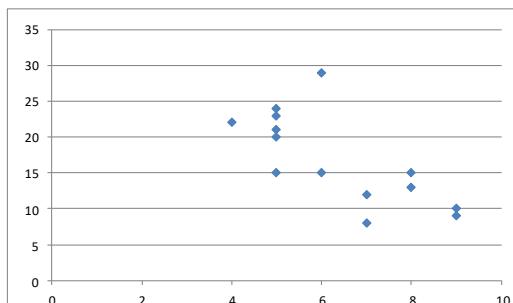
**10)** התרשימים הבא מתאר דיאגרמת פיזור.



איזה טענה נכונה?

- א. בתרשימים מוצג הקשר בין שני משתנים.
- ב. בתרשימים מוצג הקשר בין 9 משתנים.
- ג. בתרשימים מוצג הקשר בין 10 משתנים.
- ד. אין לדעת כמה משתנים מוצגים בתרשימים.

בגרף הבא מתוארת דיאגרמת פיזור של שני משתנים :



X - (משתנה בלתי תלוי בציר האופקי)  
ו- Y (משתנה תלוי).

במדגם התקבל  $r^2 = 0.52$ .

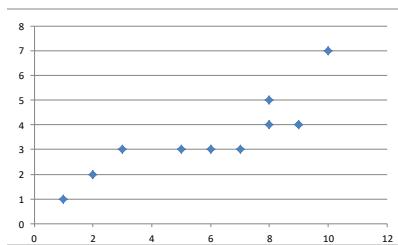
**11)** לאור הנתונים המופיעים בדיאגרמה, איזה מבחן הערכים הבאים מתאים להיות התוצאה של  $r$  ?

- א. -0.52
- ב. 0.72
- ג. -0.72
- ד. 0.52

**12)** אם מקדם המתאים בין שני משתנים הוא 1, אז :

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

**13)** להלן דיאגרמת פיזור :  
מה יהיה מקדם המתאים בין שני המשתנים ?



- א. 1
- ב. 0.85
- ג. 0.15
- ד. 0

**14)** בבדיקה קשר בין שני משתנים התקבל :  $r = -1$ .  
א. קיימת נוסחה ליניארית הקושרת בין כל התצפיות.  
ב. לא קיים קשר בין שני המשתנים.  
ג. ככל משתנה אחד נוטה לרדת גם לשני יש נטייה לרדת.  
ד. קיים קשר בין שני המשתנים, אך לא ניתן לדעת מאייזה סוג.

15) לפי הפטגס "רחוק מהעין, רחוק מהלב", יש קשר \_\_\_\_ בין קרבה פיזית לקרבה נפשית.

- א. חיובי
- ב. שלילי
- ג. אפסי
- ד. לא ניתן לדעת.

16) מבחן אמייר הינו מבחן מיוון באנגלית של המרכז הארצי לבחינות והערכתה. הציון המינימלי בבחינה הינו 150 והמаксימלי הינו 250. בקורס הכנה לבחן השתתפו 19 תלמידים. להלן הציונים שלהם על פי פلت שהתקבל:

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

יש להוסיף עמודה נוספת לצד עמודות הציונים שטראה לכל תלמיד כמה נקודות חסרות לו כדי להשלים לציוון המקסימלי בבחינה.

מה יהיה מקדם המתאים בין שתי העמודות (תלמיד, מקדם המתאים בין הציון לבין הנקודות החסרות)?

- א. -1
- ב. 1
- ג. -0.5
- ד. 0.5

17) מקדם המתאים בין שטחי דירה למחר שלחם חושב ונמצא 1.2. מה נובע לכך?

- א. ככל שהדירה גדולה יותר בשטחה כך היא יקרה יותר.
- ב. ככל שהדירה קטנה יותר בשטחה כך היא זולה יותר.
- ג. לא קיים קשר בין שטח הדירה למחר שדייה.
- ד. מצב כזה שמתואר הנתונים לא אפשרי.

18) אם ניקח 10 אנשים וונרשום לכל אדם את הגובה במטר וכמה כו' את הגובה בס"מ. מה יהיה מקדם המתאים בין גובה האדם במטר לגובה האדם בס"מ?

- א. 1
- ב. 0
- ג. -1
- ד. לא ניתן לדעת.

- 19) נמצא מתאימים חיובי בעוצמה גבוהה בין  $X$  – ציון בගראות בלשון ל  $Y$  – ציון בගראות במתמטיקה. אילו מהמשפטים הבאים נכון?
- א. ניתן לומר שאחת מהסיבות להבדלים שיש לסטודנטים במתמטיקה נובעים מההבדלים שיש להם בלשון.
  - ב. קיימת נוסחה של קו ישר שקשורה בין ציון בගראות במתמטיקה לציון בගראות בלשון.
  - ג. ללא יוצא מן הכלל, ניתן להגיד שככל תלמיד שמציל יותר מטלמיד אחר בלשונו גם יצליח יותר מאותו תלמיד במתמטיקה.
  - ד. אף אחד מהטענות שהוצעו אינה בהכרח נכונה.
- 20) עברו סדרה של תצפיות מדדו את  $X$  ואת  $Y$ . נמצא שעבור כל התצפיות שהערך של  $Y$  ירד הערך של  $X$  בהכרח ירד ללא יוצא מן הכלל. מוקדם המתאים של פירסון יהיה בהכרח:
- א. 1
  - ב. -1
  - ג. 0
  - ד. אף אחת מהתשובות.

### תשובות סופיות

- ב. הקשר חלקי, כיון הקשר שלילי.  
ב. ספורט ולשון.

- (1) א. שעות בילוי.  
(2) א. להלן טבלה:

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

- ב. ראה גרפ' בפתרון וידאו.  
ד. מקדם המתאים לא היה משתנה.  
ב. לא ישפיע על מקדם המתאים.  
ב. קטן.

- (3) א. ב"ת- מס' שעות התרגול, תלוי- ציון.  
ג. קשר לינארי חיובי חלקי.

- (4) א. עלות.  
(5) א. חלקי, חיובי.  
(6) ד'. ד'. א'. א'. ג'. (10)  
(11) ג'. א'. א'. (14) (15) (12) (13) ב'. ד'. ד'. א'. א'. (16) (17) (18) (19) ד'. ד'. א'. א'. (20)

## מדדי קשר – מדד הקשר הlienاري (פירסון) – רקע

המטרה היא לבדוק האם קיים קשר (קורלציה, מתאים) של קו ישר בין שני משתנים כמותיים. מבחינת סולמות המדיידה קשר בין סולמות רוחניים ומנה. בדרך כלל,  $X$  הוא המשתנה המסביר (הבלתי תלוי) ו- $Y$  הוא המשתנה המוסבר (התלויה).

**דוגמה:**

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

**שלב ראשון:** נהוג לשרטט דיאגרמת פיזור. זו דיאגרמה שנוננת אינדיקטיבית ויזואלית על טיב הקשר בין שני המשתנים.

**דוגמה:**

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

בבנייה של 5 דירות בדקנו את הנתונים הבאים :

$X$  - מספר חדרים בדירה.  $Y$  - מס' נפשות הגרות בדירה.

להלן התוצאות שהתקבלו :

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

**שלב שני:** מחשבים את מקדם המתאים (מדד הקשר) שבזוק עד כמה קיים קשרlienاري בין שני המשתנים. המدد (נקרא גם מדד הקשר של פירסון) מכמת את מה שנראה בשלב הראשון רק בעין.

המדד בודק את כיוון הקשר (חיובי או שלילי) ואת עוצמת הקשר (חלש עד חזק).

מקדם מתאים זה מקבל ערכאים בין 1- -1.

מקדם מתאים 1- או 1 אומר שקיים קשרlienاري מוחלט ומלא בין המשתנים שניינו לבטא על ידי הנוסחה :  $y = bx + a$ .

### מתאים חיובי מלא (מקדם מתאים 1):

קיים קשר לנארוי מלא בו השיפוע  $b$  יהיה חיובי ואילו מתאים שלילי מלא אומר斯基ים קשר לנארוי מלא בו השיפוע  $b$  שלילי (מקדם מתאים 1-).

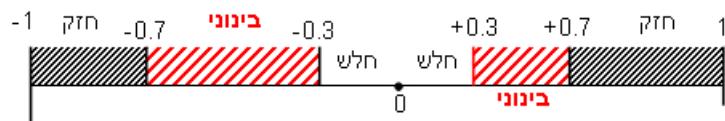
**מתאים חיובי חלקי:**

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

**מתאים שלילי חלקי:**

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

ככל שערך מקדם המתאים קרוב לאפס נאמר שעוצמת הקשר חלה יותר וככל שמקדם המתאים רחוק מהאפס נאמר שעוצמת הקשר חזקה יותר :



מקדם המתאים יסומן באות  $r$ .

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

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

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

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

$$\text{מקדם המתאים הلينארי: } r_{xy} = \frac{COV(x, y)}{S_x \cdot S_y}$$

## שאלות

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

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

א. שרטטו דיאגרמת פיזור לנ נתונים. מה ניתן להסיק מהדיאגרמה על טיב הקשר בין מספר החיסורים של תלמיד לציונו? מיהו המשתנה הבלתי תלוי ומיהו המשתנה התלווי?

ב. חשבו את מzdך הקשר של פירסון. האם התוצאה מתוישבת עם תשובה לסעיף א'?

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

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

**2)** במחקר רפואי רצוי לבדוק האם קיים קשר בין רמת ההורמון X בدم החולים לרמת ההורמון Y שלו. לצורך כך מדדו את רמת ההורמוניים ההלו עבור חמישה חולים. להלן התוצאות שהתקבלו:

א. מה הממוצע של כל רמת ההורמו?

ב. מהו מקדם המתאים בין ההורמוניים? ומה המשמעות ההתואמת?

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

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

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

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

א. חשב את מzdך הקשר הלינארי בין X ל- Y. מיהו המשתנה התלווי?

ב. מה המשמעות של התוצאה שקיבלת בסעיף א'?

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

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

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

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

חשבו את מודד הקשר הlienari בין  $X$  ל-  $Y$ .

5) במוסד אקדמי ציון ההתאמה מחושב כך: מכפילים את הציון הממוצע בוגרות ב-3 ומחיתנים 2 נקודות. ידוע שעבור 40 מועמדים סטיטית התקן של ממוצע הציון בוגרות הייתה 2. מה מגדם המתאים בין ציון ההתאמה לציון הממוצע בוגרות שלהם?

- 6)
- הלו רשימה טענות, לגבי כל טענה קבעו נכון/לא נכון ונמקו.
  - א. מתוויך דירות המיר מחירי דירות מדולר לשקל. נניח שдолר אחד הוא 3.5 נק. אם מתוויך הדירות יחשב את מודד הקשר של פירסון בין מחיר הדירה בשקלים למחיר הדירה בדולרים הוא יקבל 1.
  - ב. לסדרה של נתונים התקבל  $S_x = S_y = 1$ ,  $\bar{X} = \bar{Y}$ . לכן, מודד הקשר של פירסון יהיה 1.
  - ג. אם השונות המשותפת של  $X$  ושל  $Y$  הינה 0 אז בהכרח גם מגדם המתאים של פירסון יהיה 0.

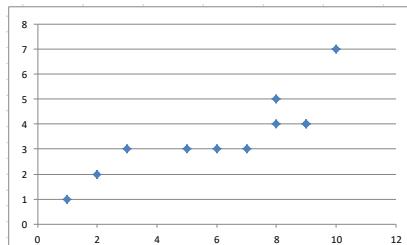
#### שאלות רב-ברירה:

- 7) נמצא שקיים מגדם מתאים שלילי בין הציון בעברית לחשבון בבחינה לכן:
- א. הדבר מעיד שהציונים בכתה היו שליליים.
  - ב. ככל שהציון של תלמיד יורך 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 ונבדק ביום מסוים המחיר שלהם בדולרים והמחיר שלהם בש"ח (באותנו היום ערך הדולר היה-2.4₪). מהו מקדם המתאים בין המחיר בדולר למחיר בש"ח?

- א. 1
- ב. 0
- ג. 4.2
- ד. לא ניתן לדעת.

9) להלן דיאגרמת פיזור:  
מה יהיה מקדם המתאים בין שני המשתנים?



- א. 1
- ב. 0.85
- ג. 0.15
- ד. 0

### תשובות סופיות

- 1) א. משתנה תלוי : ציון, משתנה ב"ת : מס' חיסורים. ראה דיאגרמה בוידאו. ניתן להסיק שקיימים קשרلينי שלילי וחלייני בין מספר החיסורים לציון התלמיד.  
 ב.  $r_{xy} = -0.9325$ .  
 ג. הקשר ישאר לינארי שלילי חלייני אך עצמתו תחלש.
- 2) א.  $r_{xy} = 0.96$       ב.  $\bar{x} = 15.4$  ,  $\bar{y} = 16$  .  
 3) א. 0.8  
 4) ב. 0.8  
 5) ג. 1.  
 6) א. נכון.  
 ב. לא נכון.  
 ג. נכון.
- 7) א. ג'.  
 8) א'.  
 9) ב'.

# bijustisjka le-refoat shinim

## פרק 14 - יסודות ההסתברות

תוכן העניינים

1. כללי .....

- 58 .....

## הגדירות יסודיות:

**רקע:**

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

**מרחב מדגם:** כלל התוצאות האפשרות בניסוי המקרי. לדוגמה, בהטלה קובייה:  $\{1, 2, 3, 4, 5, 6\}$ , או: מזג האויר בעוד שבועיים: {נאה, שרבי, מושלג, גשם, מעונן, חלקית, אביך}.

**מאורע:** תת קבוצה מתוק מרחב המדגם. מסומן באותיות: A, B, C. בהטלה קובייה למשל, המאורע 'לקבל לפחות 5' יסומן:  $A = \{5, 6\}$ . המאורע 'לקבל תוצאה זוגית' יסומן:  $B = \{2, 4, 6\}$ .

**גודל מרחב המדגם:** מספר התוצאות האפשרות למרחב המדגם. בהטלה קובייה למשל נקבע:  $|\Omega| = 6$ .

**גודל המאורע:** מספר התוצאות האפשרות במאורע עצמו. למשל, בהטלה הקובייה האירועים הקודמים יסומנו:  $|A| = 2$ ,  $|B| = 3$ .

**מאורע משלים:** מאורע המכיל את כל התוצאות האפשרות למרחב המדגם פרט לתוצאות במאורע אותו הוא משלים. למשל, בהטלה הקובייה:  $\bar{A} = \{1, 2, 3, 4\}$ ,  $\bar{B} = \{1, 3, 5\}$ .

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

**הסתברות במרחב מודגם אחיד:** במרחב מודגם אחיד הסיכוי למאורע יהיה :

$$P(A) = \frac{|A|}{|\Omega|}$$

דוגמה : מה הסיכוי בהטלה קובייה לקבל לפחות 5 ?

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2}{6}$$

דוגמה : מה הסיכוי בהטלה קובייה לקבל תוצאה זוגית ?

$$P(B) = \frac{|B|}{|\Omega|} = \frac{3}{6}$$

**הסתברות במרחב לא אחיד:** תחושב לפי השכיחות היחסית :

$$\frac{f}{n}$$

דוגמה :

להלן התפלגות הציונים בכיתה מסוימת :

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

מה ההסתברות שתלמיד אקרי שנבחר בכיתה קיבל את הציון 8 ?

$$\frac{f}{n} = \frac{5}{25} = 0.2$$

מה ההסתברות שתלמיד אקרי שנבחר בכיתה יכשל ?

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

**הסתברות למאורע משלים :** הסתברות לקבלת המשלים של המאורע ביחס למרחב המודגם :

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

להיות מחושב לפי הסיכוי להכשל :

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

**שאלות:**

- 1)** מהאותיות E, F ו-G יש ליצור מילה בת 2 אותיות, לא בהכרח בת משמעות.  
 א. הרכיבו את כל המילים האפשריות.  
 ב. רשמו את המקרים למאורע:  
 .i. במילה נמצאת האות E.  
 .ii. במילה האותיות שונות.  
 ג. רשמו את המקרים למאורע  $\bar{A}$ .
- 2)** מטילים זוג קוביות.  
 א. רשמו את מרחב המדגם של הניסוי. האם מרחב המדגם אחיד?  
 ב. רשמו את כל האפשרויות לאיורים הבאים:  
 .i. סכום התוצאות 7.  
 .ii. מכפלת התוצאות 12.  
 ג. חשבו את הסיכויים לאיורים שהוגדרו בסעיף ב'.
- 3)** נבחר באקראי ספרה מבין הספרות 0-9.  
 א. מה ההסתברות שהספרה שנבחרה גדולה מ-5?  
 ב. מה ההסתברות שהספרה שנבחרה היא לכל היותר 3?  
 ג. מה ההסתברות שהספרה שנבחרה היא אי זוגית?
- 4)** להלן התפלגות מספר מקלט טלוויזיה עבור כל משפחה ביישוב מסוים:

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

- נבחרה משפחה באקראי מהיישוב.  
 א. מה ההסתברות שאין מקלטים למשפחה?  
 ב. מה ההסתברות שיש מקלטים למשפחה?  
 ג. מה ההסתברות שיש לפחות 3 מקלטים למשפחה?

- 5)** להלן התפלגות מספר המכוניות למשפחה ביישוב "עדן":

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

- נבחרה משפחה אקראיית מן היישוב.  
 א. מה ההסתברות שאין לה מכוניות?  
 ב. מה ההסתברות שבבעלות המשפחה לפחות 3 מכוניות?  
 ג. מה הסיכוי שבבעלותה פחות מ-3 מכוניות?

- 6) נתיל מטבע רגיל 3 פעמים. בצד אחד של המטבע מוטבע עץ ובצד השני פלי.  
 א. רשמו את מרחב המדגמים של הניסוי. האם מרחב המדגם הוא אחיד?  
 ב. רשמו את כל האפשרויות לאיורים הבאים:  
 .i. התקבל פעם אחת עץ.  
 .ii. התקבל לפחות פלי אחד.  
 ג. מהו המאורע המשלימים ל-D?  
 ד. חשבו את הסיכויים לאיורים שהוגדרו בסעיפים ב-ג.

### תשובות סופיות:

$$\text{.} \Omega = \{EE, EF, EG, FE, FF, FG, GE, GF, GG\} \quad (1)$$

$$\text{.} A = \{EE, EF, EG, FE, GE\}, B \{EF, EG, FE, FG, GE, GF\}$$

$$\text{.} \bar{A} = \{FF, FG, GF, GG\}$$

$$\text{.} \Omega = \begin{Bmatrix} (1,1) & (2,1) & (3,1) & (5,1) & (4,1) & (6,1) \\ (1,2) & (2,2) & (3,2) & (4,2) & (5,2) & (6,2) \\ (1,3) & (2,3) & (3,3) & (4,3) & (5,3) & (6,3) \\ (1,4) & (2,4) & (3,4) & (4,4) & (5,4) & (6,4) \\ (1,5) & (2,5) & (3,5) & (4,5) & (5,5) & (6,5) \\ (1,6) & (2,6) & (3,6) & (4,6) & (5,6) & (6,6) \end{Bmatrix} \quad (2)$$

$$\text{.} A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}, C = \{(2,6), (3,4), (4,3), (6,2)\}$$

$$\text{.} \frac{1}{9} \text{ הסיכוי ל-B-A :} \quad \frac{1}{6} : \text{A}$$

$$\text{.} 0.5 \quad \text{.} 0.4 \quad \text{.} 0.4 \quad (3)$$

$$\text{.} 0.32 \quad \text{.} 0.78 \quad \text{.} 0.22 \quad (4)$$

$$\text{.} 0.8 \quad \text{.} 0.2 \quad \text{.} 0.1 \quad (5)$$

$$\text{.} \Omega = \{PPP, PPE, PEP, EPP, PEE, EPE, EEP, EEE\} \quad (6)$$

$$\text{.} A = \{PPE, PEP, EPP\}, D = \{PPP, PPE, PEP, EPP, PEE, EPE, EEP\}$$

$$\text{.} \bar{D} = \{EEE\}$$

$$\text{.} \frac{1}{8} \quad (7)$$

## bijustisjka le-refoat shinim

פרק 15 - פועלות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכלולים

תוכן העניינים

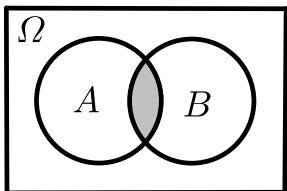
1. כללי .....

62 .....

## פעולות בין מאורעות (חיתוך ואיחוד) – מאורעות זרים ומכילים:

**רעיון:**

**פעולה חיתוך:**

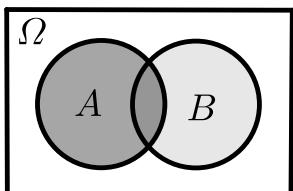


נותנת את המשותף בין המאורעות הנחטכים.

חיתוך בין המאורע  $A$  למאורע  $B$  יסומן כך:  $A \cap B$ .  
מדובר בתוצאות שנמצאות ב- $A$  וגם ב- $B$ .

**דוגמה:**

.  $A = \{5, 6\}$  בהטלת קובייה, למשל, האפשרויות לקבל לפחות 5 הן:  
.  $B = \{2, 4, 6\}$  האפשרויות לקבל תוצאה זוגית הן:  
.  $A \cap B = \{6\}$  החיתוך שביניהם הוא:



נותנת את כל האפשרויות שנמצאות לפחות באחת מהמאורעות, ומסומנת:  $A \cup B$ .

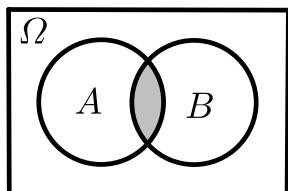
הפעולה נותנת את אשר נמצא ב- $A$  או  $B$ .  
כלומר, לפחות אחד מהמאורעות קורה.

**דוגמה:**

.  $A = \{5, 6\}$  בהטלת קובייה האפשרויות לקבל לפחות 5 הן:  
.  $B = \{2, 4, 6\}$  האפשרויות לקבל תוצאה זוגית הן:  
.  $A \cup B = \{2, 4, 5, 6\}$  האפשרויות לקבל לפחות 5 וגם תוצאה זוגית הן:

**דוגמה (הפתרון נמצא בהקלטה):**  
סטודנטים ניגש בסיסטר לשני מבחנים. מבחן בסטטיסטיקה ו מבחן בכלכלת. ההסתברות שלו לעبور את המבחן בסטטיסטיקה הוא 0.9, ההסתברות שלו לעبور את המבחן בכלכלת הוא 0.8 וההסתברות לעبور את המבחן בסטטיסטיקה ובכלכלת היא 0.75.  
מה ההסתברות שלו לעبور את המבחן בסטטיסטיקה בלבד?  
מה ההסתברות שלו להיכשל בשני המבחנים?  
מה ההסתברות לעبور לפחות מבחן אחד?

**נוסחת החיבור לשני מאורעות:**



ההסתברות של איחוד מאורעות תחושב ע"י הקשר הבא :

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

**חוקי דה מורגן לשני מאורעות:**

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

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

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

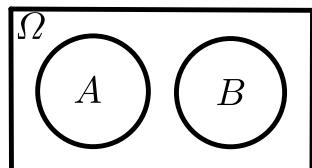
$$P(A \cup B) = 1 - P(\overline{A} \cap \overline{B})$$

**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:**



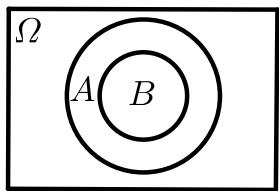
מאורעות זרים הם כאשר אין להם אף איבר משותף :  $A \cap B = \emptyset$ . כמובן, הם לא יכולים להתרחש בו זמינות.

ההסתברות של חיתוך המאורעות היא אפס :  $P(A \cap B) = 0$ .

ההסתברות של איחוד המאורעות תחושב :  $P(A \cup B) = P(A) + P(B)$

**דוגמה :**

בהתלט קובייה, האפשרויות לקבל לפחות 5 הן :  $A = \{5, 6\}$  והאפשרות לקבל 3 היא :  $B = \{3\}$ , ולכן החיתוך ביניהם הוא אפס, כאמור :  $A \cap B = \emptyset$ .

**מאורעות מוכליים:**

נתונים שני מאורעות  $A$  ו-  $B$ , השונים מאפס.  
 נאמר שהמאורע  $B$  מוכל במאורע  $A$  אם כל איברי  
 המאורע  $B$  כלולים במאורע  $A$  ונרשום:  $B \subset A$ .  
 מאורע  $A$  מכיל את מאורע  $B$  כל התוצאות שנמצאות ב-  $B$   
 מוכלות בתחום מאורע  $A$ .

קשר זה מסומן באופן הבא:  $B \subset A$

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

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

למשל:  
 $A = \{2, 4, 6\}$   
 $B = \{2, 4\}$

**שאלות:**

- 1)** מהאותיות  $E$ ,  $F$  ו-  $G$  יוצרים מילה בת 2 אותיות – לא בהכרח בת משמעות. נגידר את המאורעות הבאים :  
 A - במילה נמצאת האות  $E$ .  
 B - במילה אותיות שונות.  
 א. רשמו את כל האפשרויות לחיתוך  $A$  עם  $B$ .  
 ב. רשמו את כל האפשרויות לאיחוד של  $A$  עם  $B$ .
- 2)** תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ובבחן בסטטיסטיקה. נגידר את המאורעות הבאים :  
 A - עברו את המבחן בסטטיסטיקה.  
 B - עברו את המבחן בכלכלה.  
 היעזרו בפעולות חיתוך, איחוד ומשלים בלבד כדי להגידר 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 בכלכללה.
- 3)** נתבקשתם לבחור ספרה באקראי. נגידר את  $A$  להיות הספרה שנבחרה היא זוגית. נגידר את  $B$  להיות הספרה שנבחרה קטנה מ-5.  
 א. רשמו את כל התוצאות למאורעות הבאים :  
 $A \cup B$  ,  $A \cap B$  ,  $\bar{B}$  ,  $B$  ,  $A$   
 ב. חשבו את ההסתברויות לכל המאורעות מהסעיף הקודם.
- 4)** נסמן ב-  $\Omega$  את מרחב המדגמים וב-  $\phi$  קבוצה ריקה.  
 נתון כי  $A$  הינו מאורע בתוך מרחב המדגמים.  
 להלן מוגדרים מאורעות שפטرونום הוא  $\Omega$  או  $\phi$  או  $A$ .  
 קבעו עבור כל מאורע מה הפתרון שלו :  
 $A \cup \bar{A}$  ,  $\bar{\phi}$  ,  $A \cap \bar{A}$  ,  $A \cup \Omega$  ,  $A \cap \Omega$  ,  $A \cup \phi$  ,  $A \cap \phi$  ,  $\bar{A}$

**5) הוגדרו המאורעות הבאים:**

A - אדם שגובהו מעל 1.7 מטר

B - אדם שגובהו מתחת ל-1.8 מטר.

קבעו את גובהם של האנשים הבאים:

- . A  $\cap$  B
- . A  $\cup$  B
- .  $\bar{A} \cap B$
- .  $\bar{A} \cup \bar{B}$
- .  $\bar{A} =$
- . h.

**6) נגדיר את המאורעות הבאים:**

A - אדם דובר עברית.

B - אדם דובר ערבית.

C - אדם דובר אנגלית.

השתמשו 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 שדוברים שתי שפות במדויק (מהשפות הנ"ל).

**7)** שני מפלגות רצות לכינסת הבאה. מפלגת "גדר" תעבור את אחוז החסימה בהסתברות של 0.08 ומפלגת "עתיד" תעבור את אחוז החסימה בהסתברות של 0.20. בהסתברות של 76% שני המפלגות לא תעבורנה את אחוז החסימה.

א. מה ההסתברות שלפחות אחת מהמפלגות תעבור את אחוז החסימה?

ב. מה ההסתברות שתשתי המפלגות תעבורנה את אחוז החסימה?

ג. מה ההסתברות שרק מפלגת "עתיד" תעבור את אחוז החסימה?

**8)** במקום העבודה מסויים 40% מהעובדים הם גברים. כמו כן, 20% מהעובדים הם אקדמיים. 10% מהעובדים הין נשים אקדמיות.

א. איזה אחוז מהעובדים הם גברים אקדמיים?

ב. איזה אחוז מהעובדים הם גברים או אקדמיים?

ג. איזה אחוז מהעובדים הם נשים לא אקדמיות?

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 עלתה ביום מסוים.

חשבו את ההסתברויות הבאות לגבי שתי המניות הללו ביום מסוים :

א. שתי המניות עלנה.

ב. שאף אחת מהמניות לא עלנה.

ג. שמניה A בלבד עלה.

10) מטילים זוג קופיות, אדומה ושחורה. נגידר את המאורעות הבאים :

A - בקובייה האדומה התקבלה התוצאה 4 ובשחורה 2.

B - סכום התוצאות משתי הקופיות הוא 6.

C - מכפלת התוצאות בשתי הקופיות היא 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) עבר המאורעות A ו- B ידועות ההסתברויות הבאות :  $P(A)=0.6$

$$P(\bar{A} \cap \bar{B})=0.1, P(B)=0.3$$

א. האם A ו- B מאורעות זרים?

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

12) מטבח הווטל פעמיים. נגידר 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 הkartiyim ריקים. אדם קיבל כרטיס אקראי.

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

ב. מה ההסתברות לא לזכות בפרס?

**14)** נתון כי :  $P(A) = 0.3$  ,  $P(B) = 0.25$  ,  $P(A \cup B) = 0.49$

א. חשבו את הסיכוי ל-  $P(A \cap B)$

ב. האם  $A$  ו-  $B$  מאורעות זרים?

ג. מה ההסתברות שرك  $A$  יקרה או שرك  $B$  יקרה?

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

מה הסיכוי למאורע  $A$  ומה ההסתברות למאורע  $B$  ?

### תשובות סופיות:

$$\text{א. } A \cap B = \{EG, EF, FE, GE\} \quad (1)$$

$$\text{ב. } A \cup B = \{EG, EF, EE, FE, GE, EG, GF\}$$

$$\text{ג. } \bar{B} \quad \text{ד. } \bar{A} \cap \bar{B} \quad \text{ה. } A \cup B \quad \text{ט. } A \cap B \quad \text{ו. } A \cap \bar{B} \quad \text{ז. } B \cap \bar{A} \quad \text{א. } \bar{B} \quad (2)$$

$$\text{, } \bar{B} = 5, 6, 7, 8, 9, B = 0, 1, 2, 3, 4, A = 0, 2, 4, 6, 8 \quad (3)$$

$$\text{. } A \cup B = 0, 2, 4, 6, 8, 1, 3, A \cap B = 0, 2, 4$$

$$\text{. } P(A \cup B) = 0.7, P(A \cap B) = 0.3, P(\bar{B}) = 0.5, P(B) = 0.5, P(A) = 0.5 \quad \text{ב.}$$

$$\text{, } A \cup \Omega = \Omega, A \cap \Omega = A, A \cup \phi = A, A \cap \phi = \phi, \bar{\bar{A}} = A \quad (4)$$

$$\text{. } A \cup \bar{A} = \Omega, \bar{\phi} = \Omega, A \cap \bar{A} = \phi$$

$$\text{ב. } A \cup B : \text{ כל גובה אפשרי} \quad \text{א. } \text{גובה בין } 1.7 \text{ ל- } 1.8 \quad (5)$$

$$\text{ג. } \bar{A} \cup \bar{B} \quad \text{ט. } \text{גובה לכל היוטר } 1.7 \text{ או לפחות } 1.8.$$

$$\text{ה. } \bar{A} : \text{גובה מעל } 1.7 \quad A = \bar{\bar{A}}$$

$$\text{. } A \cup B \cup C \quad \text{ג. } \text{. } A \cap \bar{B} \cap \bar{C} \quad \text{ב. } \text{. } A \cap B \cap C \quad \text{א. } \bar{C} \quad (6)$$

$$\text{. } (A \cap B \cap \bar{C}) \cup (B \cap C \cap \bar{A}) \cup (A \cap C \cap \bar{B}) \quad \text{ה. } \text{. } \bar{C} \quad \text{ט. }$$

$$\text{. } P(B \cap \bar{A}) = 0.16 \quad \text{ג. } \text{. } P(A \cap B) = 0.04 \quad \text{ב. } \text{. } P(A \cup B) = 0.24 \quad \text{א. } (7)$$

$$\text{. } P(\bar{A} \cap \bar{B}) = 50\% \quad \text{ג. } \text{. } P(A \cup B) = 50\% \quad \text{ב. } \text{. } P(A \cap B) = 10\% \quad \text{א. } (8)$$

$$\text{. } P(A \cup \bar{B}) = 0.3 \quad \text{ג. } \text{. } P(\bar{A} \cap \bar{B}) = 0.3 \quad \text{ב. } \text{. } P(A \cap B) = 0.2 \quad \text{א. } (9)$$

$$\text{. } \text{ט. לא.} \quad \text{ג. כן.} \quad \text{ב. כן.} \quad \text{א. לא.} \quad (10)$$

$$\text{. } P(\bar{A} \cap B) = 0.3 \quad \text{ב. כן.} \quad \text{א. כן.} \quad (11)$$

(12) הטענה הנכונה היא ג'.

$$\text{. } 0.95 \quad \text{ב. } 0.05 \quad \text{א. } (13)$$

$$\text{. } P((A \cap \bar{B}) \cup (B \cap \bar{A})) = 0.43 \quad \text{ג. } \text{. } P(A \cap B) = 0.06 \quad \text{א. } (14)$$

$$\text{. } P(B) = \frac{1}{5}, P(A) = \frac{2}{5} \quad (15)$$

# bijustisjka le-refoat shinim

פרק 16 - הסתברות מותנית-במרחב מודגש אחד

תוכן העניינים

1. כללי .....

70 .....

## הסתברות מותנית – במרחב מדגם אחד:

**רקע:**

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

הסתברות של  $A$  בהינתן  $B$ -  $B$  כבר קרה :

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

**דוגמה (פתרון בהקלטה):**

נטיל קופייה.

**נגיד :**

$A$  - התוצאה זוגית.

$B$  - התוצאה גזולה מ-3.

נרצה לחשב את :

**שאלות:**

1) נבחרה ספרה זוגית באקראי. מה הסיכוי שהספרה גדולה מ-6?

2) יוסי הטיל קובייה. מה הסיכוי שקיבל את התוצאה 4, אם ידוע שההתוצאה שהתקבלת זוגית?

3) הוטלו צמדקוביות. נגיד:   
 A - סכום התוצאות בשתי ההצלות הינו 7.  
 B - מכפלת התוצאות 12.  
 חשבו את  $P(A|B)$ .

4) מطبع הוטל פעמיים. ידוע שהתקבל לכל היוטר ראש אחד, מה הסיכוי שהתקבלו שני ראשים?

5) זוג קובייות הוטלו והתקבלו שההתוצאות זהות. מה הסיכוי שלפחות אחת התוצאות 5?

6) זוג קובייות הוטלו והתקבל לפחות פעמיים אחד 4. מה הסיכוי שאחת התוצאות 5?

7) נבחרה משפחה בת שני ילדים, מהם אחד הוא בן. מה ההסתברות שבמשפחה שני בני בקרבת הילדים?

8) נבחרה משפחה בת שלושה ילדים, ונתנו שהילד האמצעי בן. מה הסיכוי שיש בנות בקרבת הילדים?

**תשובות סופיות:**

.0.2 **(1)**

. $\frac{1}{3}$  **(2)**

.0.5 **(3)**

.0 **(4)**

. $\frac{1}{6}$  **(5)**

. $\frac{2}{11}$  **(6)**

. $\frac{1}{3}$  **(7)**

. $\frac{3}{4}$  **(8)**

# bijustisjka le-refoat shinim

פרק 17 - הסתברות מותנית - מרחב לא אחד

תוכן העניינים

1. כללי .....

73 .....

## הסתברות מותנית – מרחב לא אחד:

**רקע:**

.  $P(A|B) = \frac{P(A \cap B)}{P(B)}$  הסיכוי שמאורע  $A$  יתרחש, בהינתן שמאורע  $B$  כבר קרה :

במונח : הסיכוי לחיותך של שני המאורעות, זה הנשאל וזה הנتوן שהתרחש.

במקרה : הסיכוי למאורע נתון שהתרחש.

**דוגמה (פתרון בהקלטה) :**

נבחרו 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, בקרוב 15% מהמשפחות שתי המכוניות הן מתוצרת אירופאית. אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדש אירופאי?

**שאלות:**

- 1)** תלמיד ניגש בסմיסטר לשני מבחנים: מבחן בכלכלה ו מבחן בסטטיסטיקה :  
נגידיר את המאורעות הבאים :  
 A - עבר את המבחן בסטטיסטיקה.  
 B - עבר את המבחן בכלכלה.  
 כמו כן נתון שהסיכוי לעبور את המבחן בכלכלה הנו 0.8, הסיכוי לעبور את המבחן בסטטיסטיקה הנו 0.9 והסיכוי לעبور את שני המבחנים הנו 0.75.  
 חשבו את הסיכויים למאורעות הבאים :  
 א. התלמיד עבר בסטטיסטיקה, מה ההסתברות שהוא עבר בכלכלה?  
 ב. התלמיד עבר בכלכלה, מה ההסתברות שהוא עבר בסטטיסטיקה?  
 ג. התלמיד עבר בכלכלה, מה ההסתברות שהוא נכשל בסטטיסטיקה?  
 ד. התלמיד נכשל בסטטיסטיקה, מה ההסתברות שהוא נכשל בכלכלה?  
 ה. התלמיד עבר לפחות מבחן אחד, מה ההסתברות שהוא עבר את שניהם?
- 2)** במדינה שתי חברות טלפונ סוללארי : "סופט" ו"בל". 30% מההתושבים הבוגרים רשומים אצל חברת "בל", 60% מההתושבים הבוגרים רשומים אצל חברת "סופט" ול-15% מההתושבים הבוגרים אין טלפון סוללארי כלל.  
 א. איזה אחוז מההתושבים הבוגרים רשומים אצל שתי החברות?  
 ב. נבחר אדם רשום אצל חברת "סופט", מה ההסתברות שהוא רשום גם אצל חברת "בל" ?  
 ג. אם אדם לא רשום אצל חברת "בל", מה ההסתברות שהוא כן רשום בחברה "סופט" ?  
 ד. אם אדם רשום אצל חברת אחת בלבד, מה ההסתברות שהוא רשום בחברה "סופט" ?
- 3)** במכלה שני חניותים : חניון קטן וחניון גדול. בשעה 00:08 יש סיכוי של 60% שבחניון הגדל יש מקום, סיכוי של 30% שבחניון הקטן יש מקום וסיכוי של 20% שבחניון החניונים יש מקום.  
 א. מה ההסתברות שיש מקום בשעה 00:08 רק בחניון הגדל של המכלה?  
 ב. ידוע שבחניון הקטן יש מקום בשעה 00:08, מה הסיכוי שבחניון הגדל יש מקום?  
 ג. אם בשעה 00:08 בחניון הגדל אין מקום, מה ההסתברות שבחניון הקטן יהיה מקום?  
 ד. נתון שלפחות באחד מהחניותים יש מקום בשעה 00:08, מה ההסתברות שבחניון הגדל יש מקום?

4) נלקחו 200 שכירים ו-100 עצמאים. מתוך השכירים 20 הם אקדמיים, ומ当中 העצמאים 30 הם אקדמיים.

א. בנו טבלה שכיחות משותפת לנוטונים.

ב. נבחר אדם אקרי מה ההסתברות שהוא שכיר?

ג. מה ההסתברות שהוא שכיר ולא אקדמי?

ד. מה ההסתברות שהוא שכיר או אקדמי?

ה. אם האדם שנבחר הוא עצמאי מהי ההסתברות שהוא אקדמי?

ו. אם האדם שנבחר הוא לא אקדמי, מה ההסתברות שהוא שכיר?

### תשובות סופיות:

$$\text{ה. } 0.789 \quad \text{ד. } 0.5 \quad \text{ג. } 0.0625 \quad \text{ב. } 0.9375 \quad \text{א. } 0.833 \quad (1)$$

$$\text{. } 0.6875 \quad \text{ד. } 0.786 \quad \text{ג. } 0.0833 \quad \text{ב. } .5\% \quad \text{א. } (2)$$

$$\text{. } \frac{6}{7} \quad \text{ד. } 0.25 \quad \text{ג. } \frac{2}{3} \quad \text{ב. } 0.4 \quad \text{א. } (3)$$

$$\text{. } \frac{23}{30} \quad \text{ד. } 0.6 \quad \text{ג. } \frac{2}{3} \quad \text{ב. } \text{א. להלן טבלה: } (4)$$

סה"כ	לא אקדמי	אקדמי	
שכיר	200	180	20
עצמאי	100	70	30
סה"כ	300	250	50

$$\text{ה. } 0.3 \quad \text{ו. } 0.72$$

## bijustisjka le-refoat shinim

פרק 18 - דיאגרמת עצים - נסחת ביש ונוסחת הסתברות השלמה

תוכן העניינים

1. כללי .....

76 .....

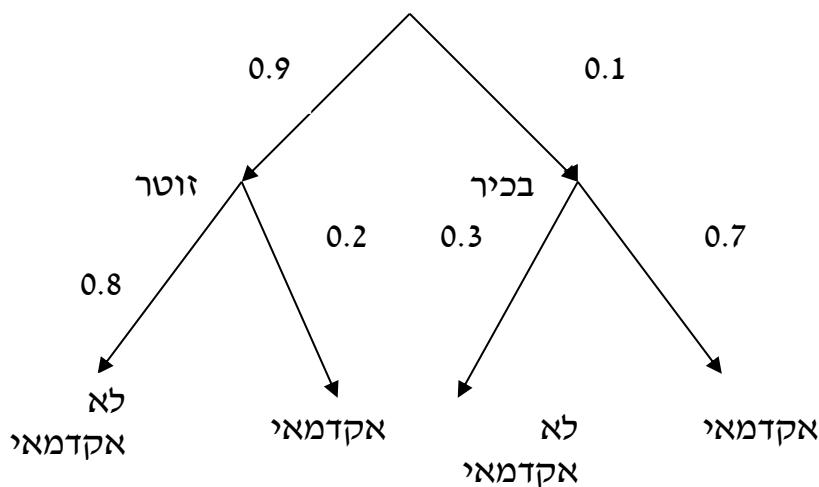
## דיאגרמת עצים – נוסחת הביס והסתברות השלמה:

---

נשתמש בשיטה זו כאשר יש תרגיל שבו התרחשויות המאורעות היא בשלבים, כך שכל תוצאה של כל שלב תלולה בשלב הקודם, פרט לשלב הראשון:

דוגמה:

בחברה מסוימת 10% מוגדרים בכירים והיתר מוגדרים זוטרים. מבין הבכירים 70% הם אקדמיים ומ בין הזוטרים 20% הם אקדמיים. נشرط עז שיתאר את הנתונים, השלב הראשון של העז אינו מותנה בכללם ואילו השלב השני מותנה בשלב הראשון.



כדי לקבל את הסיכוי לענף מסוים נכפיל את כל ההסתברויות על אותו ענף.  
נבחר אדם באקראי מאותה חברה.

- 1) מה הסיכוי שהוא בכיר אקדמי ?  $0.1 \cdot 0.7 = 0.07$ .
- 2) מה הסיכוי שהוא זוטר לא אקדמי ?  $0.9 \cdot 0.8 = 0.72$ .

כדי לקבל את הסיכוי לכמה ענפים נחבר את הסיכויים של כל ענף  
(רק אחרי שבתווך הענף הכפלו את ההסתברויות).

- 3) מה הסיכוי שהוא אקדמי ?  $0.25 + 0.9 \cdot 0.2 = 0.25 + 0.18 = 0.43$ .
- 4) נבחר אקדמי מה ההסתברות שהוא עובד זוטר?  
מדובר כאן על שאלה בהסתברות מותנה ולכן נשתמש בעיקרון של הסתברות  
モותנה :  $P(zutar | academay) = \frac{0.9 \cdot 0.2}{0.25} = \frac{0.18}{0.25} = 0.72$

**נוסחת ההסתברות השלמה:**

בהינתן  $B$ , מאורע כלשהו, וחלוקת של מרחב המדגמים  $\Omega$  ל-  $A_1, \dots, A_n$  כך ש- $\Omega = \bigcup_i A_i$ ,

$$\text{אזי: } P(B) = \sum_{i=1}^n P(A_i) \cdot P\left(\frac{B}{A_i}\right)$$

**נוסחת בייס:**

$$\cdot P\left(\frac{A_j}{B}\right) = \frac{P(A_j)P\left(\frac{B}{A_j}\right)}{\sum_{i=1}^n P(A_i) \cdot P\left(\frac{B}{A_i}\right)}$$

**שאלות:**

- 1) בשקית סוכריות 4 סוכריות תות ו-3 לימון. מוצאים באקראי סוכריה.  
אם היא בטעם תות אוכלים אותה ומוצאים סוכריה נוספת, ואם היא  
בטעם לימון מוחזרים אותה לשקית ומוצאים סוכריה נוספת.  
א. מה הסתברות שהסוכריה הראשונה שהוצאה בטעם תות  
והשנייה בטעם לימון?  
ב. מה הסתברות שהסוכריה השנייה בטעם לימון?
- 2) באוכלוסייה מסוימת 30% הם ילדים, 50% בוגרים והיתר קשישים. לפי נתוני  
משרד הבריאות הסיכוי שילד יחלה בשפעת משך החורף הוא 80%, הסיכוי  
شمボגר יחלה בשפעת משך החורף הוא 40% והסיכוי שקשיש יחלה בשפעת  
במשך החורף הוא 70%.  
א. איזה אחוז מהאוכלוסייה הינו קשישים שלא יחלו בשפעת במשך החורף?  
ב. מה אחוז האנשים שיחלו בשפעת במשך החורף?  
ג. נבחר אדם שחלה במשך החורף בשפעת, מה הסתברות שהוא קשיש?  
ד. נבחר ילד, מה הסתברות שהוא לא יחלה בשפעת במשך החורף?
- 3) בצד א' 5 כדורים כחולים ו-5 כדורים אדומים. בצד ב' 6 כדורים כחולים ו-4  
כדורים אדומים. בוחרים באקראי כד, מוצאים ממנו כדור ומבליל להחזירו  
מוצאים כדור נוסף.  
א. מה הסתברות שני ה כדורים שייצאו יהיו בצבעים שונים?  
ב. אם ה כדורים שהווצאו הם בצבעים שונים, מה הסתברות שהכדור השני  
שהווצה יהיה בצבע אדום?
- 4) חברת סלולר מסוגת את לקוחותיה לפי 3 קבוצות גיל: נוער, בוגרים ופנסיונרים.  
נתון כי: 10% מה לקוחות בני נוער, 70% מה לקוחות בוגרים והיתר פנסיונרים.  
מתוך בני הנוער 90% מוחזיקים בסמארט-פון, מתוך האוכלוסייה הבוגרת ל-70%  
יש סמארט-פון ומתוך אוכלוסיית הפנסיונרים 30% מוחזיקים בסמארט-פון.  
א. איזה אחוז מלקוחות החברה הם בני נוער עם סמארט-פון?  
ב. נבחר לקוח אקראי ונטען שיש לו סמארט-פון. מה הסתברות שהוא פנסיון?  
ג. אם לקוח אין סמארט-פון, מה הסתברות שהוא לא בן נוער?

- (5) כדי להתקבל למקום עבודה יש לעבור שלושה מבחנים. המבחנים הם בשלבים, ככלומר לאחר כישלון במבחן מסוים אין אפשרות לגשת למבחן הבא אחריו. 70% מהמטופדים עוברים את המבחן הראשון. מתוכם, 50% עוברים את המבחן השני. מבין אלה שעוברים את המבחן השני 40% עוברים את המבחן השלישי.
- מה ההסתברות להתקבל לעבודה?
  - מועדן לא התקבל לעבודה. מה ההסתברות שהוא נכשל במבחן הראשון?
  - מועדן לא התקבל לעבודה. מה ההסתברות שהוא עבר את המבחן השני?
- (6) משרד הבריאות פרסם את הנתונים הבאים:
- מתוך אוכלוסיית הילדים והנוער 80% חולמים בשפעת בזמן החורף.  
מתוך אוכלוסיית המבוגרים (עד גיל 65) 60% חולמים בשפעת בזמן החורף.  
30% מההתושבים הם ילדים ונוער. 50% הם מבוגרים. היתר קשיים.  
כמו כן נתון ש68% מהאוכלוסייה תחלה בשפעת בחורף.
- מה אחוז החולים בשפעת בקרב האוכלוסייה הקשישה?
  - נבחר אדם שלא חלה בשפעת, מה ההסתברות שהוא לא קשיש?
- (7) רצאר שנמצא על החוף צריך לקלוט אנייה הנמצאת ב-1-4 האזוריים : A, B, C, D, E.  
אם האנייה נמצאת באזור A הרצאר מזזה אותה בסיכון 0.8, סיכון זה פוחת ב-0.1 כל שהאנייה מתקרבת באזור. כמו כן נתון שהסתברות חצי האנייה נמצאת באזור D, בהסתברות 0.3 באזור C, באזור B היא נמצאת בסיכון 0.2, אחרת היא נמצאת באזור A.
- מה הסיכון שהאנייה מתגלה ע"י הרצאר?
  - אם האנייה התגלתה ע"י הרצאר, מה ההסתברות שהיא נמצאת באזור C?
  - אם האנייה התגלתה ע"י הרצאר, מה הסיכון שהיא לא נמצאת באזור B?
- (8) סימפטום X מופיע בהסתברות של 0.4 במחלה A, בהסתברות של 0.6 במחלה B ובಹסתברות של 0.5 במחלה C. סימפטום X מופיע אך ורק במקרים הללו, אדם לא יכול לחנות בו יותר ממחלה אחת מפני המחלות הללו. קלינייקה מגיעים אנשים כדלקמן: 8% חולמים במחלה A, 10% במחלה B, 2% במחלה C והיתר בריאות. כמו כן נתון שבמחלה A, סימפטום X מתגלה בסיכון של 80%, ובמחלות C, B הסימפטום מתגלה בסיכון של 90% בכל מקרה.
- מה ההסתברות שאדם הגיעו קלינייקה וגילו אצל סימפטום X?
  - אם התגלה אצל אדם סימפטום X, מה ההסתברות שהוא חולה במחלה A?
  - אם לאדם יש את סימפטום X, מה ההסתברות שהוא חולה במחלה A?
  - אם לא הגיעו אצל אדם סימפטום X, מה ההסתברות שהוא בריאות?

**תשובות סופיות:**

.0.2 .ד	.0.241 .ג	.58% .ב	.6% .א	(1) .2
		.0.5 .ב	.0.544 .א	(3)
.0.9722 .ג	.0.09375 .ב	.9% .א	(4)	
.0.2442 .ג	.0.3488 .ב	.0.14 .א	(5)	
	.0.8125 .ב	.70% .א	(6)	
.0.7543 .ג	.0.3158 .ב	.0.57 .א	(7)	
.0.8778 .ד	.0.3137 .ג	.0.2889 .ב	.0.0886 .א	(8)

# bijustisjka le-refoat shinim

פרק 19 - תלות ואי תלות בין מאורעות

תוכן העניינים

1. כללי .....

81 .....

## תלות ואי תלות בין מאורעות:

**רעיון:**

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

הדבר גורר גם ההפק:  $P(A|B) = P(A)$ , כלומר, גם  $A$  אינו תלוי 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,

חוקרים מבצעו שני ניסויים בלתי תלויים הסيكוי להצלחה בניסוי הראשון הוא 0.7 והסיקוי להצלחה בניסוי השני הוא 0.4.

א. מה הסיקוי להצלחה בשני הניסויים יחדיו?

כיוון שהמאורעות הללו בלתי תלויים :

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

ב. מה הסיקוי להיכשל בשני הניסויים?

באופן דומה :

$$P(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B}) = (1 - 0.7)(1 - 0.4) = 0.18$$

**שאלות:**

- 1)** נתון:  $P(A) = 0.2$ ,  $P(B) = 0.5$ ,  $P(A \cup B) = 0.6$ .  
האם המאורעות הללו בלתי תלויים?
- 2)** תלמיד הגיע לשני מבחנים שהצלחתם לא תליה זו בזו.  
הסיכוי שלו להצלחה בבחן הראשון הוא 0.7 והשני 0.4.  
א. מה הסיכוי להצלחה בשני המבחנים יחד?  
ב. מה הסיכוי שנכשל בשני המבחנים?
- 3)** במדינה מסוימת יש 8% אבטלה, נבחרו באקראי שני אנשים מהמדינה.  
א. מה ההסתברות שניהם מובטלים?  
ב. מה ההסתברות שלפחות אחד מהם מובטלי?
- 4)** מוצר צריך לעבור בהצלחה ארבעה בדיקות בלתי תלויות לפני שיוקו, אחרת  
הוא נפסל ולא יוצא לשוק. הסיכוי לעبور 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)** במדינה מסוימת יש 8% אבטלה, נבחרו באקראי חמישה אנשים מהמדינה.  
א. מה ההסתברות שכולם מובטלים במדגם?  
ב. מה ההסתברות שלפחות אחד מהם מובטלי?

**תשובות סופיות:**

- (1) כנ.
- (2) א. 0.18      ב. 0.28
- (3) א. 0.1536      ב. 0.0064
- (4) א. 0.9984      ב. 0.5904
- (5) א. 0.3409      ב. 0.08<sup>5</sup>

# bijustisjka le-refoat shinim

## פרק 20 - הערכת כלים אבחנתיים

תוכן העניינים

1. הערכת כלים אבחנתיים .....

83 .....

## הערכת כלים אבחנתיים:

**רקע:**

אנו מנסים לאבחן תכונה מסוימת באמצעות כלים מסוימים (למשל, לאבחן האם לאדם יש קורונה באמצעות ערכת אבחון ביתיית). נגידר מודדים סטטיסטיים שונים שנותנים אינדיקציה לאיכות הכלים האבחנה.

נסמן ב-*A*: האדם קיבל תשובה חיובית, כלומר אובחן כבעל התכונה.

נסמן ב-*B*: האדם הוא בעל התכונה.

**רגישות – Sensitivity :**

.  $Sensitivity = P(A|B)$  ההסתברות שאדם בעל התכונה קיבל תשובה חיובית, כלומר :

**סגוליות – Specificity :**

.  $Specificity = P(\bar{A}|\bar{B})$  ההסתברות שאדם ללא התכונה קיבל תשובה שלילית, כלומר :

**דוגמה (פתרון בהקלטה):**

5,000 נשים השתתפו במחקר שבו הון השתמשו בurveה לבדיקת הירינו ביתיית של חברות מסויימת ביום ה-14 של המחזור החודשי. בדיעבד מתוך 5,000 הנשים 80 היו בהירינו. 70 נשים קיבלו תשובה חיובית ממערכת הבדיקה הביתיית. 65 מהן אכן היו בהירינו. מה הסגוליות ומה הרגישות של בדיקת ההירינו הביתיית?

**הערך המנbaşı החיובי – Positive Predictive Value :**

ההסתברות שאדם שקיבול תשובה חיובית הוא אכן בעל התכונה.  
 הערך המנbaşı החיובי נקרא גם יכולת אבחון (יכולת דיאגנוטית) :  $PPV = P(B|A)$ .

**הערך המנbaşı השלילי – Negative Predictive Value :**

ההסתברות שאדם שקיבול תשובה שלילית אינו בעל התכונה :  $NPV = P(\bar{B}|\bar{A})$

**המשך הדוגמה (פתרון בהקלטה):**

מהו הערך המנבי החיובי ומהו הערך המנבי השיליי של ערכת הבדיקה הביתית?

**הסתברות לתוצאה חיובית מדוימה – False Positive :**

.  $FP = P(A|\bar{B})$  ההסתברות שאדם שאינו בעל התכונה קיבל תשובה חיובית :

**הסתברות לתוצאה שלילית מדוימה – False Negative :**

.  $FN = P(\bar{A}|B)$  ההסתברות שאדם בעל התכונה קיבל תשובה שלילית :

**המשך הדוגמה (פתרון בהקלטה):**

מה ההסתברות לתוצאה חיובית מדוימה ומה ההסתברות לתוצאה שלילית מדוימה בבדיקה באמצעות ערכת הבדיקה הביתית?

**שאלות:**

- 1)** 2,000 גברים נבדקו בשיטה חדשה לגילוי סרטן המעי הגס. מתוך 150 גברים שביופסיה הוכיחה בזודאות שהם חולמים 100 נמצאו חולמים באמצעות הבדיקה החדשה. 80 גברים שהוכיחו שהם בראים באמצעות ביופסיה קיבלו תשובה חיובית באמצעות השיטה החדשה. מצאו את הסגוליות, הרגשות, הערך המנבה החיובי, הערך המנבה השיליי, ההסתברות לתוצאה חיובית מדומה.
- 2)** הסיכוי של ערכת בדיקה ביתית של חברת קאנו לגילותמחלה מסוימת הוא 98%. לאדם בריא יש סיכוי של 5% לקבל בבדיקה תשובה חיובית. 5% ממשתמשי הערכה חולמים במחלה.
- מה הרגשות של הערכה?
  - מה הסגוליות של הערכה?
  - מה יכולת האבחן של הערכה?
- 3)** הסיכוי של תהליך אבחן של הפרעת קש布 לזהות סטודנטים בעלי הפרעת קש布 הוא 0.95. הסיכוי שלו לזהות בעיות גם סטודנטים ללא הפרעת קש布 כבעלי הפרעת קש布 הוא 0.01. ידוע ש-7 מהסטודנטים סובלים מהפרעת קש布.
- מה ההסתברות לתשובה חיובית מדומה?
  - מה ההסתברות לתשובה שלילית מדומה?
- 4)** נרוכה בדיקה בשיטה חדשה לגילוי מלנומה בקרב 800 נבדקים – 400 נבדקים עם ביופסיה מוחחת של מלנומה ויתר הנבדקים ידועים ללא מלנומה. 30 מהנבדקים נמצאו בבדיקה החדשה חולמים במלנומה. מתוך ה-30 רק 10 חולמים באמת.
- מה הרגשות של הבדיקה החדשה?
  - מה הסגוליות של הבדיקה החדשה?
  - מה יכולת הדיאגנוזטיבית של הבדיקה החדשה?
- 5)** קבועה של נבדקים שידעו ש-25% מהם נשאי HIV נבדקה בבדיקה חדשה המאפשרת לקבל תוצאה באופן מיידי. בבדיקה החדשה נמצא 20% מהנבדקים נשאי HIV. הסגוליות של הבדיקה החדשה היא 76%.
- מה הרגשות של הבדיקה החדשה?
  - מה ה-PPV של הבדיקה החדשה?
  - מה ה-NPV של הבדיקה החדשה?

**תשובות סופיות:**

- 1) סגוליות: 0.9568, רגישות:  $\frac{5}{9}$ , הערך המנба החיוובי: 0.0043.  
 הערך המנба השילי: 0.9725, ההסתברות ל贤א חיוובית מודומה: 0.0043.
- |                 |         |       |     |
|-----------------|---------|-------|-----|
| א. 0.5078       | ב. 0.95 | .0.98 | (2) |
| ג. .0.05        | ב. 0.01 | .0.01 | (3) |
| . $\frac{1}{3}$ | ב. 0.95 | .0.25 | (4) |
| ג. 0.7125       | ב. 0.1  | .0.08 | (5) |

# bijustisjka le-refoat shinim

## פרק 21 - שאלות מסכמת בהסתברות

תוכן העניינים

1. כללי .....

87 .....

## שאלות מסכימות בהסתברות:

### שאלות:

- 1)** 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.
- מה ההסתברות שמשפחה אקראייה בת שתי מכוניות תהיה ללא מכוניות מתוצרת אירופה?
  - מה ההסתברות שלפחות מכונית אחת תהיה אירופאית?
  - ידוע שלמשפחה יש מכונית אירופאית. מה ההסתברות שההיא מתוצרת אירופאית?
  - אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית?
- 2)** במדינת "שומקס" 50% מהחלב במרקולים מיוצר במחלבה א', 40% במחלבה ב' והיתר במחלבה ג'. 3% מתוצרת מחלבה א' מגיעה חמוצה למרקולים ואילו במחלבה ב' 10%. כמו כן ידוע שבמדינה "שומקס" בסך הכל 7.5% מהחלב חמוץ.
- איזה אחוז מהחלב שmagiu למרקול ממחלב ג' חמוץ?
  - אם נרכש חלב חמוץ במרקול. מה הסיכוי שהוא יוצר במחלבה ג'?
  - ברכישת חלב נמצא שאיןו חמוץ. מה הסיכוי שהוא יוצר במחלבה א'?
  - האם המאורעות: "חלב חמוץ" ו-"יוצר במחלבה א'" בלתי תלויים?
- 3)** רוני ורונה יצאו לבנות במרקז בילויים עם מספר אפשרויות בילוי: בהסתברות של 0.3 הם ייצאו לבאולינג, בהסתברות של 0.5 הם ייצאו לבית קפה ובಹסתברות של 0.7 הם ייצאו לפחות לאחד מהם (באולינג/קפה).
- מה ההסתברות שהם ייצאו רק לבאולינג?
  - האם המאורעות "lezat lebauling" ו-"lezat libet kafe" זרים?
  - האם המאורעות "lezat lebauling" ו-"lezat libet kafe" תלויים?
  - מה ההסתברות שיום אחד הם ייצאו רק לבאולינג וביום לאחר מכן לא ייצאו אף אחד מהמקומות?

4) 70% מהנבחנים בסטטיסטיקה עוברים את מועד א'. כל מי שלא עבר את מועד א' ניגש לעשות מועד ב', מתוכם 80% עוברים אותו. בין אלה שנכשלים בשני המועדים 50% נרשמים לקורס מחדש, והיתר פורשים מהתוואר.

א. מה הסיכוי שסטודנט אקראי עבר את הקורס?

ב. אם סטודנט אקראי עבר הקורס, מה הסיכוי שעבר במועד ב'?

ג. מה אחוז הסטודנטים שפורשים מהתוואר?

ד. נבחרו 2 סטודנטים אקראים רוניית וינאי, מה ההסתברות שרונית עברה במועד א' ושינאי עבר במועד ב'?

5) באוכלוסייה מסוימת 40% הם גברים והיתר הן נשים. מבין הגברים 10% מובטלים. בסך הכל 13% מהאוכלוסייה מובטלת.

א. מה אחוז האבטלה בקרב הנשים?

ב. נבחר אדם מובטל, מה ההסתברות שהוא אישה?

ג. נגידיר את המאורעות הבאים : A - נבחר אדם מובטל, B - נבחר גבר. האם המאורעות הללו זרים? והאם הם בלתי תלויים?

6) בתיבה 10 מטבעות, מתוכם 7 מטבעות לגברים (ראש, זנב) ו-3 מטבעות שבשני צדיהם טבוע ראש. אדם בוחר באקראי מטבע ומטיל אותו פעמיים. נסמן ב-A את הטלת הראשונה הראשונה בראש, וב-B את הטלת השנייה בראש.

א. חשבו את הסיכויים למאורעות A ו-B.

ב. האם המאורע A ו-B בלתי תלויים?

ג. ידוע שהטלת הראשונה התקבל ראש, מה ההסתברות שהמטבע שהוטל הוא מטבע הוגן?

7) ערן מעוניין למכור את רכבו והוא מפרסם מודעה באינטרנט ומודעה בעיתון. מבין אלה שמעוניינים לרכוש רכב משומש 30% יראו את המודעה באינטרנט, 50% יראו את המודעה בעיתון ו-72% יראו את המודעה לפחות אחת מהמדיות.

א. מה אחוז האנשים, לפחות שמעוניינים לרכוש רכב משומש, שיראו את 2 המודעות?

ב. אם אדם ראה את המודעה באינטרנט, מה ההסתברות שהוא לא ראה את המודעה בעיתון?

ג. האם המאורעות : "לראות את המודעה באינטרנט" ו-"לראות את המודעה בעיתון" בלתי תלויים?

ד. אדם שראה את המודעה באינטרנט בלבד יתקשר לערן בהסתברות של 0.7, אם הוא ראה את המודעה בעיתון בלבד הוא יתקשר לערן בהסתברות של 0.6. ואם הוא ראה את שתי המודעות הוא יתקשר לערן בהסתברות של 0.9.

i. מה ההסתברות שאדם מעוניין לרכוש רכב משומש יתקשר לערן?

ii. אדם מעוניין לרכוש רכב משומש התקשר לערן. מה ההסתברות שהוא ראה את שתי המודעות?

**תשובות סופיות:**

- |            |                     |               |           |     |
|------------|---------------------|---------------|-----------|-----|
| .0.5       | .0.6                | .0.75         | .0.25     | (1) |
| ד. תלויים. | ג. 0.524            | ב. 0.267      | א. 0.2    | (2) |
| .0.06      | ג. תלויים.          | ב. אינם זרים. | א. 0.2    | (3) |
| .0.168     | ג. 0.03             | ב. 0.255      | א. 0.94   | (4) |
|            | ג. לא זרים ותלויים. | ב. 0.692      | א. 15%    | (5) |
|            | ג. 0.5384           | ב. תלויים.    | א. 0.65   | (6) |
| .0.478     | ג. תלויים.          | ב. 0.733      | א. 8%     | (7) |
|            |                     |               | .0.15 .ii | ד.  |

# bijustisjka le-refoat shinim

## פרק 22 - המשטנה המקרי הבדיקה - פונקציית ההסתברות

תוכן העניינים

1. כללי ..... 90

## המשתנה המקרי הבודד – פונקציית הסתברות:

**רקע:**

**משתנה מקרי בודד:**

משתנה מקרי בודד הינו משתנה היכול לקבל כמה ערכים בודדים בהסתברויות שונות.

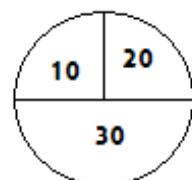
מתארים את המשתנה המקרי על ידי פונקציית הסתברות.

**פונקציית הסתברות:**

פונקציה המתאימה לכל ערך אפשרי של המשתנה את ההסתברות שלו.  
סכום ההסתברויות על פונקציית ההסתברות חייב להיות 1.

**דוגמה (פתרון בהקלטה):**

בקייםנו יש רולטה כמתואר בשרטוט:



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה ב-₪.  
בנו את פונקציית ההסתברות של סכום הזכיה במשחק בודד.

**שאלות:**

- 1)** ידוע שבישוב מסוים התפלגות מספר המכוניות למשפחה היא :  
 50 משפחות אין מכוניות במכונית.  
 70 משפחות עם מכונית אחת.  
 60 משפחות עם 2 מכוניות.  
 20 משפחות עם 3 מכוניות.  
 בוחרים באקראי משפחה מהיישוב, נגידר את X להיות מספר המכוניות של המשפחה שנבחרה. בנו את פונקציית ההסתברות של X.
- 2)** מהאותיות : A , B , C יוצרים קוד דו תוווי.  
 א. כמה קודים ניתן ליצור?  
 ב. רשמו את כל הקודים האפשריים.  
 ג. נגידר את X להיות מספר הפעמים שהאות B מופיעה בקוד.  
 בנו את פונקציית ההסתברות של X.
- 3)** תלמיד ניגש בסמסטר לשני מבחנים : מבחן בכלכלה ו מבחן בסטטיסטיקה. כמו כן, נתון שהסיכוי לעبور את המבחן בכלכלה הינו 0.8, הסיכוי לעبور את המבחן בסטטיסטיקה הינו 0.9 והסיכוי לעبور את שני המבחנים הינו 0.75. יהי X מספר המבחנים שהסטודנט עבר. בנו את פונקציית ההסתברות של X.
- 4)** הסיכוי לזכות במשחק מסוים הינו 0.3. אדם משחקים את המשחק עד אשר הוא מנצח אך ככל מקרה הוא לא משחקים את המשחק יותר מ-4 פעמים.  
 נגידר את X להיות מספר הפעמים שהוא שיחק את המשחק.  
 בנו את פונקציית ההסתברות של X.

**תשובות סופיות:**

(1) להלן טבלה :

3	2	1	0	$X$
0.1	0.3	0.35	0.25	$P(X)$

(2) להלן טבלה :

2	1	0	$X$
$\frac{1}{9}$	$\frac{4}{9}$	$\frac{4}{9}$	$P(X)$

(3) להלן טבלה :

2	1	0	$X$
0.75	0.20	0.05	$P(X)$

(4) להלן טבלה :

4	3	2	1	$X$
0.343	0.147	0.21	0.3	$P(X)$

# bijustisjka le-refoat shinim

פרק 23 - המשטנה המקרי הבדיקה - תוחלת - שונות וסטיות תקן

תוכן העניינים

1. כללי ..... 93

## המשתנה המקרי הבודד – תוחלת, שונות וסטיית תקן:

**רקע:**

**תוחלת:**

ממושיע של פונקציית ההסתברות, אם נבצע את התהילה אינסוף פעמים כמו בדוגמה נקבל. התוחלת היא צפיי של המשתנה המקרי.

$$\text{מגדירים תוחלת באופן הבא : } \mu = E(X) = \sum_i x_i P(x_i)$$

**שונות:**

תוחלת ריבועי השונות מהתוחלת – נותן אינדיקציה על הפיזור והסיכון של פונקציית ההסתברות.

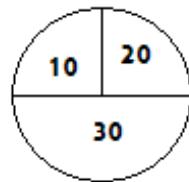
$$\text{מגדירים שונות באופן הבא : } V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = \sigma^2$$

**סטיית תקן:**

. שורש של השונות – הפיזור המוצע הצפוי סביב התוחלת. מסומנים :  $\sigma$  .  $STD = \sigma$

**דוגמה :**

בקזינו רולטה כמורה בשרטוט. אדם מסובב את הרולטה וזוכה בסכום הרשות על הרולטה ב-₪ . הסתברות לקבלת הסכומים השונים :



30	20	10	X
0.5	0.25	0.25	$P(X)$

$$E(X) = 10 \cdot 0.25 + 20 \cdot 0.25 + 30 \cdot 0.5 = 22.5 = \mu$$

$$V(X) = \sum_i (x_i - \mu)^2 P(x_i) =$$

$$= (10 - 22.5)^2 \cdot 0.25 + (20 - 22.5)^2 \cdot 0.25 + (30 - 22.5)^2 \cdot 0.5 = 68.75 = \sigma^2$$

כדי לחשב את סטיית התקן נוציא שורש לשונות :  $\sigma_x = \sqrt{V(X)} = \sqrt{68.75} = 8.29$

**שאלות:**

**1)** אדם משחק במשחק מזל.

נגדיר את  $X$  להיות סכום הזכיה.

להלן פונקציית ההסתברות של  $X$ :

40	20	0	-30	$X$
0.2	0.3	0.1	0.4	$P(X)$

מהי התוחלת, השונות וסטיית התקן של  $X$ ?

**2)** בישוב מסוים שני סניפי בנק: בנק פועלים ובנק לאומי. מתוך האוכלוסייה הבוגרת בישוב, ל-50% חשבו בנק בסניף הפועלים, ל-40% חשבו בנק בסניףלאומי ול-20% מההתושבים הבוגרים אין חשבו באף אחד מהסניפים. יהיו  $X$  מס' סניפי הבנק שלבוגר בישוב יש בהם חשבו. חשבו את  $E(X)$ .

**3)** ידוע של- 20% מהמשפחות יש חיבור לווייני בبيתם. בסקר אדם מחפש לראיין משפחה המחברת לוויין. הוא מטלפון באקראי למשפחה וממשיך עד אשר הוא מגיע למשפחה המחברת לוויין. בכל מקרה הסוקר לא יתקשר ליותר מ-5 משפחות. נגדיר את  $X$  להיות מספר המשפחות שאלייהן האדם יתקשר. א. בנו את פונקציית ההסתברות של  $X$ . ב. חשבו את התוחלת וסטיית התקן של  $X$ .

**4)** לאדם צורו מפתחות. בצרור 5 מפתחות אשר רק אחד מתאים לדלת של ביתו. האדם מנסה את המפתחות באופן מקרי. לאחר שניסיה מפתח מסוים הוא מוציא אותו מהצרור כדי שלא ישמש בו שוב. נסמן ב- $X$  את מספר הניסיונות עד שהדלת תפתח. א. בנו את פונקציית ההסתברות של  $X$ . ב. חשבו את התוחלת והשונות של  $X$ .

### תשובות סופיות:

- 1) תוחלת : 2 , שונות : 7.96
- 2) .0.9
- 3) א. ראו סרטון . ב. תוחלת : 3.36 , סטיית תקן : 1.603.
- 4) א. ראו טבלה :

5	4	3	2	1	X
0.2	0.2	0.2	0.2	0.2	$P(X)$

# bijustitsiKA لרפואת SHINIM

פרק 24 - המשטנה המקרי הבדיקה - טרנספורמציה לינארית

תוכן העניינים

1. כללי ..... 96

## המשתנה המקורי הבודד – טרנספורמציה לינארית:

**רקע:**

טרנספורמציה לינארית היא מצב שבו מבצעים הכפלת קבוע ו/או הוספה של קבוע על המשתנה המקורי (כולל גם חלוקה של קבוע והחסרה של קבוע).

בניסוח מתמטי נאמר כי אם משתנה אקראי  $Y$  מוצג ע"י משתנה אקראי  $X$  כאשר  $a, b$  הם קבועים כלשהם:  $Y = aX + b$ , אז מתקיימים:

$$\cdot E(Y) = aE(X) + b \quad (1)$$

$$\cdot V(Y) = a^2 \cdot V(X) \quad (2)$$

$$\cdot \sigma_Y = |a| \sigma_X \quad (3)$$

**שלבי העבודה:**

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

**דוגמה – הרולטה:**

במשחק לנוטני שאלת הרולטה נתנו שעלות השתתפות במשחק 15 ש". מהי התוחלת והשונות של הרווח במשחק?

**פתרון (בחקלה):**

$$\cdot E(X) = 22.5 = \mu, \quad V(X) = 68.75 = \sigma^2$$

### שאלות:

- 1)** סטודנט ניגש ל-5 קורסים הסמסטר. נניח שכל קורס שסטודנט מסיים מזכה אותו ב-4 נקודות אקדמיות. חשבו את התוחלת והשונות של סך הנקודות שיצבור הסטודנט כאשר נתון שתוחלת מספר הקורסים שישים היא 3.5 עם שונות 2.
- 2)** תוחלת סכום הזכיה במשחק מזל הינה 10 עם שונות 3. הוחלט להכפיל את סכום הזכיה במשחק. עלות השתתפות במשחק הינה 12.  
מה התוחלת ומהי השונות של הרווח במשחק?
- 3)** תוחלת של משתנה מקרי הינה 10 וסטיית התקן 5. הוחלט להוסיף 2 למשתנה ולאחר מכן להעלות אותו ב-10%.  
מהי התוחלת ומהי סטיית התקן לאחר השינוי?
- 4)**  $X$  הינו משתנה מקרי. כמו כן נתון ש- $E(X) = 4$  ו- $V(X) = 3$ .  
 $Z$  הינו משתנה מקרי חדש, עבורו:  $X - 7 = Y$ . חשבו את:  $E(Y)$  ו- $V(Y)$ .
- 5)** אדם החליט לבטא את רכבו; שווי הרכב 100,000 נק' להלן התוצאות האפשריות והסתברותן: בהסתברות של 0.001 תהיה תביעה טוטאליסט (כל שווי הרכב).  
בבסתברות של 0.02 תהיה תביעה בשווי מחצית משווי הרכב.  
בבסתברות של 5% תהיה תביעה בשווי רבע משווי הרכב.  
אחרת אין תביעה בכלל. החברה מאפשרת תביעה אחת בשנה.  
נסמן ב- $X$  את גובה התביעה השנתית, באלפיות.  
א. בנו את פונקציית ההסתברות של  $X$ .  
ב. חשבו את התוחלת והשונות של גובה התביעה.  
ג. פרמיית הביטוח היא 4,000 נק'.
- מהי התוחלת ומהי השונות של רווח חברת הביטוח לביטוח הרכב הנ"ל?

**תשובות סופיות:**

- (1) תוחלת : 14 , שונות : 32 .
- (2) תוחלת : 8 , שונות : 12 .
- (3) תוחלת : 13.2 , סטיית תקן : 5.5 .
- (4) תוחלת : 3 , שונות : 3 .
- (5) א. להלן טבלה :  
 ב. תוחלת : 2350 , שונות :  $85,727.5^2$

0	25	50	100	X
0.929	0.05	0.02	0.001	$P(X)$

ג. תוחלת : 1650 , שונות :  $85,727.5^2$

# bijustisjka le-refoat shinim

פרק 25 - התפלגיות בדידות מיוחדות - התפלגות ביןומית

תוכן העניינים

99 ..... 1. כללי .....

## התפלגיות בדידות מיוחדות – התפלגותBINOMIAL:

**רקע:**

נגידר את המושג ניסוי ברנולי:  
 ניסוי ברנולי הנה ניסוי שיש לו שתי תוצאות אפשריות: "הצלחה" ו"כישלון".  
 למשל מוצר פגום או תיקין, אדם עובד או מובטל, עץ או פלי בהטלה מטבח וכדומה.  
 בהתפלגותBINOMIAL חוזרים על אותו ניסוי ברנולי  $n$  פעמים באופן בלתי תלוי זה זהה.  
 מגדירים את  $X$  להיות מספר ההצלחות שהתקבלו בסך הכל. נסמן ב-  $P$  את הסיכוי  
 להצלחה בניסוי בודד, וב-  $Q$  את הסיכוי לכישלון בניסוי בודד.  
 אז נגיד ש:  $X \sim B(n, p)$ .

**פונקציית ההסתברות של  $X$ :**

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

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

לבודל:  $\binom{n}{k}$  ניתן לחשב באמצעות המחשבון.

**תוחלת:**  $E(X) = np$

**שונות:**  $V(X) = npq$

שימוש לב, כדי ליזות שמדובר בהתפלגותBINOMIAL צריכים להתקיים כל התנאים הבאים:

- 1) חוזרים על אותו ניסוי ברנולי באופן בלתי תלוי זה זהה.
- 2) חוזרים על הניסוי  $n$  פעמים.
- 3)  $X$  – מוגדר במספר ההצלחות המתקבלות בסך הכל.

**דוגמה (פתרון בהקלטה):**

במדינה מסוימת ל- 80% מהתושבים יש רישיון נהיגה.

נבחרו 10 תושבים אקרים מהמדינה.

- א. מה ההסתברות שבודיק ל- 9 מהם יש רישיון נהיגה?
- ב. מה ההסתברות של לפחות 9 מהם יש רישיון נהיגה?
- ג. מהי התוחלת ומהי סטיית התקן של מספר התושבים שנבדקו  
ושיש להם רישיון נהיגה?

**שאלות:**

- 1)** במדינה 10% מהאוכלוסייה מובטלה. נבחרו 5 אנשים באקראי מאותה אוכלוסייה. נגידר את  $X$  להיות מספר המובטלים שהתקבלו במדגם.
- מהי ההתפלגות של  $X$  ?
  - מה ההסתברות שהיא בדיקן מובטל אחד?
  - מה ההסתברות שכולם יעבדו במדגם?
  - מה ההסתברות שלושה 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 סמארטפון. נבחרו 10 אנשים באקראי. נגידר את  $X$  כמספר האנשים שנדרגו עם סמארטפון.
- מהי ההתפלגות של  $X$  ? הסבירו.
  - מה ההסתברות שבמדגם ל-8 אנשים יש סמארט-פון?
  - מה ההסתברות שבמדגם לפחות ל-9 יהיו סמארט-פון?
  - מה תוחלת ומה סטיית התקן של מספר האנשים שנדרגו ולהם סמארט-פון?
- 3)** בבית הימורים יש שורה של 6 מכונות מזל מאותו סוג. משחק במכונת מזל כזו עולה 5 ש". ההסתברות לזכות ב-20 ש"כ בכל אחת מהמכונות היא 0.1 וההסתברות להפסיד את ההשקה היא 0.9 בכל מכונה. מהי ממוצע כניסה לבית ההימורים ומכניס 5 ש"כ לכל אחת מ-6 המכונות.
- מה ההסתברות שיפסיד בכל המכונות?
  - מה ההסתברות שיזכה בבדיקה בשתי מכונות?
  - מה ההסתברות שיזכה ביותר בסך מה-30 ש"כ שהשקייע?
  - מהו התוחלת וסטיית התקן של הרוחות נטו של המהמר (הזכויות בניכוי ההשקה)?
- 4)** במדינה מסוימת התפלגות ההשכלה בקרב האוכלוסייה מעל גיל 30 היא כזו :

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

נבחרו 20 אנשים אקריםים מעל גיל 30.

- מה ההסתברות ש-5 מהם אקדמיים?
- מה תוחלת של מס' בעלי ההשכלה הנמוכה?

- 5) במכלה מסוימת 20% מהסטודנטים גרים בת"א. מבין הסטודנטים שגרים בת"א 30% מגיעים ברכבם, ומ בין הסטודנטים שלא גרים בת"א 50% מגיעים ברכבם למכלה.
- א. השומר בשער המכלה בודק לכל סטודנט את תיוקו בהיכנסו למכלה.  
 מה ההסתברות שבקרב 5 סטודנטים שנבדקו ע"י השומר רק 1 מתוכם הגיע למכלה ברכבו?
- ב. בהמשך לסייע הקודם מה ההסתברות שרוב הסטודנטים בקרוב ה-5 הגיעו למכלה ברכבם?
- 6) ב מבחן אמריקאי 20 שאלות. סטודנט ניגש לבחון והסıcıוי שהוא יודע שאלה כלשהי הוא 0.8. אם הוא לא יודע הוא מוחש את התשובה.  
 לכל שאלה 4 תשובות אפשריות שركacha אחת מהן נכון.  
 א. מה הסיכוי לענות על שאלה מסוימת נכון?  
 ב. מה הסיכוי שיענה נכון על בדיקת 16 שאלות?  
 ג. על כל שאלה שענה נכון התלמיד מקבל 5 נקודות, על כל שאלה שגגה מופחתת נקודה, מה התוחלת ומהי השונות של ציון התלמיד?
- 7) 5% מקו היוצר פגום. המוצריים נארזים בתוך קופסת קרטון. בכל קופסה 10 מוצרים שונים. הקופסאות נארזות בתוך מכולה. בכל מכולה 20 קופסאות.  
 א. מה ההסתברות שב קופסה אקראית לפחות מוצר אחד?  
 ב. מה התוחלת ומהי סטיית התקן של מספר הקופסאות במכלה בהן לפחות מוצר אחד?

**תשובות סופיות:**

- .0.59049 ג. .0.32805 ב. . $X \sim B(n=5, p=0.1)$  א. (1)
- .0.45 ו. תוחלת: 0.5, שונות: 0.40954 ה. 0.0729 ד.
- .1.449 ז. תוחלת: 7, סטיית תקן: 0.1493 ג. 0.2335 ב. (2)
- .0.1143 ג. 0.0984 ב. 0.5314 א. (3)
- .14.697 ד. תוחלת: -18, סטיית תקן: 0.2 0.1789 א. (4)
- .0.4253 ב. 0.1956 א. (5)
- .91.8 ג. תוחלת: 82, שונות: 0.182 ב. 0.85 א. (6)
- .2.193 ב. תוחלת: 8.025, סטיית תקן: 0.401 א. (7)

# bijustisjka le-refoat shinim

פרק 26 - התפלגיות רציפות מיוחדות - התפלגות נורמלית

תוכן העניינים

1. כללי ..... 103

## התפלגיות רציפות מיוחדות – התפלגות נורמלית:

---

**רקע:**

התפלגיות נורמלית הינה התפלגות של משתנה רציף. ישנו משתנים רציפים מסוימים שנחוג להתייחס אליהם כנורמליים כדוגמת זמן ייצור, משקל תינוק ביום היולדו ועוד. פונקציית הצפיפות של ההתפלגות הנורמלית נראה כmo פעמון:



לעוקמה זו קוראים גם עקומה גאוס ועוקמה אחת נבדלת מהשנייה באמצעות הממוצע וסטיית התקן שלה.

אליה הם הפרמטרים שמאפיינים את ההתפלגות:  $N(\mu, \sigma^2)$ .

$$\cdot f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

נוסחת פונקציית הצפיפות:

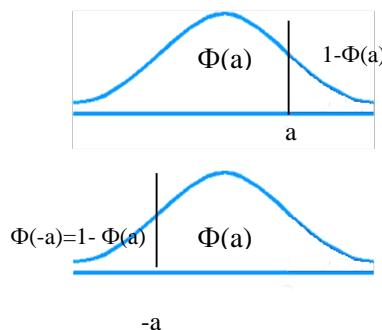
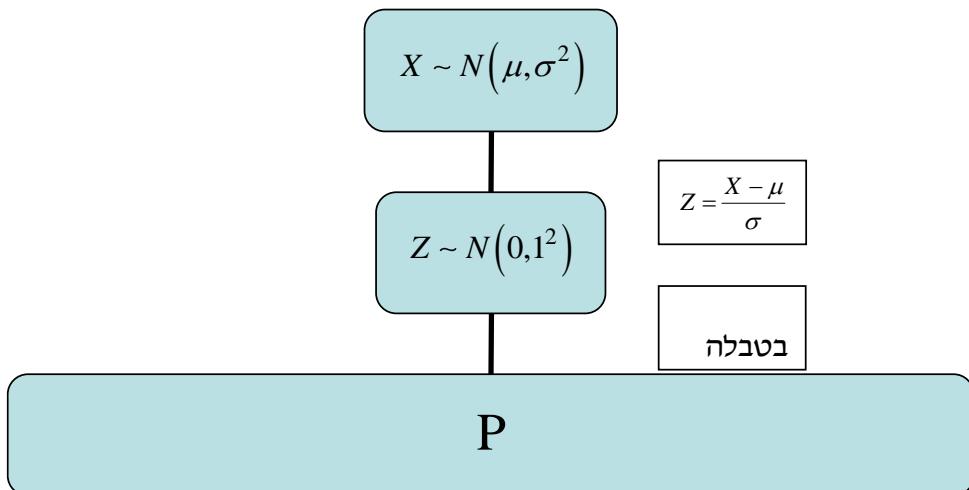
כדי לחשב הסתברויות בהתפלגות נורמלית יש לחשב את השטחים הרלוונטיים שמשתח על עוקמה. כדי לחשב שטחים אלה נמייר כל ההתפלגות נורמלית להתפלגות נורמלית סטנדרטית על ידי תהליך הנקרא תקנון. ההתפלגות נורמלית סטנדרטית היא ההתפלגות נורמלית שהממוצע שלה הוא אפס וסטיית התקן היא אחת, והיא מסומן באות  $Z$ :  $Z \sim N(0, 1^2)$ .

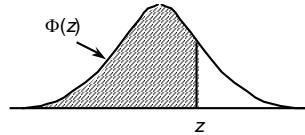
$$\cdot Z = \frac{X - \mu}{\sigma}$$

תהליך התקנון מבוצע על ידי הנוסחה הבאה:

אחרי התקנון מקבלים ערך הנקרא ציון תקן. ציון התקן משמשו בכמה סיטuatיות תקן הערך סוטה מהממוצע.

לאחר חישוב ציון התקן של ערך מסוים נזיררים בטבלה של ההתפלגות הנורמלית הסטנדרטית לחישוב השטח הרצוי, ובאופן כללי בהתאם להסכמה הבאה:



**טבלת התפלגות המცטברת 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	0.99999



דוגמה (הਪתרוון בהקלטה) :

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

- (1) מה אחוז חפיסות השוקולד ששוקלות מתחת ל-110 גרם?
- (2) מה אחוז חפיסות השוקולד השוקלות מעל 110 גרם?
- (3) מה אחוז חפיסות השוקולד השוקלות מתחת ל-92 גרם?
- (4) מהו המשקל ש-90% מהחפיסות בכו הייצור שוקלים פחות מהם?

**שאלות:**

- 1)** הגובה של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 170 ס"מ וסטיית תקן של 10 ס"מ.
- מה אחוז האנשים שגובהם מתחת ל-182.4 ס"מ?
  - מה אחוז האנשים שגובהם מעל 190 ס"מ?
  - מה אחוז האנשים שגובהם בדיקן 173.6 ס"מ?
  - מה אחוז האנשים שגובהם מתחת ל-170 ס"מ?
  - מה אחוז האנשים שגובהם לכל היותר 170 ס"מ?
- 2)** נתון שהזמן שלוקח לטרופה מסוימת להשפיע מתפלג נורמלית עם ממוצע של 30 דקות ושונות של 9 דקות רביעות.
- מהי פרופורציה המקרים בהן הטרופה תעוזר יותר מאשר משעה?
  - מה אחוז מהמקרים שהבחן הטרופה תעוזר בין 35 ל-37 דקות?
  - מה הסיכוי שהטרופה תעוזר בדיקן תוך 36 דקות?
  - מה שיעור המקרים שהבחן ההשפעה של הטרופה 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.
- מה אחוז האנשים שמשקלם נמוך מ-55 ק"ג?
  - מהי פרופורציה האנשים באוכלוסייה שמשקלם לפחות 50 ק"ג?
  - מהי השכיחות היחסית של האנשים באוכלוסייה שמשקלם בין 60 ל-70 ק"ג?
  - לאיזה חלק מהאוכלוסייה משקל הסוטה מהמשקל הממוצע بلا יותר מ-4 ק"ג?
  - מה הסיכוי שאדם אكري ישקל מתחת ל-140 ק"ג?
- 4)** משקל תינוקות ביום היולדם מתפלג נורמלית עם ממוצע של 3300 גרם וסטיית תקן 400 גרם.
- מצאו את העשרון העליון.
  - מצאו את האחוזון ה-95.
  - מצאו את העשרון התחתון.

- 5) ציוני מבחן אינטלייגנציה מתפלגים נורמליות עם ממוצע 100 ושונות 225.
- מה העשירון העליון של הציונים בבחן האינטלייגנציה?
  - 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 נורמליות עם סטיית תקן של 20 מ"ל, וננתן ש-33% מהבקבוקים בעלי נפח שעולה על 508.8 מ"ל.
- מה ממוצע נפח משקה בבקבוק?
  - 5% מהבקבוקים המזוכרים עם הנפח הגבוה ביותר נשלחים לבדיקה, החל מאייה נפח שלוחים בקבוק לבדיקה?
  - 1% מהבקבוקים עם הנפח הקטן ביותר נתרמים לצדקה, מהו הנפח המקסימלי לצדקה?
- 7) אורך חיים של מכשיר מתפלג נורמליות. ידוע שמחצית מהמכשירים חיים פחות מ-500 שעות, כמו כן ידוע ש-67% מהמכשירים חיים פחות מ-544 שעות.
- מהו ממוצע אורך חי מכשיר?
  - מהי סטיית התקן של אורך חי מכשיר?
  - מה הסיכוי שמכשיר אקראי יהיה פחות מ-460 שעות?
  - מהו המאיון העליון של אורח חי מכשיר?
  - 1% מהמכשירים בעלי אורך החיים קצר ביותר נשלחים לבדיקה מעמיקה. מהו אורך החיים המקסימלי לשילוח מכשיר למעבדה?
- 8) להלן שלוש התפלגויות נורמליות של שלוש קבוצות שונות ששורתטו באותה מערכת צירים. ההתפלגויות מוספרו כדי להבדיל ביניהן.
- לאיזו ההתפלגות הממוצע הגבוה ביותר?
  - במה מבין המינים הבאים ההתפלגות 1 ו-2 זהות?
    - בעשירון העליון.
    - בממוצע.
    - בשונות.  - לאיזו ההתפלגות סטיית התקן הקטנה ביותר?
    - .1
    - .2 .ii
    - .3 .iii
    - .iv אין לדעת.



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

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

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

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

ד. מה הסיכוי ששבוע (חמשה ימי עבודה) בדיקק פעמי אחת יהיה זמן הנסעה לפחות שלושת רביעי השעה?

10) אורך שיר אקראי המשודר ברדיו מתפלג נורמלית עם תוחלת של 3.5 דקות וסטיית תקן של שלושים שניות.

א. מה ההסתברות שאורך של שיר אקראי המונגן ברדיו יהיה בין 3 ל-2.5 דקות?

ב. מהו הטווח הבין רביעוני של אורך שיר המשודר ברדיו?

ג. ביום מסוים מנוגנים 200 שירים ברדיו. כמה שירים מתוכם תצפה שייהיו באורך הנמוך מ-3.5 דקות?

ד. בשעה מסוימת שודרו 8 שירים. מה ההסתברות שרבע מהם בדיקק היו ארוכים מ-4 דקות והיתר לא?

### תשובות סופיות:

.50%	ה.	.50%	ד	.0	ג	.2.28%	ב.	.89.25%	א.	(1)
.68.26%	ד	.0%	ג	.3.76%	ב.	.0%	א.	(2)		
.0.383	ד	.39.44%	ג	.89.44%	ב.	.26.43%	א.	(3)		
						.100%	ה.			
		.2787.2	ג	.3958	ב.	.3812.8	א.	(4)		
.87.4	ד	.112.6	ג	.80.8	ב.	.119.2	א.	(5)		
		.453.48	ג	.532.9	ב.	.500	א.	(6)		
.733	ד	.0.3446	ג	.100	ב.	.500	א.	(7)		
						.267	ה.			
		.1	ג	ב. במוצע.		.3	א.	(8)		
.0.3975	ד	.0.8563	ג	.0.0228	ב.	.0.1587	א.	(9)		
.0.25	ד	.100	ג	.0.675	ב.	.0.1359	א.	(10)		

# bijustisjka le-refoat shinim

פרק 27 - השקה סטטיסטית - הקדמה

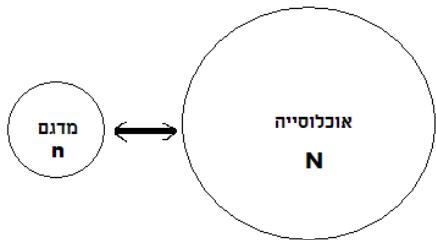
תוכן העניינים

1. כללי .....

111 .....

## הסקה סטטיסטית – הקדמה:

**רקע:**



**אוכלוסייה:**  
קבוצה שאליה מפנים שאלת מחקרית.  
למשל, חברת תרופות שמעוניינת לפתח תרופה  
למחלת הסוכרת מתעניינת באוכלוסיות חוליות  
הסוכרת בעולם.

**مثال:**

חלק מתוך האוכלוסייה.  
למשל, אם נדגים באקראי 10 אנשים מתוך חוליות הסוכרת אז זהו מثال מתוך  
אוכלוסיות חוליות הסוכרת.

במקרים רבים אין אפשרות לחקור את כל האוכלוסייה כיון שאין גישה לכולה,  
היא גדולה מדי, אנו מוגבלים בזמן ובאמצעים טכניים ולכן מבצעים מוגן במטרה  
לבצע הסקה סטטיסטית מהמוגן לאוכלוסייה.  
הדגימה בקורס תהיה דגימה מקראית - הכוונה לדוגמה שבה לכל תצפית באוכלוסייה  
יש את אותו סיכוי להיכל במדגם.

**סטטיטיסטי:**

מודל המוחש בעל המוגן.

**פרמטר:**

מודל המתאר את האוכלוסייה.

### הסימונים לפרמטר וסטטיסטי הם שונים:

סטטיסטי (מדגם)	פרמטר (אוכלוסייה)	
$\mu$	$\bar{X}$	משמעות
$P$	$\hat{p}$	פרופורציה (שכיחות יחסית)

פרמטר הוא גודל קבוע גם אם אנו לא יודעים אותו סטטיסטי הוא משתנה ממוגן למדגם ולכן יש לו התפלגות הנקראות התפלגות הדגימה.

### דוגמה (פתרון בהקלטה):

6% מאזרחי המדינה תומכים בהצעת החוק של חבר הכנסת מסוים. הוחלט לדגום 200 אזרחים ומתוכם לבדוק מהו אחוז התומכים בהצעת החוק.

- א. מי האוכלוסייה?
- ב. מה המשתנה?
- ג. מה הפרמטרים?
- ד. מהו גודל המדגם?
- ה. מהו הסטטיסטי שמתכוונים להוציא ממדגם?
- ו. האם הפרמטר או הסטטיסטי הוא משתנה מקרי?

**שאלות:**

- 1)** מתוך כלל הסטודנטים במכלה שסימנו סטטיסטיקה א' נדגוו שני סטודנטים. נתון שסכום הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מי האוכלוסייה?
  - מה המשטנה?
  - מהם הפרמטרים?
  - מהו גודל המדגם?
- 2)** להלן התפלגות מספר מקלט טלויזיה למשפחה בישוב "העוגן".  
נגידר את X להיות מספר המקלטים של משפחה אקראית. מתכנים לדגום מאוכלוסייה זו 4 משפחות ולהתבונן במספר מקלט טלויזיה במדגם.
- מיי האוכלוסייה ומהו המשטנה הנחקר?
  - מהו הסטטיסטי שיילקח מהדגם ומה סימונו?

מספר משפחות	מספר מקלטים
0	50
1	250
2	350
3	300
4	50
	סך הכל $N = 1000$

- 3)** נתון כי 20% מהשכירים במדינה הם אקדמיים. נבחרו באקראי 10 שכירים באותו אוכלוסייה ומתכנים לפרסם את מספר האקדמאים שנדגוו.
- מיי האוכלוסייה?
  - מה המשטנה באוכלוסייה?
  - מהם הפרמטרים?
  - מהו הסטטיסטי?

**תשובות סופיות:**

- 1)** א. כלל הסטודנטים במכלה שסימנו סטטיסטיקה א'.      ב. ציון.  
ג. ממוצע : 78, סטיית תקן : 15.      ד. 2.
- 2)** א. האוכלוסייה : 1000 משפחות בישוב העוגן, המשטנה הנחקר : מס' מקלטים.  
ב.  $\bar{X} = \text{ממוצע מדגם}$ .
- 3)** א. השכירים במדינה.  
ב. השכלה : אקדמי, לא אקדמי.  
ג. מס' האקדמאים באוכלוסייה : 0.2.

# bijustisjka le-refoat shinim

## פרק 28 - התפלגות הדגימה ומשפט הגבול המרכזי

### תוכן העניינים

114 .....	1. התפלגות ממוצע המדגם ומשפט הגבול המרכזי .....
121 .....	2. התפלגות פרופורצית ההצלחות במדגם .....
125 .....	3. התפלגות מספר ההצלחות במדגם - קירוב נורמלי להתפלגות הבינומית .....

## התפלגות ממוצע המדגם ומשפט הגבול המרכזי:

**רקע:**

$$\bar{x} = \frac{\sum x_i}{n}$$

בפרק זה נדון בהתפלגות של ממוצע המדגם :

מכיוון שמדובר למדגם אנו יכולים לקבל ממוצע מדגם שונה, אזי ממוצע המדגם הוא משתנה מקרי ויש לו התפלגות. גורמים המتأרים בהתפלגות כלשהי או אוכלוסייה כלשהי נקראים פרמטרים. להלן רישימה של פרמטרים החשובים לפרק זה : ממוצע האוכלוסייה נסמן ב-  $\mu$  (נקרא גם תוחלת).

שונות אוכלוסייה נסמן ב-  $\sigma^2$ . סטיית תקן של אוכלוסייה :  $\sigma$ .

### תכונות ההתפלגות:

ממוצע כל ממוצעי המדגם האפשריים שווה לממוצע האוכלוסייה :  $E(\bar{x}) = \mu_{\bar{x}} = \mu$  . שונות כל ממוצעי המדגם האפשריים שווה לשונות האוכלוסייה מחולק ב-  $n$  .

$$V(\bar{x}) = \frac{\sigma^2}{n}$$

תמונה זו נcona רק במדד מקרי :

יש יחס הפוך בין גודל המדגם לבין שונות ממוצעי המדגם.

אם נוציא שורש לשונות נקבל סטיית תקן של ממוצע המדגם שנקרהת גם

$$\sigma(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

טעות תקן :

### דוגמה (פתרון בהקלטה) :

השכר הממוצע במשק הינו 9000 נס עם סטיית תקן של 4000. דגמו באקרים 25 עובדים.

א. מיהי אוכלוסיית המחקר? מהו המשטנה הנחקר?

ב. מהם הפרמטרים של האוכלוסייה?

ג. מה התוחלת ומהי סטיית התקן של ממוצע המדגם?

**דוגמה מההתפלגות נורמללית:**

אם נדגם מתוך אוכלוסייה שהמשנה בה מתפלג נורמלית עם ממוצע  $\mu$  ושונות  $\sigma^2$ .

$$\text{ממוצע המדגם גם יתפלג נורמלית: } \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right), Z_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma}$$

**דוגמה (פתרון הבדיקה):**

משקל תינוק ביום היולדו מתפלג נורמלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.

מה ההסתברות שבמדגם של 4 תינוקות אקראיים בעת הולדתם המשקל הממוצע של התינוקות יהיה מתחת ל-3.5 ק"ג?

**משפט הגבול המרכזי:**

אם אוכלוסייה מתפלגת כלשהו עם ממוצע  $\mu$  ושונות  $\sigma^2$  אז עבור מדגם מספיק

$$\text{גדול } (n \geq 30) \text{ ממוצע המדגם מתפלג בקרוב לנורמלי: } \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

**דוגמה (פתרון הבדיקה):**

משקל חפיסת שוקולד בקוו ייצור מתפלג עם ממוצע 100 גרם וסטיית תקן של 4 גרם.

דגמו מכו הייצור 36 חפיסות שוקולד אקראיות.

מה ההסתברות שהמשקל הממוצע של חפיסות השוקולד שנדגו ייה מתחת ל-102 גרם?

### שאלות:

- 1) מתוך כלל הסטודנטים במכלה שסיוומו סטטיסטיקה א נדגו שני סטודנטים. נתון שסכום הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מיהי האוכלוסייה?
  - מה המשנה?
  - מהם הפרמטרים?
  - מהו גודל המדגם?
  - מהו תוחלת ממוצע המדגים?
  - מהי טעות התקן?

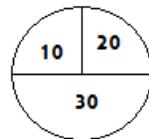
2) להלן התפלגות מספר מקלט טלויזיה למשפחה בישוב מסוים :

מספר משפחות	מספר מקלטים
0	500
1	2500
2	3500
3	3000
4	500
	סך הכל $N = 10000$

- א. בנו את פונקציית ההסתברות של  $X$ .
- ב. חשבו את התוחלת, השונות וסטיית התקן של  $X$ .
- ג. אם נדגו 4 משפחות מהישוב עם החזרה מה תהיה התוחלת, מהי השונות ומהי סטיית התקן של ממוצע המדגים?
- 3) אם נטיל קובייה פעמים ונתבונן בממוצע התוצאות שיתקבלו, מה תהיה התוחלת ומה תהיה סטיית התקן של ממוצע זה?
- 4) משקל תינוק ביום היולדו מתפלג נורמלית עם ממוצע 3400 גרם וסטיית התקן של 400 גרם.
- א. מה ההסתברות שתינוק אكري בעת הלידה ישקל פחות מ-3800 גרם?  
 נתון כי ביום מסוים נולדו 4 תינוקות.
- ב. מה ההסתברות שהמשקל הממוצע שלהם יעלה על 4 ק"ג?
- ג. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-5.2 ק"ג?
- ד. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת بلا יותר מ-50 גרם?
- ה. הסבירו ללא חישוב כיצד התשובה לשיעיף הקודם הייתה משתנה אם היה מדובר על יותר מ-4 תינוקות?

- 5) הגובה של המתגוייסים לצה"ל מתפלג נורמללית עם תוחלת של 175 ס"מ וסטיית תקן של 10 ס"מ. ביום מסויים התגוייסו 16 חיילים.
- מה ההסתברות שהגובה הממוצע שלהם יהיה לפחות 190 ס"מ?
  - מה ההסתברות שהגובה הממוצע שלהם יהיה בדיק 180 ס"מ?
  - מה ההסתברות שהגובה הממוצע שלהם יסטה מתחולת הגבהים לפחות מ-5 ס"מ?
  - מהו הגובה שבהסתברות של 90% הגובה הממוצע של המדגים יהיה נמוך ממנו?
- 6) הזמן הממוצע שלוקח לאדם להגיע לעבודתו 30 דקות עם שונות של 16 דקות רבועות. האדם נוסע לעבודה במשך שבוע 5 פעמים. לצורך הפטIRON הניחו שזמן הנסעה לעבודה מתפלג נורמלית.
- מה ההסתברות שבמשך שבוע משך הנסעה הממוצע יהיה מעל 33 דקות?
  - מהו הזמן שבהסתברות של 90% ממוצע משך הנסעה השבועי יהיה גבוה ממנו?
  - מה ההסתברות שמספרם משך הנסעה השבועי יהיה מרוחק מ-30 דקות לפחות 2 דקות?
  - כיצד התשובה לسؤال הקודם הייתה משתנה אם האדם היה נוסע לעבודה 6 פעמים בשבוע?
- 7) נפח היין בבקבוק מתפלג נורמללית עם תוחלת של 750 סמ"ק וסטיית תקן של 10 סמ"ק.
- בארכו 4 בקבוקי היין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארכו יהיה בדיק 755 סמ"ק?
  - בארכו 4 בקבוקי היין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארכו יהיה יותר מ-755 סמ"ק?
  - בארכו 4 בקבוקי היין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארכו יהיה לפחות 755 סמ"ק?
  - בקבוקי היין שבארכו נזוגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?
- 8) משתנה מתפלג נורמללית עם תוחלת 80 וסטיית תקן 4.
- מה ההסתברות שמספרם המדגים יסטה מתחולתו ללא יותר מichiיה כאשר גודל המדגים הוא 9?
  - מה ההסתברות שמספרם המדגים יסטה מתחולתו ללא יותר מichiיה כאשר גודל המדגם הוא 16?
  - הסביר את ההבדל בתשובות של שני הטעיפים.

**9)** בקזינו ישנה רולטה. על הרולטה רשומים המספרים הבאים כמוראה בشرطו:



- אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.
- בנו את פונקציית ההסתברות של סכום הזכיה במשחק בודד.
  - מה התוחלת ומה השונות של סכום הזכיה?
  - אם האדם י משחק את המשחק 5 פעמים מה התוחלת ומה השונות של ממוצעו סכום הזכיה ביחס לשחקנים?
  - אם האדם י משחק את המשחק 50 פעם מה ההסתברות שבסה"כ יזכה ב-1050 נס ומעלה?

**10)** לפי הערכות הלשכה המרכזית לסטטיסטיקה השכר הממוצע במשק הוא 8000 נס עם סטיית תקן של 3000 נס.  
מה ההסתברות שבמדגם מקרי של 100 עובדים השכר הממוצע יהיה יותר מ-8500 נס?

**11)** קובייה הוטלה 50 פעמים.  
מה ההסתברות שהממוצע של התוצאות יהיה לפחות 3.72?

**12)** אורך צינור שפועל מייצר הינו עם ממוצע של 70 ס"מ וסטיית תקן של 10 ס"מ.  
א. נלקחו באקראי 100 מוטות, מה ההסתברות שסכום 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 המינימלי, כדי שההסתברות של 5% ממוצע המדגם יהיה קטן מ-69 ס"מ. היעזרו במשפט הגבול המركזי.

**13)** נתון משתנה מקרי בדיד בעל פונקציית ההסתברות הבאה:

2	4	6	8	X
$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$P(X)$

מתוך התפלגות זו נלקח מדגם מקרי בגודל 50.  
מה הסיכוי שסכום המדגם יהיה קטן מ-5?

**14)** נתון ש-  $X \sim N(\mu, \sigma^2)$ . דגמו 5 תכפיות מאותה ההתפלגות והתבוננו בממוצע המדגם  $\bar{X}$  לכן:  $P(\bar{X} > \mu)$

- א. 0.
- ב. 0.5
- ג. 1
- ד. לא ניתן לדעת.

**15)** נתון ש-  $X$  מתפלג כלשהו עם תוחלת  $\mu$  ושונות  $\sigma^2$ .  
החליטו לבצע מדגם בגודל 200 מתוך ההפלגות הנתונה לפי משפט הגבול המרכזי מתקיים (בחרו בתשובה הנכונה):

- א.  $X \sim N\left(\mu, \frac{\sigma^2}{200}\right)$
- ב.  $\mu \sim N\left(\mu, \frac{\sigma^2}{200}\right)$
- ג.  $\bar{X} \sim N(\mu, \sigma^2)$
- ד.  $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{200}\right)$

**16)** נתון ש-  $X \sim N(\mu, \sigma^2)$ . אם נדgos  $n$  תכפיות מתוך ההתפלגות ונגידיר:  
אזי (בחרו בתשובה הנכונה):

- א.  $\mu$  ו-  $\bar{X}$  יהיו משתנים מקרים.
- ב.  $\mu$  יהיה משתנה מקרי ו-  $\bar{X}$  קבוע.
- ג.  $\bar{X}$  יהיה משתנה מקרי ו-  $\mu$  קבוע.
- ד.  $\mu$  ו-  $\bar{X}$  יהיו קבועים.

**תשובות סופיות:**

- 1) א. כלל הסטודנטים במכללה שסימנו סטטיסטיקה א. ב. ציון.  
 ג. ממוצע : 78 , סטיית תקן : 15 .  
 ד. 2 .  
 ה. 10.6 .  
 ו. 1 .

א. להלן טבלה :

4	3	2	1	0	X
0.05	0.3	0.35	0.25	0.05	$P(X)$

$$\sigma = 0.973 , \sigma^2 = 0.9475 , \mu = 2.05 .$$

$$\sigma(\bar{X}) = 0.486 , \sigma_{\bar{x}}^2 = 0.2369 , \mu_{\bar{x}} = 2.05 .$$

$$\sigma(\bar{X}) = 1.21 , \mu_{\bar{x}} = 3.5 \quad (3)$$

- .0.1974.ד .0.ג .0.0013.ב .0.8413.א (4)  
 .178.205 .ד .0.9544 .ג .0. .0. .0. א (5)  
 ד. התשובה הייתה קטנה.  
 .0.2628 .ג .27.71 .ב .0.0465 .א (6)  
 .0.5 .ד .0.1587 .ג .0.1587 .ב .0. .0. א (7)  
 .0.6826 .ב .0.5468 .א (8)

א. להלן טבלה :

30	20	10	X
0.5	0.25	0.25	$P(X)$

- ג. התוחלת : 13.75 , השונות : 22.5 .  
 ב. התוחלת : 68.75 , השונות : 22.5 .  
 ד. 0.8997  
 .0.0475 (10)  
 .0.1814 (11)  
 .271 .ג .0.0228 .ב .0.9772 .א (12)  
 .0.5 (13)  
 (14) ב'.  
 (15) ד'.  
 (16) ג'.

## התפלגות פרופורציית ההצלחות במדגם:

**רקע:**

בפרק זה נדון בהתפלגות הדגימה של פרופורציית המדגם.  
א - מספר ההצלחות במדגם (למשל, מספר המובטלים במדגם).

$$\hat{p} = \frac{y}{n} - \text{פרופורציית ההצלחות במדגם.}$$

למשל, שיעור המובטלים במדגם -  $n = 200$  -  
מספר המובטלים :  $Y = 20$ .

$$\text{פרופורציית המובטלים במדגם : } \hat{p} = \frac{20}{200} = 0.1$$

נסמן ב-  $p$  את שיעור ההצלחה באוכלוסייה וב-  $q$  את שיעור הכישלונות באוכלוסייה.  
נבע מדגם מקרי (הנחה שהתצפויות בלתי תלויות זו בזו) ונتابון בהתפלגות של  
פרופורציית המדגם.

**התוחלת, השונות וסטיית התקן של פרופורציית המדגם:**

$$E(\hat{p}) = p , V(\hat{P}) = \frac{pq}{n}$$

משפט הגבול המרormalי עבור הפרופורציה המדגמית :

$$\text{אם : } Z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} . \hat{p} \sim N\left(p, \frac{pq}{n}\right) , np \geq 5 \& nq \geq 5 , \text{ אז :}$$

**הערות:**

- התנאים לקרוב הנורמללי הם נזילים, כלומר משתנים ממרכה אחד לשני.  
התנאי שהציגי כאן הוא הפופולרי ביותר :

$$1. n \cdot p \geq 5$$

$$2. n \cdot (1-p) \geq 5$$

- ישנים מרצים שנוטנים את התנאי המחייב הבא :

$$1. n \cdot p \geq 10$$

$$2. n \cdot (1-p) \geq 10$$

- וישנים מרצים המשתמשים בתנאי:  $(n \geq 30)$ .

- תאלצו לבדוק מהו התנאי שנותנו לכם בכיתה כדי לעبور לנורמלית.

- כיוון שפרופורציה אינה חייבות להיות מספרשלם בהכרח לא נהוג לבצע כאן תיקון רציפות.

**דוגמה (פתרון בהקלטה):**

לפי נתוני משרד החינוך בעיר ירושלים ל- $60\%$  מתלמידי התיכון זכאים לתעודת בגרות. נדגו  $200$  תלמידי תיכון.

- מה ההסתברות שהשכיחות היחסית ( $\hat{p}$ ) של הזכאים לבגרות במדגם עלה על  $60\%$ ?
- מה ההסתברות שפרופורציות הזכאים לבגרות במדגם עלה על  $70\%$ ?

## שאלות:

- 1)** במדינה מסוימת 10% מכלל האוכלוסייה הינם מובטלים. נדגוו באקראי 140 אנשים מהמדינה.
- מה התוחלת ומהי השונות של פרופורציות המובטלים שנדגוו?
  - מה ההסתברות שבמדגם לפחות 10% יהיו מובטלים?
  - מה ההסתברות שלכל היוטר 9% מהמדגם יהיו מובטלים?
- 2)** נניח כי 30% מהאוכלוסייה תומכים בהצעת חוק מסוימת. אם נדגוו מהאוכלוסייה 200 איש. חשבו את ההסתברויות הבאות :
- פחות 35% יתמכו בהצעת החוק במדגם.
  - כל היהת 25% יתמכו בהצעת החוק במדגם.
  - יותר מ-27% יתמכו בהצעת החוק במדגם.
- 3)** לפי נתוני משרד התקשרות 40% מהאוכלוסייה מחזיקים בטלפון נייד מסווג "סמארטפון". נדגוו 400 אנשים מהאוכלוסייה.
- מה ההסתברות שבמדגם לכל היוטר ל-40% יש סמארטפון?
  - מה ההסתברות שבמדגם לרוב יש סמארטפון?
  - מה ההסתברות שפרופורציית בעלי הסמארטפון במדגם תסטה מהפרופורציה באוכלוסייה ללא יותר מ-4%?
  - כיצד התשובה לשיעיף הקודם הייתה משתנה אם היינו מגדילים את גודל המדגם?
- 4)** נתון כי 80% מבתי האב מחוברים לאינטרנט. נדגוו 400 בתים אב אקרים.
- מה ההסתברות שלפחות 340 מהם מחוברים לאינטרנט?
  - מה ההסתברות שפרופורציית המחוברים לאינטרנט במדגם תסטה מהפרופורציה האמיתית ביותר מ-4%?
  - כמה בתים אב יש לדוגם כדי שהסתטיה בין הפרופורציה המדגמית לפרופורציה האמיתית לא תעלה על 3% בהסתברות של 90%?
  - מהו העשירון התיכון של התפלגות פרופורציית המדגם?
- 5)** נתון שציוני פסיקומטרי מתפלגים נורמלית עם תוחלת 500 וסטיית תקן 100. ב"מועדון ה-700" נכללים נבחנים שמקבלים ציון מעל 700 בפסיכומטרי. מה הסיכוי שבמועדון בו נבחנו 2000 נבחנים אקרים יהיו לפחות 3% המשתייכים למועדון?

6) נתון ש- $\hat{P} = \frac{X}{n}$ , ונגידיר את המשתנה הבא :

$$\text{א. הוכיחו ש } E(\hat{P}) = p, V(\hat{P}) = \frac{p(1-p)}{n}$$

ב. מה  $p$  המביא את  $V(\hat{P})$  להיות מקסימום?

### תשובות סופיות:

1) א. התוחלת : 0.1, השונות : 0.00064. ב. 0.5.

2) א. 0.0618. ב. 0.0618. ג. 0.8238.

3) א. 0.5. ב. 0.8968. ג. 0.8.

4) א. 0.0062. ב. 0.0456. ג. 0.481. ד. 0.77436.

5) 0.0154

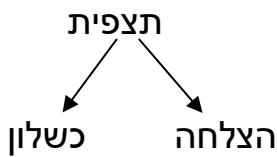
6) א. שאלת הוכחה. ב. 0.5.

## התפלגות מספר ההצלחות במדגם – קירוב נורמלי להתפלגות הבינומית:

**רקע:**

### תזכורת על התפלגותBINOMIAL:

בפרק זה נדונו בהתפלגות מספר ההצלחות במדגם אקראי (תצפויות בלתי תלויות זו בזו).  
את מספר ההצלחות במדגם מסמן ב-  $Y$ .  
מחלקים כל תצפית במדגם להצלחה או כישלון.



כעת מה שמשתנה מתצפית הוא משתנה דיכוטומי (משתנה שיש לו שני ערכיים).  
הסיכוי להצלחה מסומן עם הפרמטר  $p$  וכישלון מסומן ע"י הפרמטר:  $q = 1 - p$ .  
מבצעים מדגם אקראי בגודל  $n$ :  $Y \sim B(n, p)$ .

$$\text{פונקציית ההסתברות של ההתפלגות הבינומית היא: } p(y=k) = \binom{n}{k} p^k q^{n-k}$$

$$\text{תוחלת: } E(y) = np$$

$$\text{שונות: } V(y) = npq$$

### קירוב נורמלי עבור התפלגותBINOMIAL:

אם לפניו התפלגותBINOMIAL:  $Y \sim B(n, p)$ , ומתקיים ש:

$$\begin{aligned} 1. \quad & n \cdot p \geq 5 \\ 2. \quad & n \cdot (1-p) \geq 5 \\ & y \sim N(np, npq) \\ & Z_y = \frac{y - np}{\sqrt{npq}} \end{aligned}$$

אנו:

**תיקון רציפות:**

כאשר משתמשים בקירוב הנורמלי להתפלגות הבינומית יש לבצע תיקון רציפות. הסיבה שעוברים כאן מההתפלגות בדידה להתפלגות נורמלית שהיא התפלגות רציפה. על פי הכללים הבאים :

$$\cdot p(Y = a) \cong p\left(a - \frac{1}{2} \leq Y \leq a + \frac{1}{2}\right) . 1$$

$$\cdot P(Y \leq a) \cong P(Y \leq a + 0.5) . 2$$

$$\cdot P(Y \geq a) \cong P(Y \geq a - 0.5) . 3$$

**הערות:**

- התנאים למעבר מבינומי לנורמלי הם נזילים, כלומר משתנים ממרצה אחד לשני. התנאי שהציגי כאן הוא הפופולרי ביותר :

$$\cdot n \cdot p \geq 5 . 1$$

$$\cdot n \cdot (1-p) \geq 5 . 2$$

- ישנו מרצים שנוטנים את התנאי המחייב הבא :

$$\cdot n \cdot p \geq 10 . 1$$

$$\cdot n \cdot (1-p) \geq 10 . 2$$

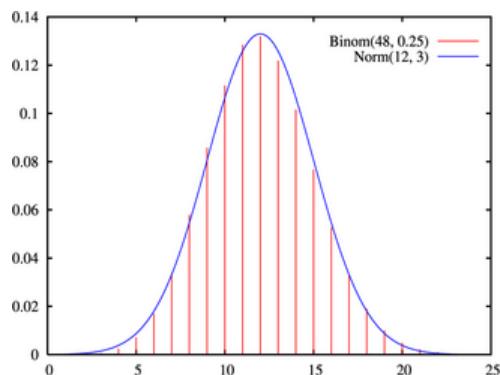
- וישנו מרצים שה坦אי שהם נתונים הוא :  $(n \geq 30)$ .

- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעبور מההתפלגותBINOMIAL לנורמלית.

- הערכה נוספת היא לגבי תיקון רציפות. ישנו מרצים שלא מחיברים לבצע תיקון רציפות שהמדגמים גדולים (בדרך כלל מעל 100 תצפיות) בפתרונות שאציג תמיד אבצע תיקון רציפות במעבר מבינומי לנורמלי כיון שכח הפתרון יהיה יותר מדויק (בכל מקרה שהמדגמים גדולים העניין זניח).

**דוגמה (הפתרון בהקלטה):**  
 נתון שבקרב אוכלוסיית הנוער 25% זוקקים למשקפיים. נדגוו באקראי 48 בני נוער.

1. מה הסיכוי שבדיווק 14 מהתוכם יהיו זוקקים למשקפיים?
2. מה הסיכוי שלכל היוטר 13 מהתוכם זוקקים למשקפיים?



**שאלות:**

**1)** נתון ש-20% מאוכלוסייה מסוימת אקדמאית. נבחרו באקראי 10 אנשים באותה אוכלוסייה.

א. מה ההסתברות שלשלושה מהם אקדמיים?

ב. מה ההסתברות שלכל היותר אחד מהם אקדמי?

ג. מה התוחלת ומהי סטיית התקן של מספר האקדמאים במדגם?

**2)** בפועל 10% מהמושרים פגומיים. נלקחו 100 מושרים באקראי מקו הייצור.

א. מה ההסתברות שנציגו לפחות 6 מושרים פגומיים?

ב. מה ההסתברות שמספר המושרים הפגומיים יהיה לכל היותר 11 במדגם?

**3)** ציוני פסיקומטרי בקרבת הנרשמים למוסד מסוים מתפלגים נורמלית עם ממוצע 500 וסטיית תקן 100. למוסד מסוים הוחלט לקבל אך ורק סטודנטים שקיבלו מעל 600 בפסיכומטרי. 100 סטודנטים אקדמיים נרשמו למוסד. מה ההסתברות שלפחות 20 יתקבלו?

**4)** מטילים מטבח 50 פעמים.

א. מה ההסתברות לקבל לכל היותר 30 ע齊ים?

ב. מה ההסתברות לקבל 28 ע齊ים לפחות הבינומית ולפי הקירוב הנורמלי?

**5)** במתוס מקום ל-400 נוסעים. נרשמו לטיסה 430 אנשים (overbooking). מנתונים סטטיסטיים ידוע שהסיכון שאדם שנרשם לטיסה אכן יגיע הוא 0.9.

א. מה ההסתברות שלא יהיו מקומות ישיבה לכל האנשים שהגיעו לטיסה?

ב. מה צריך להיות גודל המתוס כדי שבסיכוי שלפחות 95% המתוס יספק לכמות הנרשמים?

**6)** מפעלי לייצור ארטיקים טוען שהסיכון שארטיק שהוא מייצר יהיה פגום הוא 0.01. מוכר הזמן 1000 ארטיקים מהמפעל. מה ההסתברות שהמוכר קיבל לפחות 980 ארטיקים תקינים אם טענת המפעל מוצדקת?

**7)** מהמר מטיל קובייה הוגנת 100 פעמים. בכל הטלה, אם מתקבל תוצאה זוגית בקובייה המהמר זוכה בשקל. אחרת, המהמר משלם שקל. המהמר הטיל את הקובייה 100 פעמים מה הסיכון שהרווח של המהמר יהיה לכל היותר 10?

### תשובות סופיות:

- |         |                                               |         |               |
|---------|-----------------------------------------------|---------|---------------|
| .1.2649 | ג. התוחלת : 2, סטיית התקן :                   | .0.3758 | ב. 0.201 (1)  |
|         |                                               | .0.6915 | ב. 0.9332 (2) |
|         |                                               |         | .0.1611 (3)   |
|         | ב. בינומית - 0.0788, קירוב לנורמלית - 0.0778. |         | א. 0.9406 (4) |
|         |                                               | .398    | ב. 0.015 (5)  |
|         |                                               |         | .0.9996 (6)   |
|         |                                               |         | .0.8643 (7)   |

# bijustisjka le-refoat shinim

פרק 29 - מושגי יסוד באמידה

תוכן העניינים

1. כללי .....

130 .....

## מושגי יסוד באמידה:

### רקע:

כזכור מהפגש הקודם, פרמטר הוא גודל המתאר את האוכלוסייה או התפלגות מסויימת. כמו ממוצע הגבאים בקרוב מתגisiים לצה"ל -  $\mu$ .

כמו פרופורצית התומכים במשלה בקרוב אזרחי המדינה -  $p$ .

בדרכ כל הפרמטרים הם גדלים שאינם ידועים באמת, ולכן מוצאים מוגדים במטרה לאמוד אותם. אין אפשרות לחשב אותם הניסיון הוא בהערכתו כמה הם שווים ככל שניתן.

- נסמן באופן כללי פרמטר באות  $\theta$  ואומד ב- $\hat{\theta}$ .  $\hat{\theta}$  הוא סטטיסטי המוחשב על המוגדים ובאמצעותו נאמוד את  $\theta$ .
- שגיאת אמידה:  $|\hat{\theta} - \theta|$  - ההפרש בין האומד לאמת (הפרמטר).

### דוגמה (פתרון בהקלטה):

בכנסת ה-19 קיבלת מפלגת העבודה 15 מנדטים. בערוץ 10 ברגע סגירת הקלפיות הערכו את מספר המנדטים של המפלגה להיות 17 מנדטים וזאת על סמך תוצאות מוגדים של העורך.

- א. מה הפרמטר בדוגמה זו?
- ב. מהי טעות האמידה של ערוץ 10?
- $\hat{\theta}$  יהיה אומד חסר הטיה ל- $\theta$  אם התוחלת של  $\hat{\theta}$  תהיה שווה ל- $\theta$ :  $E(\hat{\theta}) = \theta$ .
- טעות התקן של אומד היא סטיית התקן שלו, כלומר:  $\sigma(\hat{\theta}) = S.E.$

**פרמטרים מרכזיים ואומדיים שלחה:****ממוצע האוכלוסייה  $\mu$ :**

$$\text{האומד הנקודתי שלו יהיה: ממוצע המדגמים: } \bar{x} = \frac{\sum x}{n}$$

$$\text{. } \sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}} = SE \text{ . } E(\bar{x}) = \mu \text{ . } \text{כמו כן, טעות התקן: } \mu$$

**פרופורציה באוכלוסייה  $p$ :**

$$\text{האומד הנקודתי שלו יהיה: פרופורציה במדגם: } \hat{p} = \frac{y}{n}$$

$$\text{. } \sigma(\hat{P}) = \sqrt{\frac{p \cdot (1-p)}{n}} \text{ . } E(\hat{P}) = p \text{ . } \text{כמו כן טעות התקן: } p$$

**שונות האוכלוסייה  $\sigma^2$ :**

$$\text{האומד הנקודתי שלו יהיה: } S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$\text{. } S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} \text{ . } \sigma^2 \text{ . } E(S^2) = \sigma^2 \text{ . } \text{ולכן } S^2 \text{ הינו אומד חסר הטיה ל- } \sigma^2$$

**הערה:** אומד הוא הנוסחה הכללית לאמידת הפרמטר ואומדן הוא הערך הספציפי שהתקבל במדגם מסוים.

**דוגמה (פתרון בהקלטה):**

נdagmo 10 משפחות בתל אביב ונבדק עבור כל משפחה מספר הילדים שלה.  
להלן התוצאות שהתקבלו: 3, 1, 3, 2, 1, 4, 5, 2, 1, 2. אמדדו באמצעות אומדיים חסרי הטיה את הפרמטרים הבאים:

1. ממוצע מספר הילדים למשפחה בתל אביב.
2. שונות מספר הילדים למשפחה בתל אביב.
3. פרופורציית המשפחות בנות שני ילדים.

**שאלות:**

- 1)** מתוך 500 טירונים, נמצאו 120 בעלי שברי הליכה. נתנו שהטיסוי שטירון יהיה עם שבר הליכה הוא 0.25.
- מהי האוכלוסייה המוצגת בשאלת? מהם הפרמטרים שלה?
  - מהי טעות התקן של האומד כשהמדגם בגודל 500?
  - מהו האומדן לפרמטר?
  - מהי טעות האמידה?
- 2)** לפי נתונים היכרנו, מקרר צורך ממוצע 2400 וואט לשעה עם סטיית התקן של 500 וואט לשעה.
- במדגם של 25 מקרים של היכרן התקבל ממוצע של 2342 וואט לשעה.
- מהי האוכלוסייה המוצגת בשאלת? מהם הפרמטרים שלה?
  - מהי טעות התקן של האומד?
  - מהו האומדן לפרמטר?
  - מהי טעות האמידה?
- 3)** נדגו עשרה מתגיים לzech"l. גובהם נמדד בס"מ. להלן התוצאות שהתקבלו: 168, 184, 192, 180, 171, 177, 187, 168, 177 ו-175.
- מצאו אומדן חסר הטיה לגובה הממוצע של מתגייסי zech"l.
  - מצאו אומדן חסר הטיה לשונות הגבהים של מתגייסי zech"l.
  - מצאו אומדן חסר הטיה לפ羅פּוֹרְצִיוֹת המתגיים בגובה של לפחות 180 ס"מ.
- 4)** נדגו 20 שכירים באקראי. עברו כל שכיר נמדד השכר באלפי שקלים.
- להלן התוצאות שהתקבלו:  $\sum_{i=1}^{20} X_i^2 = 1502.2$ ,  $\sum_{i=1}^{20} X_i = 162$
- AMDו את השכר הממוצע של השכירים במשק.
  - AMDו את סטיית התקן של שכר השכירים במשק.
- 5)** במטרה לאמוד את ממוצע האוכלוסייה, נדגו תכיפות בלתי תלויות מהאוכלוסייה וחישבו את הממוצע שלהם. מהי טעות התקן?
- סטיית התקן של האוכלוסייה.
  - סטיית התקן של ממוצע האוכלוסייה.
  - סטיית התקן של המדגם.
  - סטיית התקן של ממוצע המדגם.

6) משקל הממוצע של אוכלוסייה מסוימת הוא 75 ק"ג עם שונות של 25 . אם יבחרו כל המדגמים האפשריים בגודל 10 מאוכלוסייה זו סטיית התקן של ממוצעי המדגמים תהיה :

- .א. 3.
- .ב. 2.5
- .ג. 1.581
- .ד. אין מספיק נתונים לדעת.

7) במדגם מקרי, متى סכום ריבועי הסטיות מהממוצע,  $\sum_{i=1}^n (x_i - \bar{x})^2$ , מחולק ב- $n-1$ ?  
 א. כאשר  $n$  קטן.  
 ב. כאשר תצפיות המדגם אינן בלתי תלויות.  
 ג. כאשר האוכלוסייה אינה מתפלגת נורמללית.  
 ד. כאשר מעוניינים באומד חסר הטיה לשונות האוכלוסייה ממנה הוצאה המדגם.  
 ה. כאשר מעוניינים לחשב את שונות התפלגות הדגימה של ממוצע המדגם.

8) מדגם מקרי מתוך אוכלוסייה בעלת ממוצע  $\mu$  לא ידוע

ושונות :  $s^2 = 64$ . טעות התקן של האומד ל-  $\mu$  היא :

- .א. 16.
- .ב. 8.
- .ג. 4.
- .ד. 2.

9) מהו אומד חסר הטיה?

- א. אומד שערכו שווה לממוצע התפלגות הדגימה שלו.
- ב. אומד שערכו שווה לערך הפרטר באוכלוסייה.
- ג. אומד שממוצע התפלגות הדגימה שלו שווה לערך הפרטר באוכלוסייה.
- ד. אומד שהסיכוי שערכו יהיה גבוה מערך הפרטר באוכלוסייה שווה לשינוי שיהיה נזוק ממנו.

**תשובות סופיות:**

- (1) א. 0.25      ב. 0.19      ג. 0.24      ד. 0.01
- (2) א. אוכלוסייה: מקרים של יצרן, תוחלת: 2400, סטיית תקן: 500.  
                 .58      .2342      ג.      ב. 100
- (3) א. 0.4.ג      ב. 64.1      ג. 177.9  
                 .3.16.ב      ב. 8.1
- (4) א.      ב.      ג. 5  
                 ד.      ג. 6  
                 ד. 7  
                 ד. 8  
                 ג. 9

# bijustisjka le-refoat shinim

פרק 30 - רוח סמרק לתחולת (מומוץ)

## תוכן העניינים

135 .....	1. רוח סמרק כשינויים האוכלוסייה ידועה
140 .....	2. קביעת גודל מוגם
142 .....	3. רוח סמרק כשינויים האוכלוסייה לא ידועה

## רוח סמך כשינויות האוכלוסייה ידועה:

**רקע:**

ממוצע המדגם הוא אומד לממוצע האוכלוסייה, אך לא באמת ניתן להבין ממנו על גודלו של ממוצע האוכלוסייה. ההסתברות שממוצע המדגם יהיה בדיקות כמו הממוצע האמתי הוא אפסי.

מה שנחוג לעשות כדי לאמוד את ממוצע האוכלוסייה, זה לבנות רוח סמך.

בנייה מרוחה בטחון שהסיכוי שהפרט  $\mu$  ייכל בתוכו הוא:  $1 - \alpha$ .

$\alpha - 1$  : נקרא רמת בטחון או רמת סמך. כך ש:  $\alpha - 1 = P(A \leq \mu \leq B)$ .

$A$  - גבול תחתון של רוח הסמך.

$B$  - הגבול העליון של רוח הסמך.

$L = B - A$  - אורך רוח הסמך.

**דוגמה (פתרון בהקלטה):**

חווק דגם 25 חיילים שנבחנו ב מבחון הפסיכומטרי. הוא בנה רוח סמך לממוצע הציונים ב מבחון הפסיכומטרי ב קרב אוכלוסיית החיילים ו קיבל בין 510 ל-590. רוח הסמך בונה ברמת סמך של 95%.

1. מהי אוכלוסיית המחקר?
2. מה המשתנה באוכלוסייה?
3. מה הפרט שהחווק רצה לאמוד?
4. מהו רוח הסמך?
5. מה אורך רוח הסמך?
6. מהי רמת הביטחון של רוח הסמך?

בפרק זה נרצה לבנות רוח סמך לתוחלת ( $\mu$ ) במקהה ש- $\sigma^2$  (שונות האוכלוסייה) ידועה. פרמטר אותו נרצה לאמוד:  $\mu$ .

אומד נקודתי:  $\bar{x}$ .  
תנאים לבניית רוח הסמך:  $N \sim X$  או  $n \geq 30$ .

$\sigma^2$  (שונות האוכלוסייה) ידועה.

$$\text{נוסחה לרוח הסמך: } \bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

#### דוגמה (פתרון בהקלטה):

על פי נתונים היצרנו אורך חיי סוללה מתפלג נורמלית עם סטיית תקן של 1 שעה. מעוניינים לאמוד את תוחלת חיי סוללה. נציגו באקראי 4 סוללות, אורך החיים הממוצע שהתקבל הוא 13.5 שעות. בנו רוח סמך ברמת סמך של 95% לתוחלת אורך חיי סוללה.

$$\text{שגיאת האמידה המקסימלית: } \varepsilon = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

ע- נוthen את שגיאת האמידה המקסימלית, דבר שנראה גם טעות סטטיסטית, טעות דגימה.

#### דוגמה (פתרון בהקלטה):

בשימוש לשאלת עם הסוללות. מה ניתן להגיד בביטחון של 95% על שגיאת האמידה?

קשרים מתמטיים ברוח הסמך:

- אורך רוח הסמך הוא פערם שגיאת האמידה המקסימלית:  $L = 2\varepsilon$ .
- ממוצע המדגמים נופל תמיד באמצע רוח הסמך:  $\bar{X} = \frac{A+B}{2}$ .
- ככל שמספר התצפיות ( $n$ ) גבוהה יותר, כך יש יותר אינפורמציה ולכן האומד יותר מדויק, ולכן מקבל רוח סמך יותר קצר.
- ככל שרמת הביטחון ( $\alpha - 1$ ) גבוהה יותר, כך:  $\varepsilon = z_{1-\frac{\alpha}{2}}$  גבוהה יותר, ורוח הסמך יותר ארוך.

**שאלות:**

- 1)** חוקר התענין למד את השכר המומוצע במשק. על סמך מוגם הוא קבע שבביטחון של 95% כי השכר המומוצע במשק נع בין 9200 ל-9800 ₪.
- מי האוכלוסייה במחקר?
  - מה המשנה הנחקר?
  - מה הפרמטר שאוטו רוצים למד?
  - מה רוח הסמך לפרמטר?
  - מה רמת הסמך לפרמטר?
  - מה אורך רוח הסמך?
  - מה הסיכוי שטעות הדגימה תעלה על 300 ₪?
- 2)** מעוניינים למד את התפוקה היומיית המומוצעת של מפעל מסוים ברמת סמך של 95%. בדוגמאות אקראי של 100 ימים התקבלה תפוקה ממוצעת 4950 מוצרים ביום. לצורך פתרון הנח שטית התקן האמצעית ידועה ושויה 150 מוצרים ביום. בנו את רוח הסמך.
- 3)** מעוניינים למד את ממוצע אורך החיים של מכשיר. מנתוני היצרך ידוע שאורך החיים מתפלג נורמלית עם סטיית תקן של 20 שעות. נגמו 25 מכשירים ונמצא כי ממוצע אורך החיים שלהם היה 230 שעות.
- בנו רוח סמך ברמת סמך של 90% לאורך החיים המומוצע של מכשיר.
  - בנו רוח סמך ברמת סמך של 95% לאורך החיים המומוצע של מכשיר.
  - הסבירו כיצד ומדוע השתנה רוח הסמך.
- 4)** נגמו 200 עובדים מהמשק הישראלי. השכר המומוצע שלהם היה 9700 ₪. נניח שטית התקן של השכר במשק היא 3000 ₪.
- בנו רוח סמך ברמת סמך של 95% לתוכלת השכר במשק.
  - מה ניתן לומר בביטחון של 95% על הסטייה המרבית בין ממוצע המוגם לתוכלת השכר?
  - מה היה צריך להיות גודל המוגם אם היו רוצחים להקטין את רוח הסמך ב-50%?
  - אם היינו מגדילים את גודל המוגם ובונים רוח סמך באותה רמת סמך האם היה ניתן לטעון בביטחון רב יותר שרוח הסמך מכיל את הפרמטר?

- 5) בנו רוח סמך לממוצע הציוניים של מבחן אינטלייגנציה. ידוע שסטיטית התקן היא 15 והמדד מtabסס על 100 תוצאות. רוח הסמך שהתקבל הוא (105,99). שזרו את :
- ממוצע המדגמים.
  - שגיאת האמידה המקסימאלית.
  - רמת הסמך.
- 6) זמן החלה מאנגינה מתפלג עם סטיטית התקן של יומיים. חברת תרופות מעוניינת לחקור אנטיביוטיקה חדשה שהיא פיתחה. במחקר השתתפו 60 אנשים שחלו באנגינה וקיבלו את האנטיביוטיקה החדשה. בממוצע הם החלימו לאחר 4 ימים.
- בנו רוח סמך לתוחלת זמן ההחלמה תחת האנטיביוטיקה החדשה ברמת סמך של 90%.
  - מה הייתה קורה לאורך רוח הסמך אם היה תקציב להגדלת גודל המדגמים פי 4? הסבירו.
  - מה הייתה קורה לאורך רוח הסמך אם היו בונים את רוח הסמך ברמת סמך גדולה יותר? הסבירו.
- 7) חוקר בנה רוח סמך לממוצע וקיבל את רוח הסמך הבא :  $\mu = 82$ . נתון שסטיטית התקן בהתפלגות שווה ל-10 ושהמדד מtabסס על 16 תוצאות. התפלגות המשתנה היא נורמללית.
- מהו ממוצע המדגמים?
  - מהי רמת הסמך של רוח הסמך שנבנה?
  - מה הסיכוי ששגיאת האמידה באמידת ממוצע האוכלוסייה תעלה על 5?
- 8) חוקר בנה רוח סמך לתוחלת כאשר השונות בהתפלגות ידועה ברמת סמך של 95%. אם החוקר כעת יבנה על סמך אותו נתונים נתונים רוח סמך ברמת סמך קטנה מ-95%, איזה מהמשפטים הבאים לא יהיה נכון.
- אורך רוח הסמך החדש יהיה קטן יותר.
  - גודל המדגמים יהיה כעת קטן יותר.
  - הMargin בין ממוצע המדגם לקצota רוח הסמך יהיה קטנים יותר ברוחה הסמך החדש.
  - רמת הביטחון לבנות רוחה הסמך החדש תהיה קטנה יותר.

9) חוקר בנה רוח סמך ל-  $\mu$  וקיבל:  $48 < \mu < 54$ . מה נכון בהכרח:

- א.  $\mu = 51$ .
- ב.  $\bar{X} = 6$ .
- ג.  $\bar{X} = 51$ .
- ד. אורך רוח הסמך הינו 3.

10) אייזה מהגורמים הבאים אינם משפיע על גודלו של רוח בר סמך, כאשר שונות האוכולוסייה ידועה (בחרו בתשובה הנכונה):

- א. רמת הביטחון.
- ב. סטיית התקן באוכולוסייה.
- ג. מספר המשתתפים.
- ד. סטיית התקן במדגם.

### תשובות סופיות:

- |                               |                                                                   |
|-------------------------------|-------------------------------------------------------------------|
| 1) א. העובדים במשק.           | ב. שכר ב-₪.                                                       |
| ה. $0.95 < \mu < 0.600$       | ג. $0.95 < \mu < 0.600$                                           |
| 2) $4979.4 < \mu < 4920.6$    |                                                                   |
| 3) א. $236.58 < \mu < 223.42$ | ב. $237.84 < \mu < 222.16$                                        |
| ג. ראה סרטוון.                |                                                                   |
| 4) א. $10,116 < \mu < 9284$   | ב. הסטיה המרבית בין $\bar{x}$ ל- $\mu$ היא 416 נס בביטחון של 95%. |
| ג. לא.                        | ד. $0.800$ .                                                      |
| 5) א. $0.102 < \mu < 0.09544$ | ב. $0.3 < \mu < 0.09544$                                          |
| ג. גדול.                      | ד. יקטן פי 2.                                                     |
| 6) א. $4.42 < \mu < 4.835$    | ב. $4.42 < \mu < 4.835$                                           |
| 7) א. $0.87 < \mu < 0.9544$   | ב. $0.5 < \mu < 0.87$                                             |
| 8) ב'.                        |                                                                   |
| 9) ג'.                        |                                                                   |
| 10) ד'.                       |                                                                   |

## קביעת גודל מוגן:

**רקע:**

אם מעוניינים לאמוד את ממוצע האוכלוסייה כאשר סטיטית התקן של האוכלוסייה ידועה:  $\sigma$  ברמת סמך של  $\alpha=1$  ושיגיאת אמידה שלא עלתה על  $\epsilon$  מסויים, נציב

$$n \geq \left( \frac{z_{1-\frac{\alpha}{2}} \cdot \sigma}{\epsilon} \right)^2$$

בנוסחה הבאה:

כדי להציג בנוסחה צריך שהמשתנה הנחקר يتפלג נורמלית או שהמוגן ייצא בגודל של לפחות 30 תצפיות.

**דוגמה (פתרון בהקלטה):**

חברת תעופה מעוניינת לאמוד את תוחלת משקל המטען של נוסע. נניח שמשקל מטען של נוסע מתפלג נורמלית עם סטיטית התקן של 2 ק"ג. כמה נוסעים יש לדוגם אם מעוניינים שבבביחוון של 98% הסטייה המרבית בין ממוצע המוגן לממוצע האמתי לא עליה על 0.5 ק"ג? (תשובה: 87).

**שאלות:**

- (1)** משתנה מקרי מתפלג נורמללית עם סטיטית תקן ידועה 12. מה צריך להיות גודל המדגם כדי לבנות רוח סמך ברמת סמך של 98% שאורכו לא עולה על 2?
- (2)** מעוניינים לאמוד את הדופק הממוצע של מתגייסים לצבאי. מעוניינים שבביטחון של 95% שגיאת האמידה המרבית תהיה 0.5. נניח שהדופק מתפלג נורמלית על סטיטית תקן של 3 פעימות לדקה.
- כמה מתגייסים יש לדוגום?
  - אם ניקח מדגם הגדל פי 4 מהמדד של סעיף א' ונאמוד את הממוצע באותה רמת סמך כיצד הדבר ישפייע על שגיאת האמידה?
- (3)** יהיו  $X$  משתנה מקרי עם ממוצע  $\mu$  וסטיטית תקן  $\sigma$ . חוקר רוצה לבנות רוח בר סמך ל- $\mu$  ברמת ביטחון של 0.95, כך שהאורך של הרוח יהיה  $\sigma = 0.5\sigma$ . מהו גודל המדגם הנדרש?

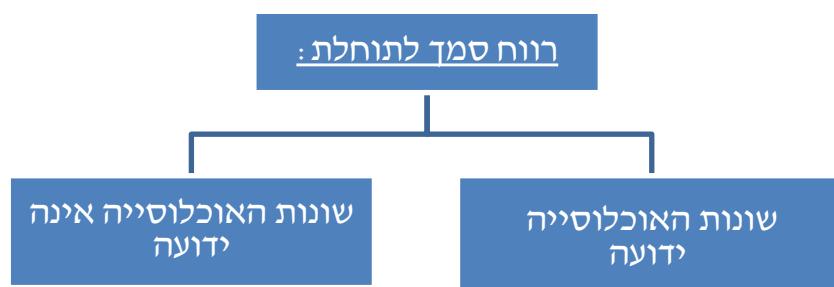
**תשובות סופיות:**

- (1) .780  
 (2) א. 139  
 (3) .  $n = 62$
- ב. הדבר יקטין את  $\sigma$  פי 2.

## רוח סמך כשינויות האוכלוסייה לא ידועה:

רקע:

בבואהנו לבנות רוח סמך לתוחלת אנו צריכים להתמקד בשני המצביעים הבאים:

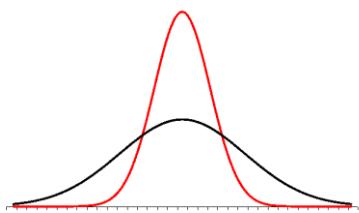


בפרק זה עוסק במקרה ששינויות האוכלוסייה  $(\sigma^2)$  אינה ידועה לנו.

מקרה יותר פרקטטי.

התנאי:  $N \sim X$  או שהמדגם גדול.

$$\text{רוח סמך: } \bar{X} \pm t_{1-\frac{\alpha}{2}}^{(n-1)} \cdot \frac{S}{\sqrt{n}}$$



$$\text{האומד לשונות: } S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$$

התפלגות T:

הינה התפלגות סימטרית בעומניות שהתוחלת שלה היא 0. ההתפלגות דומה להתפלגות Z רק שהיא רחבה ולכון הערכים שלה יהיו יותר גבוהים. התפלגות T תלויות במושג שנקרוא דרגות החופש. דרגות החופש הן:  $df = n - 1$ . ככל שדרגות החופש עולות התפלגות הופכת להיות יותר גבוהה וצרה. שדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה (פתרון בהקלטה):

הזמן שלוקח לפתרון שאלה מסוימת ביחסו מתפלג אצל תלמידי כיתה ח' נורמלית. במטרה לאמוד את תוחלת זמן הפתרון נדגומו 4 תלמידים בכיתה ח'. להלן התוצאות שהתקבלו בדקות: 4.7, 5.2, 4.6, 5.3. בנו רוח סמך ברמת סמך של 95% למומוצע זמן הפתרון לשאלת קרב תלמידי כיתה ח'.

**שאלות:**

- 1)** מחקר מעוניין לדעת כיצד תרופה מסוימת משפיעה על קצב פעימות הלב.  
ל-5 אנשים שנטלו את התרופה מדדו את הדופק והתקבל מספר פעימות לדקה : 84, 84, 88, 79, 89.  
הערה : לצורך פתרון הנח שקצב פעימות הלב מתפלג נורמלית בקירוב.  
א. בנו רוח סמך ברמת סמך של 95% לתוחלת הדופק של נוטלי התרופה הניל.  
ב. נתון שהדופק הממוצע ללא לקיחת התרופה הינו 70. לאור זאת, האם בביטחון של 95% התרופה משפיעה על הדופק?  
ג. בהמשך לסעיף א', אם היינו בונים את רוח הסמך ברמת ביטחון של 99%, כיצד הדבר היה משפיע על רוח הסמך?
- 2)** במדגם שנעשה על 25 מתגייסים לצבא האמריקאי התקבל כי גובה ממוצע של חיל הינו 178 ס"מ עם סטיית תקן :  $S = 13$  ס"מ.  
בנו רוח סמך ברמת סמך של 90% לתוחלת גובה המתגייסים לצבא האמריקאי. מה יש להניח לצורך פתרון?
- 3)** אדם מעוניין לאמוד את זמן הנסיעה הממוצע שלו לעבודה. לצורך כך הוא דוגם 5 ימים שזמן הנסעה בהם בבדיקות הוא : 27, 34, 32, 40, 30.  
א. ברמת ביטחון של 95% אמוד את זמן הנסעה הממוצע. מהי ההנחה הדורשahn לorzuch פתרון?  
ב. איך גודל רוח הסמך היה משתנה אם היו דוגמים עוד ימים?
- 4)** ציוני מבחן אינטיליגנציה מתפלגים נורמלית. נדגו 25 מבחנים והתקבל ממוצע ציונים 102 וסטיית תקן מדגמית 13.  
א. בנו רוח סמך לממוצע הציונים באוכלוסייה ברמת ביטחון של 95%.  
ב. חזרו על סעיף א' אם סטיית התקן הינה סטיית התקן האמתית של כלל הנבחנים.  
ג. הסבירו את ההבדלים בין שני השעיפים הניל.
- 5)** נשקלו 60 תינוקות אשר נולדו בשבועות-40 של החירון. המשקל נמדד בKİLOGRAMS. להלן התוצאות שהתקבלו :  $\sum_{i=1}^{60} X_i^2 = 643.19$  ,  $\sum_{i=1}^{60} X_i = 195$ .  
בנו רוח סמך ברמת סמך של 95% לתוחלת משקל תינוק ביום היולדו.

- 6) נדגו 120 אנשים אקראים מעל גיל 50. עבור כל אדם נבדק מספר שנות השכלתו. להלן התוצאות שהתקבלו :  $S = 2$ ,  $\bar{x} = 13.8$ .  
בנו רוח סמך ברמת סמך של 96% למומוצע ההשכלה של אזרחים מעל גיל 50.
- 7) שני סטטיסטיקים בנו רוח בר-סמך לאותו פרמטר  $\mu$ .  
 לכל אחד מהסטטיסטיקים מוגם אחר, אך באותו גודל 10.  
 שניהם קבעו אותה רמת סמך.  
 סטטיסטיκי א : הניח  $20 = \sigma$ .  
 סטטיסטיκי ב : חישב לפיה המוגם וקיבל  $20 = S$ .  
 למי משני הסטטיסטיקים יהיה רוח סמך ארוך יותר?  
 א. סטטיסטיκי א.  
 ב. סטטיסטיκי ב.  
 ג. אותו ארוך רוח סמך לשני הסטטיסטיקים.  
 ד. תלוי בתוצאות המוגם של כל סטטיסטיκי.
- 8) נתון ש :  $N(\mu, \sigma^2) \sim X$  ביצעו מוגם בגודל 16 וקיבלו סטיית תקן מוגמית 10.  
 אורך רוח הסמך שהתקבל הוא : 8.765. מהי רמת הביטחון של רוח הסמך?

### תשובות סופיות:

- (1) א.  $\mu < 89.72$       ב.  $\mu > 79.88$   
 ג. הוא היה גדול.      ב. כן.  
 (2) ראה בסרטון.
- (3) א. צריך להניח שהמשתנה מתפלג נורמלית.  
 ב. לא ניתן לדעת.
- (4) א.  $\mu < 107.37$       ב.  $\mu > 107.37$       ג. ראה בסרטון.  
 .96.90 <  $\mu < .96.63$       .  
 .107.10 <  $\mu < .107.37$       .  
 .3.149 <  $\mu < .3.351$       (5)  
 .13.42 <  $\mu < .14.18$       (6)
- (7) ב'.  
 (8) .90%

# bijustisjka le-refoat shinim

פרק 31 - רוח סマー לפרוורציה

תוכן העניינים

145 .....	1. רוח הסマー לפרוורציה .....
148 .....	2. קביעת גודל מוגם .....

## רוח הסמן לפרופורציה:

**רקע:**

המטרה היא לאמוד את  $P$  – פרופורציה באוכלוסייה.

**האומד הנקודתי:**

$$\hat{p} = \frac{y}{n} \quad Y - \text{מספר ההצלחות שבמדגם}.$$

$$\cdot \hat{p} \pm Z_{\frac{1-\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} : p$$

**תנאי לבניית רוח סמן:**

מדגם של לפחות 30 תוצאות (לעתים נתונים תנאי של מספר ההצלחות ומספר כשלונות לפחות 5 או לפחות 10).

**האומד לטיעות התקן:**

$$\cdot L = 2\varepsilon , \hat{P} = \frac{A+B}{2}$$

**דוגמה (פתרון בהקלטה):**

במטרה לאמוד את אחוז המובטלים במשק נגמו 200 אזרחים, מתוכם התקבלו ש-24 היו מובטלים.

א. בנו רוח סמן לאחוז המובטלים באוכלוסייה ברמת סמן של 95%.

ב. מהו האומד לטיעות התקן?

### שאלות:

- 1) נדגמו 200 דירות בעיר חיפה. 48 מהתוכן נמצאו כבעלות ממ"ד.
  - א. בנו רוח סמך ברמת סמך של 95% לאחוז הדירות בחיפה עם ממ"ד.
  - ב. על סמך סעיף א' מה ניתן לומר על שגיאת האמידה המקסימאלית?
  - ג. בהנחה ובחיפה 80 אלף דירות, בנו רוח סמך ברמת סמך של 95% למספר הדירות בחיפה עם ממ"ד בפועל.
  
- 2) במדגם של 300 אנשי היי-טק התקבל ש-180 מהם אקדמיים.
  - א. בנו רוח סמך לפרוורציה אקדמאים ברמת סמך של 95% (בקרב אנשי היי-טק).
  - ב. כיצד רוח הסמך של סעיף א' תהיה משתנה אם היינו מקטינים את רמת הסמך?
  - ג. כיצד רוח הסמך תהיה משתנה אם היינו מגדילים את גודל המדגמים?
  
- 3) במדגם של 400 נוהגים התקבל רוח סמך לפרוורציה הנוהגים החדשניים:
 
$$0.08 < p < 0.18$$
  - א. כמה נוהגים במדגם היו נוהגים חדשים?
  - ב. מהי רמת הסמך של רוח הסמך שנבנה?
  
- 4) במסגרת מערכת הבחירה בארה"ב נשאלו 840 אנשים עבר איזה מועמד יצביעו. 510 אנשים ענו כי יצביעו בעד ברק אובמה. בסקר פורסם שתתכן סטייה של  $\pm 3\%$  מתוצאות האמת. באיזו רמת ביטחון הסקר השתמש?
  
- 5) במדגם של 300 נשים בגילאי 40-45 נמצא ש-140 היו נשואות, 80 היו גרושות, 60 רווקות והיתר אלמנות.
  - א. מצאו רוח סמך ברמה של 90% לאחוז הגרושות באוכלוסייה הנחקרת.
  - ב. מצאו רוח סמך ברמה של 99% לסיcoli שבאוכלוסייה הנחקרת תמצא אישة לא נשואה?
  
- 6) ביצעו מדגם באוכלוסייה. שיעור ההצלחות במדגם היה 10% ורוח הסמך ניבנה ברמת סמך של 95%. אורכו הינו 8.3156%. מהו גודל המדגם שנלקח?

**תשובות סופיות:**

- (1) א.  $18.1\% < p < 29.9\%$ .  
 ב. בביטחון של 95% שגיאת האמידה היא לכל היותר 0.059.  
 ג. לא ניתן לדעת.
- (2) א.  $0.545 \leq p \leq 0.655$ .  
 ב. 0.997.  
 ג. לא ניתן לדעת.
- (3) א. 0.52.  
 ב. 0.925.  
 ג. לא ניתן לדעת.
- (4) א. 60.72% >  $p > 45.91\%$ .  
 ב.  $30.9\% > p > 22.5\%$ .  
 ג. לא ניתן לדעת.
- (5) א. 0.200.  
 ב. לא ניתן לדעת.

## קביעת גודל מוגן:

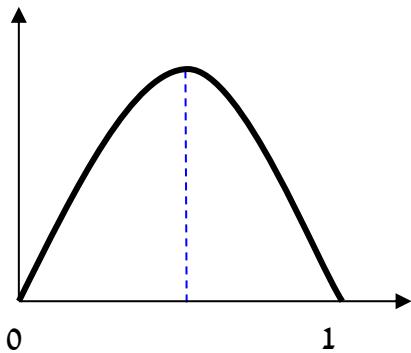
**רקע:**

בפרק זה נדונן איך קובעים גודל מוגן שבאים לאמוד פרופורציה באוכלוסייה מסוימת: החוקר קובע מראש את רמת הסמך הרצוי:  $\alpha - 1$ .

החוקר קובע מראש את הטעות הסטטיסטית המרבית שבה הוא מעוניין:  $\varepsilon$  (או את אורך רוח הסמך).

$L = 2\varepsilon$  - אורך רוח הסמך.

$\varepsilon$  - טעות אמידה מרבית: המרחק המקסימלי (הסטטיה) בין הפרמטר ( $p$ ) לאומד ( $\hat{p}$ ).



ויתעניין לדעת מהו גודל המוגן הרצוי לשם כך.

$$\text{נקלט ש: } \varepsilon = z_{\frac{1-\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

הבעיה שאין אנו יודעים את  $\hat{p}$ .

נתבונן בביטויי:  $(\hat{p} - p)^2$ .

כיוון שאין לנו ידוע מוקדם על  $\hat{p}$  נציב את המקorra השמרני ביותר שמקסם את הביטוי עבור:  $\hat{p} = 0.5$ .

$$\text{נקלט ש: } \varepsilon = z_{\frac{1-\alpha}{2}} \cdot \sqrt{0.5 \cdot 0.5} \Rightarrow n \geq \left( \frac{z_{\frac{1-\alpha}{2}}}{L} \right)^2$$

אך אם תהיה לנו אינפורמציה מוקדמת על פרופורציה נציב את הערך הקרוב ביותר ל-0.5 האפשרי.

**דוגמה (פתרון בהקלטה):**

מעוניינים לאמוד את שיעור האבטלה במשק. האמידה צריכה להתבצע ברמת סמך של 90% ועם שגיאת אמידה שלא תעלה על 4%.

- מהו גודל המוגן המינימאלי שיש לקחת?
- חזור לסעיף א' אם ידוע שהابטלה לא אמורה לעלות על 20%.

**שאלות:**

- 1)** הממשלה אומדת מדי חדש את אחוז הتمיכה בה.  
מהו גודל המדגם אשר יש לנקח אם דורשים שהאומדן לא יסטה מהאחוז האמתי באוכלוסייה ביותר מ-3%, וזאת בביטחון של 95%?
- 2)** משרד התקשורת מעוניין לדעת מה שיעור בתי האב עם אינטרנט.  
א. כמה בתים אב יש לדגום אם מעוניינים שבביטחון של 90% אורך רוח הסמק לא עולה על 8%?  
ב. חזרו על סעיף א' אם ידעו שלפני חמיש שנים לפחות 80% מבתי האב היה אינטרנט וכיום יש להניח שיש יותר אינטרנט.
- 3)** עורך טלוויזיה מעוניין לאמוד את הריאטיבינג של העורך בפריים טיים. המטרה שבביטחון של 95% הסתירה המרבית בין האומדן לריאטיבינג האמתי לא תעלה על 4%.  
א. כמה מכשירי PEOPLE METER יש להתקין לצורך האמידה?  
ב. לפי הערכה מוקדמת הריאטיבינג של העורך לא יכול לעלות על 20%.  
בנה ומכשיר כזה עולה 500 ש"ח ליחידה מה החיסכון הכספי  
מאינפורמציה זאת?
- 4)** ענו על הסעיפים הבאים:  
א. כמה אזוריים יש לדגום כדי לאמוד את אחוז הتمיכה הממשלה עם אורך רוח הסמק שלא עולה על 9% ברמת סמק של 90%?  
ב. בנה ובצע מדגם שאט גודלו חישבתם בסעיף א' והתקבל שאותו  
התמיכה הממשלה במדגם הנו 42%. בנו רוחה סמק לאחוזה התמיכה  
בממשלה ברמת סמק של 95%.  
ג. על סמך סעיף ב', האם תקבלו את הטענה ששיעור האוכלוסייה תומך הממשלה?
- 5)** משרד הבריאות מתכוון לבצע מדגם שמטרתו לבדוק את הסיכוי לחילות בשפעת עם לקיחת חיסון נגד שפעת. הוא מעוניין שבסיכוי של 98% טעות האמידה לא תעלה על 3%.  
א. כמה מחוסנים יש לדגום?  
ב. משרד הבריאות ביצע את המדגם שאט גודלו חישבת בסעיף הקודם  
וקיבל ש-15% מבין אלה שקיבלו חיסון נגד שפעת בכל זאת חלו במשיך  
ההורף בשפעת. בנו ברמת סמק של 98% את הסיכוי לחילות בחורף  
בשפעת עם לקיחת חיסון נגד שפעת.  
ג. בהמשך לסעיף הקודם. מהי טעות האמידה המרבית בביטחון של 98%?  
מדוע הוא קטן מ-3%?

**תשובות סופיות:**

- |                                                             |                   |
|-------------------------------------------------------------|-------------------|
| .1068                                                       | (1)               |
| .271                                                        | ב. (2)            |
| .108,000                                                    | ב. (3)            |
| .367 < p < 0.473                                            | ב. (4)            |
| ג. בביטחון של 0.95 ניתן להגיד ששיעור האוכלוסייה תומך במשלה. |                   |
| א. 0.15±0.02                                                | ג. ראה סרטון. (5) |

## bijostatystyka dla medycyny

rozdział 32 - Rzut oka na rozkłady (mimożeby) w różnych warunkach

### Obszarów tematycznych

- |     |                                                             |
|-----|-------------------------------------------------------------|
| 151 | 1. Cechy charakterystyczne populacji badanej                |
| 153 | 2. Cechy charakterystyczne populacji badanej i porównywanej |

## כשונות האוכלוסייה ידועות:

**רקע:**

המטרה היא לאמוד את פער התוחלות:  $\mu_2 - \mu_1$ , כולם ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי:  $\bar{x}_1 - \bar{x}_2$ .

התנאים לבניית רוח הסמן:

1.  $\sigma^2_1, \sigma^2_2$  ידועות.

2.  $X_1, X_2 \sim N_{n_1, n_2}$  או  $n_1, n_2 > 30$ .

3. שני מוגדים בלתי תלויים.

$$\text{רוח סמן: } (\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}}$$

אם הערך אפס נופל בגבולות רוח הסמן נגיד שבביטחון של  $\alpha-1$ , לא קיים הבדל בין התוחלות.

**דוגמה (פתרון בהקלטה):**

נדגמו 100 תושבים מאזור A והמשכורת הממוצעת הייתה ש 9200 ₪. כמו כן נdagמו 120 תושבים מאזור B וממוצע המשכורות שהתקבל ש 8700 ₪. לצורך פתרון נניח שטטיית התקן של המשכורות באוכלוסיות שני האזוריים היא 1800 ₪. אמדנו ברמת סמן של 90% את הפרש השכר הממוצע בין אזור A לאזור B.

**שאלות:**

- 1)** מעוניינים לבדוק האם קיים הבדל בין ממוצע ציוני הפסיכומטרי של חילילים לממוצע ציוני הפסיכומטרי של תלמידי תיכון. ידוע שצינוי הפסיכומטרי מתפלג עם נורמלilit עס טיטית תקן 100. במדגם של 16 נבחנים חילילים התקבל ממוצע 543. במדגם של 20 תלמידי תיכון התקבל ממוצע 508. בנו רוח סמך לפער תוחלות הצינויים בין חילילים לתלמידי תיכון ברמת סמך של 90%. מה ניתן להסיק מרווח סמך זה?
- 2)** צינוי IQ מתוכנים כך שיתפלגו נורמללית עם סטיטית תקן של 15. במדגם של 20 נבחנים ישראליים התקבל ממוצע צינויים 104. במדגם של 23 נבחנים אמריקאים התקבל ממוצע צינויים 99.
- א. בנו רוח סמך ברמת סמך של 95% לפער בין ישראל לארה"ב בממוצע הצינויים בבחן IQ-ו.
- ב. האם קיים הבדל בין ישראליים אמריקאים מבחינת ממוצע הצינויים?
- 3)** חברה להנדסת בנין מעוניינת להשוות ברמת הקשיות של שני סוגי ברגים. ידוע שרמת הקשיות של ברגים מתפלגת נורמלית עם סטיטית תקן של 4 יחידות. במדגם של 15 ברגים מסווג א' התקבל רמת קשיות ממוצעת של 28 יחידות ובמדגם של 12 ברגים מסווג ב' התקבל רמת קשיות ממוצעת של 25. עברו אילו רמות בטחון יקבע שאיו הבדל בין שני סוגי הברגים מבחינת ממוצע רמת הקשיות שלהם?

**תשובות סופיות:**

- (1) (-20,90).
- (2) א.  $\mu_2 - \mu_1 < 3.99$ .
- ב. לא נוכל בטחון בביטחון של 95% שקיימים הבדל בין ישראל לארה"ב.
- (3) רמות בטחון הגבוחות מ-0.9476.

## כשונות האוכלוסייה לא ידועות ובהנחה שוויון שונויות:

**רקע:**

המטרה היא לאמוד את פער התוחולות:  $\mu_2 - \mu_1$ , כולם ההבדלים של הממוצעים בין שתי האוכלוסיות.

**האומד נקודתי:**  $\bar{x}_1 - \bar{x}_2$ .

**התנאים לבניית רוח סמך:**

$$\cdot \sigma^2_1 = \sigma^2_2 \cdot 1$$

$$\cdot X_1, X_2 \sim N \cdot 2$$

3. מוגדים בلتוי תלויים.

**השונות המשוקלلت:** כיון שאנו מניחים שבין שתי האוכלוסיות השונות שווות אנו אומדים את השונות הזו על ידי שקולן שתי השונות של שני המוגדים על ידי

$$\text{הנוסחה הבאה: } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

**דרגות החופש:**  $d.f = n_1 + n_2 - 2$

$$\text{רוח סמך: } (\bar{x}_1 - \bar{x}_2) \pm t_{1-\frac{\alpha}{2}}^{n_1+n_2-2} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

אם הערך אפס נופל בגבולות רוח סמך נגד שבביטחון של  $\alpha - 1$ , לא קיים הבדל בין התוחולות.

**דוגמה (פתרון בהקלטה):**

מחקר מעוניין לבדוק האם קיים הבדל בין תל אביב לברק שבע מבחן הכנסה הממוצעת של אקדמיים. להלן תוצאות המוגדים שנעשה:

טל אביב	ברק שבע	מספר האקדמאים
10	20	9500
9500	11,000	250
200		סטיית התקן של הכנסות אקדמיים

בנו רוח סמך ברמת ביטחון של 90% להפרש תוחולות הכנסה בשני האזוריים. הניחו שהשכר מתפלג נורמלית עם אותה שוננות בכל אחד מהאזורים.

**שאלות:**

- 1) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו ל מבחן IQ.  
להלן תוצאות המדגמים:

ארה"ב	ישראל	המדינה
15	15	גודל המדגם
1470	1560	סכום הציוניים
147,560	165,390	סכום ריבועי הציוניים

מצאו רוח סמך ברמת סמך של 95% לסתיטה בין ממוצע הציוניים בישראל לממוצע הציוניים בארה"ב. רשמו את כל הנקודות הדרושות לצורך פתרון התרגילים.

- 2) להלן 4 תוצאות על משתנה  $X$  שמתפלג:  $N(\mu_x, \sigma^2)$ , ומשתנה  $Y$  שמתפלג:  $N(\mu_y, \sigma^2)$ .

X	22	20	21	25
Y	18	25	17	12

חשבו רוח סמך ל-  $\mu_x - \mu_y$  ברמת הסמך 90%, בהנחה שני המדגמים בלתי תלויים.

**תשובות סופיות:**

- 1) הנקודות:  
 1. השונות שווה.  
 2. שהציוניים מתפלגים נורמלית.  
 3. המדגמים אינם תלויים זה בזה.  
 $-5.52 < \mu_1 - \mu_2 < 17.52$   
 $-9.6 < \mu_y - \mu_x < 1.6$  (2)

## bijustisjka le-refoat shinim

פרק 33 - רוח סמרק לתחולת (מומוץ) ההפרשים במדגמים מזוגים

תוכן העניינים

1. רוח סמרק לתחולת (מומוץ) ההפרשים במדגמים מזוגים .....

## רוח סמך לתוחלת (ממוצע) ההפרשים במדגים مزוגים:

---

**רקע:**

**מבחן מזוג:** מבחן אחד שבו יש  $n$  צמדים. כל תצפית במדגם תנפק זוג ערכים :  $X$  ו-  $Y$ . ניצור משתנה חדש :  $D = y - x$ . הפרמטר שנרצה לאמוד :  $\mu_D$ . התנאים לבניית רוח הסמך :  $x, y \sim N(1)$ .

2. המבחן מזוג.

נוסחת רוח הסמך :  $\bar{D} \pm t_{\frac{\alpha}{2}}^{n-1} \frac{S_D}{\sqrt{n}}$ . כאשר דרגות החופש :  $df = n - 1$ .

**דוגמה (פתרון בהקלטה):**

מעוניינים לבדוק האם יש הבדל בין מהירות הריצות של שתי תוכנות מחשב. נלקחו 5 קבצים אקראיים והריצו אותם בשתי התוכנות :

5	4	3	2	1	הקובץ
38	46	49	48	25	זמן בתוכנה הראשונה
48	40	42	46	27	זמן בתוכנה השנייה

הניחו כי זמי הריצות מתפלגים נורמלית. נמצא רוח סמך של 95% להפרש תוחלת הזמן בין שתי התוכנות.

**שאלות:**

- 1) נדגמו 5 סטודנטים ששסיימו את הקורס סטטיסטיKA ב'. להלן הציונים בסמסטר A' ו-B':

82	75	90	68	74	<b>סמסטר A'</b>
100	76	87	84	80	<b>סמסטר B'</b>

- נניח שהציונים מתפלגים נורמללית.
- A. בנו רוח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סמסטר A' לבין סמסטר B'.
- B. האם על סמך רוח הסמך קיים הבדל בין הסמסטרים מבחינת תוחלת הציונים?
- C. מה צריך לשנות בתנאים כדי שהمدגמים יהיו בלתי תלויים?
- 2) במטרה לבדוק האם קיים הבדל בין קווי זהב לבזק מבחינת ממוצע המחרירים לשיחות בין"יל. נדגמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקט שיחת. להלן התוצאות:

חברה/ מדינה	ארה"ב	אירופה - X	יפן	סין	מצרים	פולין	הולנד	קנדה
בזק - X	1.5	2.1	2.2	3	2.1	2.1	2.2	2.1
קווי זהב - Y	1.4	2	1.9	3.1	1.9	1.9	1.9	1.9

בහנה והחרירים מתפלגים נורמלית עבור כל חברה, בנו רוח סמך ברמת סמך של 90% לתוחלת הפרש המחרירים של שתי החברות.

**תשובות סופיות:**

- 1) א.  $\mu_0 < 38$ .      ב. בביטחון של 95% לא קיים הבדל.  
ג. ראה הסבר בסרטון.
- 2)  $-0.013 < \mu < 0.185$ .

# bijustisjka le-refoat shinim

## פרק 34 - שאלות מסכמת על רוחי סマー

תוכן העניינים

1. שאלות מסכמת על רוחי סマー .....

157 .....

## שאלות מסכמת על רוחי סמך:

### שאלות:

2) 200 אנשים נשאלו כמה פעמים ביום הם שותים כוס קפה.  
להלן התפלגות התשובות:

מספר אנשים	מספר פעמים
86	0
34	1
28	2
22	3
20	4
10	5

- א. תננו רוח סמך למספר כוסות הקפה שאנשים נהגים לשותות ביום.  $\alpha = 0.05$ .
- ב. אדם השותה לפחות 4 כוסות קפה ביום נקרא "מכור לקפה". בנו רוח סמך לאחוז "המכורים לקפה".  $\alpha = 0.1$ .
- 3) חוקר בנה רוח סמך לאחוז האנשים שהתקरרו לפחות פעם אחת בשנה. רוח הסמך שהתקבל הוא:  $p < 91 < 81$ . רוח הסמך הניל הتبבס על מדגם של 500 איש.
- א. כמה אנשים במדגם טענו שככל לא התקררו השנה?
- ב. באיזו רמת סמך נבנה רוח הסמך?
- ג. בנו רוח סמך לאחוז האנשים שהתקררו לפחות פעם אחת השנה ברמת סמך של 96% על סמך תוצאות המדגם.
- 4) ציוני IQ באלה"ב מתפלגים נורמלית עם תוחלת 100. במדגם של 20 ישראלים שנבחנו ב מבחן ה-IQ התקבלו התוצאות הבאות:
- $$\sum_{i=1}^{20} x_i = 2040, \sum_{i=1}^{20} x_i^2 = 210740$$
- א. אמדנו ברמת ביטחון של 90% את ממוצע ציוני בחינת ה-IQ בישראל – מהי ההנחה הדרישה לפתרון?
- ב. על סמך רוח הסמך של סעיף א' האם תקבלו את הטענה שבישראל ממוצע הציונים שונה מארה"ב?
- ג. מה היה קורה לרוח הסמך אם היינו מגדילים את רמת הסמך שלו?

- 7) בנק מתלבט האם לפתח סניף באזור A או באזור B. לצורך פתרון נניח שסטיות התקן של המשכורת באזור A היא 1200 ובאזור B 1500. הבנק דוגם 50 אנשים מאזור A, המשכורת המומוצעת שהתקבלה במדגם היא 6,800 ₪. כמו כן, נדגמו 40 אנשים מאזור B, המשכורת המומוצעת שהתקבלה במדגם היא 6,600 ₪.
- א. בנו רוח סמך ברמת סמך של 95% להפרש המומוצעים של המשכורות בשני האזורים. האם על רוח הסמך ניתן להמליץ לבנק היכן לפתח את הסניף. אם כן, היכן?
- ב. בנו רוח סמך לתוחלת המשכורת באזור A ברמת סמך של 95%.

### תשובות סופיות:

- |                                        |    |                                      |        |
|----------------------------------------|----|--------------------------------------|--------|
| . $10.85\% \leq p \leq 19.15\%$        | ב. | . $1.21 \leq \mu \leq 1.65$          | א. (2) |
| . $83\% < p < 89\%$                    | ב. | .0.9988                              | א. (3) |
| ג. גידל.                               | ב. | . $97.4 \leq \mu \leq 106.6$         | א. (4) |
| . $6467 \leq \mu_A \leq 7133$ , לא. ב. | ב. | . $-372 \leq \mu_A - \mu_B \leq 772$ | א. (7) |

# bijustisjka le-refoat shinim

פרק 35 - מבוא לבדיקה השערות על פרמטרים

תוכן העניינים

159 .....	1. הקדמה.....
163 .....	2. סוגים טעויות.....

## הקדמה:

**רקע:**

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטי. בבדיקה השערות על פרמטרים עוסcid לפיה שלבים הבאים:

**שלב א:** נזהה את הפרמטר הנחקר.

**שלב ב:** נרשות את השערות המחקר. השערת האפס המסומנת ב- $H_0$ .

בדרך כלל השערת האפס מסמלת את אשר היה מקובל עד עכשו, את השגרה הנורמה.

השערה אלטרנטיבית (השערת המחקר) המסומנת ב- $H_1$ .

ההשערה האלטרנטיבית מסמלת את החדשנות בעצם ההשערה האלטרנטיבית בדברת על הסיבה שהמחקר נעשה היא שאלת המחקר.

**שלב ג:** נבדוק האם התנאים לביצוע התהליך מתקיימים ונניח הנחות במידת הצורך.

**שלב ד:** נרשות את כלל ההכרעה. בתהליך של בדיקת השערות יוצרים כלל שנקרה כלל הכרעה. הכלל יוצר אзорים שנקרים:

1. **אזור דחיה:**

דחיה של השערת האפס כלומר קבלה של האלטרנטיבית.

2. **אזור קבלה:**

קיבלה של השערת האפס ודחיה של האלטרנטיבית. כלל ההכרעה מתבסס על איזשהו סטטיסטי. אזור הדחיה מוכתב על ידי סיכון שלוקח החוקר מראש שנקרה רמת מובהקות ומסומן ב- $\alpha$ .

**שלב ה:** בתהליך יש ללקת לתוצאות המדגם וליחס את הסטטיסטי המתאים ולבדוק האם התוצאות נופלות באזור הדחיה או הקבלה.

**שלב ו:** להסיק מסקנה בהתאם לתוצאות המדגם.

**דוגמה (פתרון בהקלטה):**

משרד הבריאות פרסם משקל ממוצע של תינוקות ביום לידהם בישראל 3300 גרם. משרד הבריאות רוצה לחזור את הטענה שנשים מעשנות בזמן ההריון יולדות תינוקות במשקל נמוך מהממוצע. במחקר השתתפו 20 נשים מעשנות בהריון.

להלן תוצאות המדגם שבודק את המשקל של התינוקות בעת הלידה :

$$\bar{X} = 3120, S = 280, n = 20.$$

- א. מהי אוכלוסיית המחקר?
- ב. מה המשתנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

**שאלות:****בשאלות הבאות, ענו על הטעיפים הבאים:**

- א. מהי אוכלוסיית המחקר?
- ב. מה המשנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

- 1)** ממוצע הציונים בבחינת הבגרות באנגלית הנו 72 עם סטיטית תקן 15 נקודות. מורה טוען שפיתחה שיטת לימוד חדשה שתעלה את ממוצע הציונים. משרד החינוך החליט לתת למורה 36 תלמידים אקראים. ממוצע הציונים של אותם תלמידים לאחר שלמדו בשיטתו היה 75.5.
- 2)** לפי הצהרת היিירן של חברת משקאות מסויימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיטית תקן 20 סמ"ק. אגודה הרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצחרת. במדוג שעשתה אגודה הרכנים התקבל נפח ממוצע של 492 סמ"ק במדוג בגודל 25.
- 3)** במשך שנים אחדו המועמדים שהתקבלו לפיקולטה למשפטים היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. מחקר מעוניין לבדוק האם השנה מקשים על הקבלה לפיקולטה למשפטים.
- 4)** בחודש ינואר השנה פורסם שאחדו האבטלה במשק הוא 8% במדוג עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים. רוצחים לבדוק ברמת מובהקות של 5% האם אחדו האבטלה הוא כמו בתחילת השנה.

### תשובות סופיות:

- ב. ציון.
- 1) א. נבחנים בברירות באנגלית.  
 $H_0: \mu = 72$       ג. ממוצע הציונים בשיטת לימוד חדשה.  
 $H_1: \mu > 72$
- ב. נפח משקה בבקבוק של חברת מסויימת.
- 2) א. משקאות בבקבוק של חברת מסויימת.  
 $H_0: \mu = 500$       ג. ממוצע נפח המשקה בבקבוק.  
 $H_1: \mu < 500$
- ב. משתנה דיכוטומי (התקבל, לא התקבל).
- 3) א. מועמדים לפיקולטה למשפטים.  
 $H_0: p = 0.25$       ג. אחוז הקבלה.  
 $H_1: p < 0.25$
- ב. משתנה דיכוטומי (מובטל, עובד).
- 4) א. אזרחים בוגרים במשק.  
 $H_0: p = 0.08$       ג. אחוז האבטלה ביום.  
 $H_1: p \neq 0.08$

## סוגי טעויות:

**רکע:**

בתחילת בדיקת השערות יוצרים כלל שנקרא כלל הכרעה. הכלל יוצר אзорים שנקראים:

1. אזור דחיה – דחיה של השערת האפס כלומר קבלה של האלטרנטיבית.
2. אזור קבלה – קבלה של השערת האפס ודחיה של האלטרנטיבית.

כל הכרעה מתבסס על איזשהו סטטיסטי. בתחילת יש ל选取 תוצאות המדגם ולבזוק האם התוצאות נופלות באזור הדחיה או הקבלה וכן להגעה למסקנה – המסקנה היא עירובן מוגבל כיוון שהיא תלויה בכל הכרעה ובתוצאות המדגם. אם נשנה את כלל הכרעה אז אנחנו יכולים לקבל מסקנה אחרת. אם נבצע מדגם חדש אז אנחנו עלולים לקבל תוצאה אחרת. לכן יתכונו טעויות במסקנות שלנו:

		הכרעה	
מציאות	$H_0$	$H_1$	
	$H_0$	אין טעות 1	טעות מסוג 1
	$H_1$	טעות מסוג 2	אין טעות

### הגדרת הטעויות:

טעות מסוג ראשון: להכריע לדוחות את  $H_0$  למראות שבמציאות  $H_0$  נכונה.  
טעות מסוג שני: להכריע לקבל את  $H_0$  למראות שבמציאות  $H_1$  נכונה.

### דוגמה (פתרון בהקלטה):

אדם חשוד בביוץ עבירה ונتابע בבית המשפט.  
אילו סוגי טעויות אפשריות בהכרעת הדין?

**שאלות:**

- 1)** לפי הצהרת היכרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הcrcנים מTELONNATE על הפחתת נפח המשקה בבקבוק מהכמות המוצחרת. בדוגמא שעשתה אגודת הcrcנים התקבל נפח ממוצע של 492 סמ"ק במדוגים בגודל 25. בסופו של דבר הוחלט להזכיר לטובת חברת המשקאות.
- רשמו את השערות המחקר.
  - מה מסקנת המחקר?
  - אייזו סוג טעות יתכן וביצעו במחקר?
- 2)** במחקר על פרמטר מסוים הוחלט בסופו של דבר לדוחות את השערת האפס.
- אם ניתן לדעת אם בוצע טעות במחקר?
  - מה סוג הטעות האפשרית?
- 3)** לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטיית תקן 0.4. ישנה טענה שכיוום ממוצע מספר הילדים במשפחה קטן יותר. לצורך כך הוחלט לדוגם 121 משפחות. בדוגמא התקבל ממוצע 2.17 ילדים למשפחה. על סמך תוצאות המדוגם נקבע שלא ניתן לקבוע שבאופנו מובהק תוחלת מספר הילדים למשפחה קטנה כיום.
- מהי אוכלוסיות המחקר?
  - מה המשנה הנחקרה?
  - מה הפרמטר הנחקר?
  - מה השערות המחקר?
  - מה מסקנת המחקר?
  - מי סוג הטעות האפשרית במחקר?

**תשובות סופיות:**

- 1)** א.  $\mu = 500$ .  
ב.  $\mu < 500$ .
- 2)** א. לא ניתן לדעת.  
ב. טעות מסווג ראשונה.
- 3)** א. משפחות כיום.  
ב. מס' הילדים.
- ג. תוחלת מספר הילדים למשפחה כיום.  
ה. לא לדוחות את  $H_0$ . ו. טעות מסווג שנייה.
- $H_0 : \mu = 2.3$   
 $H_1 : \mu < 2.3$

# bijustisjka le-refoat shinim

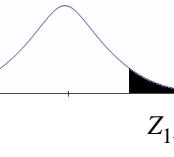
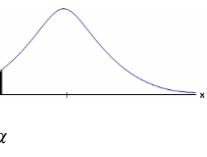
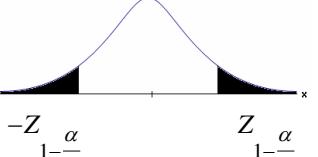
## פרק 36 - בדיקת השערות על תוחלת (ממוצע)

### תוכן העניינים

1. בדיקת השערות על תוחלת (ממוצע) כשבונות האוכלוסייה ידועה	165
2. סיכון לטעויות ועוצמה (שבונות האוכלוסייה ידועה)	169
3. קביעת גודל מוגן (שבונות האוכלוסייה ידועה)	175
4. מובהקותות תוצאה - אלף מינימלית (שבונות האוכלוסייה ידועה)	178
5. בדיקת השערות על תוחלת (ממוצע) כשבונות האוכלוסייה לא ידועה	183
6. מובהקותות תוצאה - אלף מינימלית (שבונות האוכלוסייה לא ידועה)	187
7. הקשר בין רוח סמך לבדיקה השערות על תוחלת (ממוצע)	190
8. ניתוח פלטים	192

## בדיקות השערות על תוחלת (ממוצע) כשבונות האוכלוסייה ידועה:

**רקע:**

$H_0 : \mu \leq \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu = \mu_0$	השערת האפס: השערת אלטרנטיבית:
$H_1 : \mu > \mu_0$	$H_1 : \mu < \mu_0$	1. $\sigma$ ידועה או מוגן מספיק גדול $X \sim N$ .2	
$Z_{\bar{x}} > Z_{1-\alpha}$  -דוחים את $H_0$	$Z_{\bar{x}} < -Z_{1-\alpha}$  -דוחים את $H_0$	$Z_{\bar{x}} < -Z_{\frac{1-\alpha}{2}}$ או $Z_{\bar{x}} > Z_{\frac{1-\alpha}{2}}$  -דוחים את $H_0$	<b>כל הכרעה: אזור הדחיה של <math>H_0</math>:</b>

**סטטיסטי המבחן:**  $Z_{\bar{x}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

**חלופה אחרת לכל הכרעה:**

$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$	<b>נדחה <math>H_0</math> אם מתקיים:</b>
----------------------------------------------------------------	----------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------

**דוגמה:**

יבול העגבנייהות מתפלג נורמלית עם תוחלת של 10 טון לדונם וסטיית תקן של 2.5 טון לדונם בעונה. משערים ששיטת זיוב חדשת تعالה את תוחלת היבול לעונה מבלי לשנות את סטיית התקן. נדגמו 4 חלוקות שזובלו בשיטה החדשת. היבול הממוצע שהתקבל היה 12.5 טון לדונם. בדקו את ההשערה ברמת מובהקות של 1%.

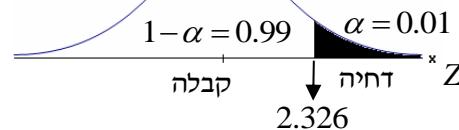
**פתרונות:**אוכלוסייה: עגבנייהות.המשתנה:  $X$  = יבול העגבנייהות בטון לעונה.הפרמטר:  $\mu$  = תוחלת היבול בשיטה החדשת.

$$\begin{aligned} H_0 : \mu &= 10 \\ H_1 : \mu &> 10 \end{aligned}$$

תנאים:

1.  $X \sim N$ .

2.  $\sigma = 2.5$ .

כל הכלעה:נדחה את  $H_0$  אם  $Z_{\bar{x}} > 2.326$ תוצאות:  $n = 4$ ,  $\bar{x} = 12.5$ 

$$\text{סטטיסטי המבחן} : Z_{\bar{x}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$\text{נzieb} : Z_{\bar{x}} = \frac{12.5 - 10}{\frac{2.5}{\sqrt{4}}} = 2 < 2.326$$

מסקנה:לא נדחה  $H_0$  (נקבל  $H_0$ ).

ברמת מובהקות של 1% לא נוכל לקבל את הטענה ששיטת הזיבול החדשת מעלה את תוחלת היבול של העגבנייהות.

**שאלות:**

- 1)** מוצע הציונים בבחינות הבגרות באנגלית הנו 72 עם סטיית תקן 15 נקודות. מורה טוען שפיתח שיטת לימוד חדשה שתעלה את ממוצע הציונים. משרד החינוך החליט לתת למורה 36 תלמידים אקראים. מוצע הציונים של אותם תלמידים לאחר לימודו בשיטתו היה 75.5. בהנחה שגם בשיטתו סטיית התקן תהיה 15 מה מסקנתכם ברמת מובהקות של 5%?
- 2)** לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודות היצרנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצחרת. במדוגם שעשתה אגודות היצרנים התקבל נפח ממוצע של 492 סמ"ק במדוגם בגודל 25.
- מה מסקנתכם ברמת מובהקות של 2.5%?
  - האם ניתן לדעת מה תהיה המסקנה עבור רמת מובהקות גבוהה מ-5%?
- 3)** מהנדס האיכות מעוניין לבדוק אם מכונה מכילה (מאופסת). המכונה כוננה לחתווך מוטות באורך 50 ס"מ. לפי נתוני היצרן סטיית התקן בחיתוך המוטות היא 0.5 ס"מ. במדוגם של 50 מוטות התקבל ממוצע אורך המוט 50.93 ס"מ. מה מסקנתכם ברמת מובהקות של 5%?
- 4)** המשקל המוצע של הספורטאים בתחום ספורט מסוים הוא 90 ק"ג, עם סטיית תקן 8 ק"ג. לפי דעת מומחים בתחום יש צורך בהורדת המשקל ובשימוש בדיאטה מסוימת לצריכה להביא להורדת המשקל. לשם בדיקתיעילות הדיאטה נלקח מדגם מקורי של 50 ספורטאים ובתום שנה של שימוש בדיאטה התברר שהמשקל המוצע במדוגם זה היה 84 ק"ג. יש לבדוק בר"מ של 10%, האם הדיאטה גורמת להורדת המשקל.
- 5)** לפי מפרט נתון, על עובי בורג להיות 4 מ"מ עם סטיית תקן של 0.2 מ"מ. במדוגם של 25 ברגים העובי המוצע היה 4.07 מ"מ. קבעו ברמת מובהקות 0.05, האם עובי הברגים מתאים למפרט. הניחו כי עובי של בורג מתפלג נורמלית וסטיית התקן של עובי בורג היא אכן 0.2 מ"מ.
- 6)** במחקר נמצא שתוצאה היא מובהקת ברמת מובהקות של 5% מה תמיד נכון? בחרו בתשובה הנכונה.
- הגדלת רמת המובהקות לא תנסה את מסקנת המחקר.
  - הגדלת רמת המובהקות תנסה את מסקנת המחקר.
  - הקטנת רמת המובהקות לא תנסה את מסקנת המחקר.
  - הקטנת רמת המובהקות תנסה את מסקנת המחקר.

- 7) חוקר ערך מבחן דו צדי ברמת מובהקות של  $\alpha$  והחליט לדחות את השערת האפס. אם החוקר היה עורך מבחן דו צדי ברמת מובהקות של  $\frac{\alpha}{2}$  אז בהכרח:
- השערת האפס הייתה נדחתה.
  - השערת האפס הייתה לא נדחתה.
  - לא ניתן לדעת מה תהיה מסקנתו במקרה זה.
- 8) שני סטטיסטיקים בדקו השערות:  $H_1: \mu > \mu_0$ ,  $H_0: \mu = \mu_0$  נגדן עברו שנות ידועה ובאותה רמת מובהקות. שני החוקרים קיבלו אותו ממוצע במדגם אך לחוקר א' היה מדגם בגודל 100 ולחוקר ב' מדגם בגודל 200.
- אם בחוקר א' החלטת לדחות את  $H_0$ , מה יהיה החלטת בחוקר ב'? נמקו.
  - אם בחוקר א' יחליט לא לדחות את  $H_0$ , מה יהיה החלטת בחוקר ב'? נמקו.

### תשובות סופיות:

- נקבל  $H_0$ , בר"מ של 5% לא קיבל את הטענה של המורה ששיטת הלימוד שלו מעלה את ממוצע הציונים.
- א. נדחה  $H_0$ , בר"מ של 2.5% קיבל את תלונת אגודות הרכנים בדבר הפחחת נפח המשקה בבקבוק.  
ב. הגדלנו את רמת המובהקות בכך אנחנו נשארים בדוחיה של  $H_0$  והמסקנה לא תשתנה.
- נדחה  $H_0$ , בר"מ של 5% נקבע שהמכונה לא מאופסת.
- נדחה  $H_0$ , בר"מ של 0.1 נקבל את הטענה שהדיאטה עיליה ומפחיתה את המשקל הממוצע.
- נקבל  $H_0$ , בר"מ של 0.05 נזכיר שתוחלת עובי הבורג מתיים למפרט.
- א'.  
ב'. ג'.
- א. לדחות.  
ב. לא ניתן לדעת.

## סיכום לטעויות ועוצמה (שינוי האוכלוסייה ידועה):

רקע:

		הכרעה	
		$H_0$	$H_1$
מציאות	$H_0$	אין טעות 1	טעות מסוג 1
	$H_1$	טעות מסוג 2	אין טעות

הגדרת הסתברויות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות) :  
 $(\text{לדוחות } H_0 = P_{H_0} (H_0 \text{ נכונה}) | \text{ לדוחות את } H_0)$

הסיכוי לבצע טעות מסוג 2 :  
 $(\text{לקבל } H_0 = P_{H_1} (H_1 \text{ נכונה}) | \text{ לקבל את } H_1)$

רמת בתרון :  
 $(\text{לקבל } H_0 = P_{H_0} (H_0 \text{ נכונה}) | \text{ לקבל את } H_0)$

עוצמה :  
 $(\text{לדוחות } H_1 = P_{H_1} (H_1 \text{ נכונה}) | \text{ לדוחות את } H_1)$

**התהlixir לחישוב סיכוי לטעות מסוג שני:**

$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	<b>השערת האפס: השערת אלטרנטיבתית:</b>
$H_1 : \mu > \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu \neq \mu_0$ <b>תנאים:</b> 1. $\sigma$ ידועה 2. או מדגם מספיק גדול $X \sim N$ .	
$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$	<b>כל הכרעה: אזור הדחיה של <math>H_0</math>:</b>
$P_{H_1} \left( \bar{X} < \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right)$	$P_{H_1} \left( \bar{X} > \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right)$	$P_{H_1} \left( \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$	<b>חישוב <math>\beta</math>:</b>

**התפלגות ממוצע המדגמים:**  $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$$\text{התקנון: } Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

**דוגמה:**

בתחילת השנה חשבו הטלפון הסלולארי הממוצע לאדם היה 200 נק' עם סטיית תקן של 80 נק' לחודש. בעקבות כניסה של חברות טלפון סלולארית חדשות מעונייניות לבדוק האם ביום ממוצע חשבו הטלפון הסלולארי פחות. לצורך בדיקה דגמו באקראי 36 אנשים וחשבו הטלפון הסלולاري שלהם היה 150 נק' בממוצע לחודש.

- רשמו את השערות המחקר ובנו כלל הכרעה במנוחי חישוב ממוצע מדגמי ברמת מובהקות של 5%.
- מה מסקנתכם? איזה סוג טעות אפשרית במסקנה?
- נניח שבמציאות ביום החשבון הממוצע הוא 160 נק'. מה הסיכוי לבצע טעות מסוג שני?
- אם נקבע את רמת המובהקות מסעיף א', כיצד הדבר ישפיע על התשובה מסעיף ג'?

**פתרונות:**א. אוכלוסייה: משלמי חשבון טלפון סלולאר Cioms.המשתנה :  $X = \text{חשבון הטלפון החדש שקלים}$ .הפרמטר :  $\mu$ .

$$\begin{aligned} H_0: \mu &= 200 \\ H_1: \mu &< 200 \end{aligned}$$

תנאים :

$$\cdot \mu = 200 \cdot 1$$

$$\cdot n = 36 \cdot 2$$

$$\cdot \bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \quad K = \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\alpha = 0.05$$

$$Z_{1-\alpha} = Z_{0.95} = 1.645$$

$$\text{נציב: שקלים } K = 200 - 1.645 \cdot \frac{80}{\sqrt{36}} = 178.07$$

ככל ההכרעה: דחה את  $H_0$  אם שקלים  $\bar{X} < 178.07$ 

ב. ברמת מובהקות של 5% נזכיר שאכן ממוצע חשבון הטלפון הסלולרי פחת מתחילת השנה.

$$\begin{aligned} H_0: \mu_0 &= 200 \\ H_1: \mu &< 200 \end{aligned}$$

ככל ההכרעה: נדחה את  $H_0$  אם  $\bar{X} < 178.07$ 

$$H_1: \bar{X} \sim N\left(160, \frac{80^2}{36}\right)$$

$$Z = \frac{178.07 - 160}{\frac{80}{\sqrt{36}}} = 1.36$$

$$\beta = P_{H_1} \left( \bar{X} > 178.07 \mid H_0 \right) = P_{H_1} \left( \bar{X} > 178.07 \right) = 1 - \phi(1.36) = 1 - 0.9131 = 0.0869$$

ד. הקטנת  $\alpha$  מגדילה את  $\beta$ .

**שאלות:**

**1)** נתון ש:  $X \sim N(\mu, \sigma^2 = 1)$ .

להלן השערות של חוקר לגבי הפרמטר  $\mu$ :  $H_0: \mu = 5$ ,  $H_1: \mu = 7$ . מעוניינים ליצור כל הכרעה המתבסס על הסמך תצפית בזדמנות כז שרמת המובהקות תהיה 5%.

א. עבור אילו ערכים של  $X$  שידגמו נדחתת השערת  $H_0$ ?

ב. מה הסיכוי לבצע טעות מסוג שני?

ג. אם במדגם התקבל ש-  $X = 6.9$  מה תהיה המסקנה ומה הטעות האפשרית?

**2)** לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטטיסטיקת תקן 0.4. מעוניינים לבדוק אם כיוון ממוצע מספר הילדים למשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה.

א. רשמו כלל הכרעה במונחי ממוצע מדגם קרייטי ברמת מובהקות של 5%.

ב. בהמשך לסעיף א' מה תהיה המסקנה ומהי הטעות האפשרית במסקנה?

ג. אם באמצעות ממוצע מספר הילדים במשפחה פחות לכדי 2.1 מהי העצמה של הכלל מסעיף א'?

**3)**להלן נתונים על תהליכי בדיקת השערות על תוחלת:

$n = 30$ ,  $\sigma = 30$ ,  $H_1: \mu \neq 200$ ,  $H_0: \mu = 225$ .

א. רשמו כלל הכרעה במונחי ממוצע מדגם קרייטי וברמת מובהקות של 10%.

ב. בהמשך לסעיף א', מהי העצמה אם התוחלת שווה ל-195?

ג. הסבירו, ללא חישוב, איך העצמה תשנה אם רמת המובהקות תהיה 5%?

**4)** מפעל לייצור צינורות מייצרת צינור שקווטרו מתפלג נורמלית עם תוחלת של 50 מ"מ וסטטיסטיקת תקן של 6 מ"מ. במחלקת ביקורת האיכות דוגמים בכל יום 81 צינורות ומודדים את קוטרם, בצד בדוק, בעזרת מבחן סטטיסטי, האם מכונת הייצור מכוקית כנדרש או שקווטר הצינורות קטן מהדרוש.

א. רשמו את ההשערות ואת כלל ההכרעה ברמת מובהקות של 5%.

ב. אם ביום כלשהו מכונת הייצור התקללה והיא מייצרת את הצינורות שתקלה לא תגלה בבדיקה האיכות? כיצד נקבעת הסתברות זו?

ג. הסבירו ללא חישוב כיצד התשובה לשעיף ב' תשנה אם רמת המובהקות תנגדל.

ד. הסבירו ללא חישוב כיצד התשובה לשעיף ב' תשנה אם התוחלת האמיתית היא 47 ולא 48 מ"מ.

- 5) להלן השערות של מחקר:  $H_0: \mu = 50$ ,  $H_1: \mu = 58$ .  
 מעוניינים לדגום 100 תכפיות. ידוע שטטיות התקן של ההתפלגות הינה 20.  
 א. בנו כלל הכרעה שהסיכוי לטעות מסוג שני בו הוא 10%.  
 מהי רמת המובהקות?  
 ב. כיצד הייתה משתנה רמת המובהקות אם (כל סעיף בפני עצמו)?  
 i. טטיות התקן הייתה יותר גדולה.  
 ii. הסיכוי לטעות מסוג שני גדול יותר.

**השאלות שלහן הן שאלות רב-ברירה, בחרו בתשובה הנכונה ביותר:**

- 6) אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו אז:  
 א. הסיכוי לטעות מסוג ראשון גדול.  
 ב. העוצמה של המבחן גדולה.  
 ג. הסיכוי לטעות מסוג שני גדול.  
 ד. תשובות א' ו-ב' נכונות.
- 7) חוקר ביצע מחקר ובו עשה טעות מסוג שני בכך:  
 א. השערת האפס נכונה.  
 ב. השערת האפס נדחתה.  
 ג. השערת האפס לא נדחתה.  
 ד. אף אחת מהתשובות לא נכונה בהכרח.

- 8) מה המצב הרצוי לחוקר המבצע בבדיקה השערה:

$\alpha$	$1 - \beta$
א. גדולה	קטנה
ב. גדולה	קטנה
ג. קטנה	גדולה
ד. קטנה	קטנה

- 9) נערך שינוי בכלל ההחלטה של בדיקת השערה מסוימת ובעקבותיו איזור דחיתת  $H_0$  קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה לכך:  
 א. הוא  $\alpha$ , והוא  $\beta - 1$ , קטן.  
 ב.  $\alpha$  יישאר ללא שינוי ואילו  $\beta - 1$  גדל.  
 ג.  $\alpha$  גדל ואילו  $\beta - 1$  קטן.  
 ד. הוא  $\alpha$  והוא  $\beta - 1$  גדלו.

**10)** ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניח של לחץ הדם בקרוב עיתונאים גבוה יותר מה ממוצע באוכלוסייה. הואלקח מדגם של 60 עיתונאים וקיים ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק של לחץ הדם בקרוב העיתונאים אינו גבוה יותר. מה הטעות האפשרית שהרופא עושה?

- א. טעות מסוג ראשון.
- ב. טעות מסוג שני.
- ג. טעות מסוג שלישי.
- ד. אין טעות במסקנותו.

### תשובות סופיות:

- (1) א. מעל 0.3594. ב. 6.645.
- (2) ג. דחינו את  $H_0$ , ת騰ן טעות מסוג ראשון.
- (3) א. נדחה  $H_0$  אם  $\bar{X} < 2.24$ . ב.  $\bar{X} > 203.29$  או  $\bar{X} < 196.71$ . ג. תקתו.
- (4) א. נדחה  $H_0$  אם  $\bar{X} < 48.9$ . ב. 0.0885. ג. תקתו. ד. תקתו.
- (5) א. 0.0033. ב. נ. רמת המובהקות הייתה קטנה. ב. נ. רמת המובהקות הייתה גבוהה.
- (6) ד. נ.
- (7) ג. נ.
- (8) ג. נ.
- (9) א. נ.
- (10) ב. נ.

## **קביעת גודל מוגן (שינוי האוכלוסייה ידועה):**

**רקע:**

השערות המחקר הן:  $H_0: \mu = \mu_0$ ,  $H_1: \mu = \mu_1$ .  
 סטיטית התקן של האוכלוסייה ידועה  $\sigma$  ומשמעותיים לבצע מחקר שרמת המובהקות  
 לא תעלה על  $\alpha$  והסיכוי לטעות מסווג שני לא עלה על  $\beta$ .

$$\text{הנוסחה הבאה נותנת את גודל המוגן הרצוי: } n \geq \left( \frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

**דוגמא:**

משרד החינוך מפעיל בגין חובה שיטת חינוך שפותחה בשנת 1995. לפי שיטת חינוך זו  
 תוחלת הציון בבחן אוצר מילים לגיל הרך הוא 70. אנשי חינוך החליטו לבדוק  
 שיטת חינוך שפותחה בהולנד הנוגנת שם תוחלת ציון אוצר מילים של 80.  
 נניח שציוני מבחן זה מתפלגים נורמלית עם  $\sigma = 17$ .  
 כדי לבדוק האם גם בישראל הפעלת שיטת החינוך ההולנדית תעבוד בגנים, רוצחים  
 לבנות מחקר ברמת מובהקות של 5%. כמו כן, מעוניינים שאמ בפעולת השיטה  
 ההולנדית תוחלת הציונים תעלה לכדי 80, המחקר יגלה זאת בסיכוי של 90%.  
 כמה ילדי גן חובה דרושים למחקר?

**פתרון:**

האוכלוסייה: ילדי גן חובה.

המשתנה:  $X$  = ציון בבחן אוצר מילים.

הפרמטר:  $\mu$ .

$$\begin{aligned} \text{השערות: } H_0: \mu &= 70 \\ H_1: \mu &= 80 \end{aligned}$$

$$X \sim N(\mu, \sigma^2 = 17^2)$$

אם בפעולת השיטה ההולנדית התוחלת תעלה ל80, נגלה זאת בסיכוי 90%.

$$n \geq \left( \frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

$$\alpha = 0.05$$

$$1 - \beta = 0.9$$

$$\mu_0 = 70$$

$$\mu_1 = 80$$

$$\sigma = 17$$

$$\begin{aligned} Z_{1-\alpha} &= Z_{0.95} = 1.645 \\ Z_{1-\beta} &= Z_{0.9} = 1.282 \end{aligned}$$

$$\begin{aligned} n \geq & \left( \frac{(1.645 + 1.282) \times 17}{70 - 80} \right)^2 = 24.76 \\ \text{נכיב: } & \cdot n_{\min} = 25 \end{aligned}$$

**שאלות:**

- 1)** ב厰ן אינטיגנצה הציוניים מתפלגים נורמללית עם סטיית תקן 8 וממוצע 100. פסיקולוג מעוניין לבדוק את הטענה שבאוכליות במצב סוציאו אקונומי נמוך תוחלת הציוניים היא 95. אם מעוניינים לגלו את הטענה בהסתברות של לפחות 99% כشرط המובהקות היא 5% מהו גודל המדגם הדרוש?
- 2)** משרד התקשורת טוענים שאדם מדבר בממוצע 180 דקות בחודש בטלפון הסלולרי. חברות הטלפון הסלולרי טוענות שאינפורמציה זו אינה נכונה ואדם מדבר בממוצע פחות : כ-160 דקות. לצורך פתרון נניח שסטיית התקן של זמן השיחה החדש ידוע ושווה ל-60 דקות. כמה אנשים יש לדגום כך שאם טענת משרד התקשורת נכונה נדחה אותה בסיכוי של 5% (איך קוראים להסתברות זאת?) כמו כן אם טענת חברות הטלולריות נכונה המחקר יגלה זאת בסיכוי של 90% (איך קוראים להסתברות זאת?).
- 3)** השערות המחקר הן :  $\mu_1 = \mu$  ,  $H_0: \mu = \mu_0$  . כמו כן נתון שהמשתנה מתפלג נורמלית עם סטיית התקן ידועה  $\sigma$  מעוניינים לבצע מחקר שרמת המובהקות לא תעלה על  $\alpha$  והסיכוי לטעות מסוג שני לא上升 על  $\beta$  . הוכיחו שגודל המדגם הרצוי לכך יהיה :
- $$\cdot n \geq \left( \frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

**תשובות סופיות:**

- .41 (1)  
.78 (2)  
(3) שאלת הוכחה.

## mobekot\_tozacha - alfa\_minimalit (shevunot) האוכלוסייה ידועה):

**רקע:**

דרך נוספת להגעה להכרעות שלא דרך כלל הכרעה, היא דרך חישוב מובהקות התוצאות :

באמצעות תוצאות המדגם מחשבים את מובהקות התוצאה שמסומן ב-  $p_v$ .  
את רמת המובהקות החוקר קובע מראש לעומת זאת, את מובהקות התוצאה החוקר יוכל לחשב רק אחרי שייהיו לו את התוצאות.

המסקנה של המחקר תקבע לפי העיקרונו הבא : אם  $\alpha \leq p_v$ , דוחים את  $H_0$ .  
mobekot\_tozacha זה הסיכוי לקבלת תוצאות המדגם וקייזוני מתוצאות אלה בהנחה השערת האפס.

(לקבל את תוצאות המדגם וקייזוני)  $\cdot p_v = P_{H_0}$

אם ההשערה היא דו צדדיות :

(לקבל את תוצאות המדגם וקייזוני)  $\cdot p_v = 2P_{H_0}$

mobekot\_tozacha היא גם האלפא המינימלית לדחיתת השערת האפס.

$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_1: \mu > \mu_0$	$H_1: \mu < \mu_0$	השערת האפס : השערה אלטרנטיבית :
$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	. $\sigma$ ידועה		
$P_{H_0}(\bar{X} \geq \bar{x})$	$P_{H_0}(\bar{X} \leq \bar{x})$	$2 \cdot P_{H_0}(\bar{X} \geq \bar{x}) \iff \bar{x} > \mu_0$ $2 \cdot P_{H_0}(\bar{X} \leq \bar{x}) \iff \bar{x} < \mu_0$			<b>Tנאים :</b> <b>p-value</b>

כאשר בהנחה השערת האפס :  
 $Z_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} , \bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$

**דוגמה:**

המשקל הממוצע של מתגייסים לצבע לפני 20 שנה היה 65 ק"ג. מחקר מעוניין לבדוק האם כיום המשקל הממוצע של מתגייסים גבוה יותר. נניח שהמשקל המתגייסים מתפלג נורמלית עם סטטיסטיקה של 12 ק"ג. במדגם של 16 מתגייסים התקבל משקל ממוצע של 71 ק"ג.

- מהי מובהקות התוצאה?
- מה המסקנה אם רמת המובהקות היא 5% ואם רמת המובהקות היא ?!

**פתרון:**

a. אוכלוסייה: המתגייסים לצבע ביום.

משתנה:  $X$  = משקל בק"ג.

פרמטר:  $\mu$ .

השערות:  
 $H_0: \mu = 65$   
 $H_1: \mu > 65$

תנאים:

.  $X \sim N$ . 1

.  $\sigma = 12$ . 2

תוצאות מדגם:

$$n = 16$$

$$\bar{X} = 71$$

$$P_V = P_{H_0} \left( \text{لتוצאות המזגם וקיצוני} \right) = P_{H_0} (\bar{X} \geq 71) = 1 - \phi(2) = 1 - 0.9772 = 0.0228$$

$$Z_{\bar{x}} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{71 - 65}{12 / \sqrt{16}} = 2$$

$$\alpha_{\min} = 0.0228$$

**שאלות:**

- 1)** להלן השערות של מחקר:  $H_0: \mu = 70$ ,  $H_1: \mu > 70$ . המשתנה הנחקר מתפלג נורמלית עם סטיטית תקן 20. במדגם מאותה אוכלוסייה התקבלו התוצאות הבאות:  $\bar{x} = 74$ ,  $n = 100$ . מהי מובהקות התוצאה?
- 2)** השכר הממוצע במשק בשנת 2012 היה 8800 נס' עם סטיטית תקן 2000. במדגם שנעשה אטמול על 100 עובדים התקבל שכר ממוצע 9500 נס'. מטרת המחקר היא לבדוק האם כיים חלה עלייה בשכר. עבור אילו רמות מובהקות שיבחר החוקר יוחלט שהלאה עלייה בשכר הממוצע במשק?
- 3)** אדם חושד שהברת ממתקים לא עומדת בהתחביבוותה, ומשקלו של חטייף מסוים אותו הוא קונה מדי בוקר נזוק מ-100 גרם. חברות הממתקים טוענת מצידה שהיא אכן עומדת בהתחביבוותה. ידוע כי סטיטית התקן של משקל החטייף היא 12 גרם. האדם מתכוון לשקלול 100 חפיפות חטייפים ולאחר מכן מכון להגיע להחלטה.  
לאחר הבדיקה הוא קיבל משקל הממוצע של 98.5 גרם.  
א. רשמו את השערות המחקר.  
ב. מהי רמת המובהקות המינימלית עבורה דוחים את השערת האפס?  
ג. מהי רמת המובהקות המקסימלית עבורה קיבל את השערת האפס?  
ד. מה המסקנה ברמת מובהקות של 5%?
- 4)** מכונה לחישוק מוטות בפעול חותכת מוטות באורך שמתפלג נורמלית עם תוחלת אליה כוונה המכונה וסטיטית תקן 2 ס"מ. ביום מסוים כוונה המכונה לחישוך מוטות באורך 80 ס"מ. אחרי האיכות מעוניין לבדוק האם המכונה מכילה. לצורך כך נדרגו מקו הייזור 16 מוטות שנחתכו אורכו הממוצע היה 81.7 ס"מ.  
א. מהי רמת המובהקות המינימלית עבורה נカリע שהמכונה לא מכילה?  
ב. אם נסיף עוד ציפוי שערכה יהיה 82 ס"מ, כיצד הדבר ישפיע על התשובה של הסעיף הקודם?  
ג. הכרע ברמת מובהקות של 5% האם המכונה מכילה.
- 5)** אם מקבלים בחישובים לפחות מינימלית (value P) קטנה מאד, סביר להניח כי החוקר ידחה את השערת האפס בקלות. נכון/לא נכון? נמק.

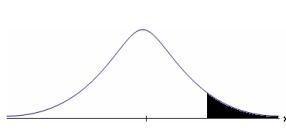
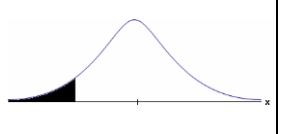
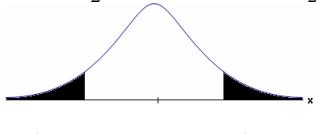
- 6) בבדיקה השערות התקבל שה-  $p-value = 0.02$ . מה תהיה מסקנת חוקר המשמש ברמת מובהקות 1%? בחרו בתשובה הנכונה.
- יקבל את השערת האפס בכל מקרה.
  - ידחה את השערת האפס מקרה.
  - ידחה את השערת האפס רק אם המבחן הנו דו צדדי.
  - לא ניתן לדעת כי אין מספיק נתונים.
- 7) מובהקות התוצאה (PV) היא גם (בחרו בתשובה הנכונה):
- רמת המובהקות המינימאלית לדוחות השערת האפס.
  - רמת המובהקות המקסימאלית לדחית השערת האפס.
  - רמת המובהקות שנקבעה מראש על ידי החוקר שטרם קיבל את תוצאות המחקר.
  - רמת המובהקות המינימאלית לאי דחית השערת האפס.
- 8) בבדיקה השערות מסוימת התקבל:  $p value = 0.0254$  לכן (בחרו בתשובה הנכונה):
- ברמת מובהקות של 0.01 אך לא של 0.05 נדחה את  $H_0$ .
  - ברמת מובהקות של 0.01 ושל 0.05 לא נדחה את  $H_0$ .
  - ברמת מובהקות של 0.05 אך לא של 0.01 נדחה את  $H_0$ .
  - ברמת מובהקות של 0.01 ושל 0.05 נדחה את  $H_0$ .

**תשובות סופיות:**

- (1) 0.0228 .  
 (2) עבר כל רמת מובהקות סבירה.  
 (3)  $H_0: \mu = 100$  .  
 $H_1: \mu < 100$  .  
 א. 0.1056 .  
 ב. 0.1056 .  
 ג. נכון שאיון כיוול.  
 ד. נכريع שיש עמידה בהתחייבות של החברה.  
 (4) א. 0.0006 .  
 (5) נכון.  
 (6) א'.  
 (7) א'.  
 (8) ג'.

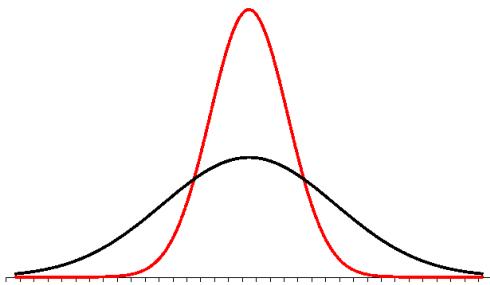
## בדיקות השערות על תוחלת (ממוצע) כשבונות האוכלוסייה לא ידועה:

**רקע:**

$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	<b>השערת האפס: השערה אלטרנטיבית:</b>
$H_1 : \mu > \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu \neq \mu_0$	
.1. $\sigma$ אינה ידועה או מוגן מספיק גדול $X \sim N$ .2			<b>תנאים:</b>
$t_{\bar{x}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha,n-1}$	$t_{\bar{x}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha,n-1}$	$t_{\bar{x}} < -t_{\frac{1-\alpha}{2},n-1}^{(n-1)}$ או $t_{\bar{x}} > t_{\frac{1-\alpha}{2},n-1}^{(n-1)}$  $-t_{\frac{1-\alpha}{2},n-1}$ $t_{\frac{1-\alpha}{2},n-1}$	<b>כל הבדיקה: אזור הדחיה של <math>H_0</math>:</b>
$\bar{X} > \mu_0 + t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} < \mu_0 - t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + t_{\frac{1-\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$ או $\bar{X} < \mu_0 - t_{\frac{1-\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$	<b>חלופה לכל הבדיקה: נדחה <math>H_0</math> אם מתקיים:</b>

$$\text{סטטיטיסטי המבחן: } t_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$$

**התפלגות T:**

הינה התפלגות סימטרית בעומניות שהתוחלת שלה היא 0. ההתפלגות דומה לתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים. התפלגות T תלויות במושג שנקרא דרגות החופש.

דרגות החופש הן:  $df = n - 1$ .

כל שדרגות החופש עולות בהתפלגות הופכת להיות יותר גבוהה וצרה. שדרגות החופש שוואות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

**דוגמה (פתרון בהקלטה):**

מפעל קיבל הזמנה לייצור משטחים בעובי של 0.1 ס"מ. כדי לבדוק האם המפעל עומד בדרישה נדגו 10 משטחים ונמצא שהעובי הממוצע הוא 0.104 עם אומדן לסטיתת תקן 0.002 ס"מ.

- מהו השערות המתקרי?
- מה ההנחה הדורשה לצורך פתרון?
- בודק ברמת מובהקות של 5%.

**שאלות:**

- 1)** משך זמן ההחלמה בלקיחת אנטיביוטיקה מסויימת הוא 120 שעות בממוצע עם סטיית תקן לא ידועה. מעוניינים לבדוק האם אנטיביוטיקה אחרת מקטינה את משך זמן ההחלמה. במדגם של 5 חולים שלקחו את האנטיביוטיקה האחראית התקבלו זמני ההחלמה הבאים: 125, 100, 95, 90, 80 שעות. מה מסקנתכם ברמת מובהקות של 5% מהי ההנחה הדרושה לצורך הפתרון?
- 2)** משרד הבריאות פרסם שמשקל ממוצע של תינוקות ביום היולדות בישראל 3300 גר'. משרד הבריאות רוצה לחקור את הטענה שנשים מעשנות בזמן ההריון يولדות תינוקות במשקל נמוך מהתמוצע. במחקר השתתפו 20 נשים מעשנות בהריון. להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:
- $$n = 20$$
- $$\bar{x} = 3120$$
- $$S = 280$$
- מה מסקנתכם ברמת מובהקות של 5% מה יש להניח לצורך פתרון?
- 3)** ציוני מבחן אינטיליגנציה מתפלגים נורמלית. באלה"ב ממוצע הציונים הוא 100. במדגם שנעשה על 23 נבחנים ישראלים, התקבל ממוצע ציונים 104.5 וסטיית התקן המדגמית 16. האם בישראל ממוצע הציונים שונה מאשר באלה"ב? הסיקו ברמת מובהקות של 5%.
- 4)** באוכלוסייה מסוימת נדגמו 10 תכפיות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{10} X_i = 750$$

$$\sum_{i=1}^{10} (X_i - \bar{X})^2 = 900$$

נתון שההתפלגות היא נורמלית.

בדוק ברמת מובהקות של 5% האם התוחלת של ההתפלגות שונה מ-80>.

- 5) ליאור ורוני העלו את אותן השערות על ממוצע האוכלוסייה. כמו כן הם התבפסו על אותן תוצאות של מדגם. ליאור השתמש בטבלה של התפלגות Z. רוני השתמשה בטבלה של התפלגות t. מה נוכל לומר בנוגע להחלטת המחקר שלהם? בחר בתשובה הנכונה.
- אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.
  - אם רוני תדחה את השערת האפס אז גם בהכרח ליאור.
  - שני החוקרים בהכרח הגיעו לאותה מסקנה.
  - לא ניתן לדעת על היחס בין דמיון השערת האפס של שני החוקרים.

- 6) נתנו ש:  $H_0: \mu = \mu_0$  ו-  $H_1: \mu < \mu_0$ .  $X \sim N(\mu, \sigma^2)$  כמו כן נתונות ההשערות הבאות:
- חוקר בדק את ההשערות הללו על סמך מדגם שככל 10 תצפיות.  $\sigma^2$  לא הייתה ידועה לחוקר. החוקר החליט לדחות את השערת האפס ברמת מובייקות של 5% לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תצפיות וشكلל את תוצאות אלה גם למדגם כך שככל עכשו 15 תצפיות. בחר בתשובה הנכונה:
- cut בברור הוא ידחה את השערת האפס.
  - cut הוא דוקא קיבל את השערת האפס.
  - cut לא ניתן לדעת מה תהיה מסקנתו.

### תשובות סופיות:

- 1) נדחה  $H_0$ .
- 2) נדחה  $H_0$ .
- 3) קיבל  $H_0$ .
- 4) קיבל  $H_0$ .
- 5) ב'.
- 6) ג'.

## mobekot\_tozacha - alfa\_minimalit (shevona) האוכלוסייה לא ידועה):

**רקע:**

נזכיר שהמסקנה של המבחן תיקבע לפי העיקרון הבא: אם  $\alpha \leq p_v$  דוחים את  $H_0$ .  
mobekot\_tozacha היא הסיכוי לקבלת תוצאות המדגם וקיצוני מהתוצאות אלה בהנחה השערת האפס.  
לקבל את תוצאות המדגם וקיצוני) .  
אם ההשערה היא דו צדדית :  
לקבל את תוצאות המדגם וקיצוני) .

mobekot\_tozacha היא גם האלפא המינימלית לדחינת השערת האפס.

$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	השערת האפס: השערה אלטרנטיבית:	
$H_1: \mu > \mu_0$	$H_1: \mu < \mu_0$	1. $\sigma$ אינה ידועה או 2. מדגם מספיק גדול $X \sim N$			
$P_{H_0}(\bar{X} \geq \bar{x})$	$P_{H_0}(\bar{X} \leq \bar{x})$	$2 \cdot P_{H_0}(\bar{X} \geq \bar{x}) \leftarrow \bar{x} > \mu_0$	$2 \cdot P_{H_0}(\bar{X} \leq \bar{x}) \leftarrow \bar{x} < \mu_0$	p-value	

$$t_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$$

$$d.f = n - 1$$

**דוגמה:**

ממוצע זמן הנסיעה של אדם לעובדה הינו 40 דקות. הוא מעוניין לבדוק דרך חלופית שאמורה להיות יותר מהירה. לצורך כך הוא דוגם 5 ימים שבהם הוא נוסע בדרך החלופית. זמני הנסיעה שקיבל בדיקות הם: 34, 40, 30, 32, 27. הנicho שזמן הנסיעה מתפלג נורמלית.

- א. רשמו את השערות המחקר.
- ב. מצאו חסמים לモבಹקות התוצאה.
- ג. מה המסקנה ברמת מובהקות של 5%?

**פתרון:**

אוכלוסייה: כלל הנסיעות לעובדה בדרך החלופית.

משתנה:  $X =$  זמן נסעה בדיקות.

תנאים:  $X \sim N$ .

פרמטר:  $\mu$ .

א. השערות:  
 $H_0: \mu = 40$   
 $H_1: \mu < 40$

ב. תוצאות המדגם:

$$n = 5, \bar{X} = \frac{\sum X_i}{n} = \frac{34 + 40 + \dots}{5} = 32.6$$

$$S^2 = \frac{\sum X_i^2 - n \cdot \bar{X}^2}{n-1} = \frac{34^2 + 40^2 + \dots - 5 \cdot 32.6^2}{5-1} = 23.4$$

$$S = \sqrt{23.4}$$

$$t_{\bar{X}} = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{32.6 - 40}{\frac{4.88}{\sqrt{5}}} = -3.39$$

$$P_V = P_{H_0} = (\bar{X} \leq 32.6) = P(t \leq -3.39)$$

$$d.f = 5 - 1 = 4$$

$$1\% < P_V < 2.5\%$$

$P_V < \alpha = 0.05$ , לכן דוחים את  $H_0$ .

מסקנה: בר"מ של 5% נכרייע שהדרך החלופית מהירה יותר.

**שאלות:**

- 1)** קוו ייצור אריזות סוכר נארזות כך שהמשקל הממוצע של אריזות הסוכר צריך להיות אחד קילוגרם. בכל יום דוגמים מקו היצור 5 אריזות במטרה לבדוק האם קו היצור תקין. בבדיקה דגמו 5 אריזות סוכר ולהלן משקלן בגרמים: 1024, 996, 1005, 997, 1008.
- רשמו את השערות המחקר.
  - מהי מובಹקות התוצאות? הצג חסמים.
  - מה המסקנה ברמת מוב hawkות של 5%?
- 2)** חוקר בדק את הטענה כי פועלים העובדים במשמרתليل האיטיים יותר מפועלים העובדים ביום. ידוע כי משך הזמן הממוצע הדרוש לייצר מוצר מסוים ביום הוא 6 שעות. בדוגמא מיקרי של 25 פועלים שעבדו במשמרתليل נמצא כי הזמן הממוצע לייצר אותו מוצר הוא 7 שעות עם סטיית תקון של 3 שעות. מהי  $\alpha$ -המינימלית שלפיה ניתן להחליט שancock העובדים במשמרתليل האיטיים יותר?
- 3)** הגובה של מתגייםים לצה"ל מתפלג נורמלית. בדוגמא של 25 מתגייםים מדדו את הגבהים שלהם בס"מ והתקבלו התוצאות הבאות:
- $$\sum (x_i - \bar{x})^2 = 2832, \bar{x} = 176.2$$
- מטרת המחקר היא לבדוק האם תוחלת הגבהים של המתגייםים גבוהה מ-174 ס"מ באופן מובהק. מהי בקרוב מוב hawkות התוצאות ועל פייה מה תהיה המסקנה ברמת מוב hawkות של 6%?

**תשובות סופיות:**

- 1)** א.  $H_0: \mu = 1000$       ב.  $20\% \leq P_v \leq 50\%$   
 $H_1: \mu \neq 1000$
- ג. ברמת מוב hawkות של 5% לא נוכל לקבוע שקו היצור אינו תקין.
- 2)**  $.10\%$
- 3)** נקבל את  $H_0, 1.01$

## הקשר בין רוח סמך לבדיקה השערות על תוחלת (מומוצע):

**רקע:**

ניתן לבצע בדיקת השערות דו צדדיות ברמת מובהקות  $\alpha$  על  $\mu$  :

$$\mu_0 : \mu = \mu_1 , H_0 : \mu \neq \mu_0$$

על ידי בניית רוח סמך ברמת סמך של  $\alpha - 1$  ל-  $\mu$  :

אם  $\mu_0$  נופל ברווח  $\leftarrow$  קיבל את  $H_0$ .

אם  $\mu_0$  לא נופל ברווח  $\leftarrow$  נדחה את  $H_0$ .

**דוגמה:**

חוקר ביצע בדיקת השערות לתוחלת. להלן השערותיו :

$$H_0 : \mu = 80 , H_1 : \mu \neq 80 , \alpha = 5\%$$

החוקר בנה רוח סמך ברמה של 90% וקיבל:  $84 < \mu < 79$ .

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

**פתרון (פתרון מלא בהקלטה):**

רוח הסמך ברמת סמך של 90% מכיל "80".

ברמת סמך של 95% רוח הסמך יגדל וכייל "80".

לכן, ברמת מובהקות של 5% קיבל  $H_0$ .

**שאלות:**

- 1)** חוקר רצה לבדוק את ההשערות הבאות:  $H_0: \mu = 90$ ,  $H_1: \mu \neq 90$ . החוקר בנה רוחח סמך לתוכלת ברמת סמך של 95% וקיבל את רוחח הסמך הבא: (87, 97). אם החוקר מעוניין לבצע בדיקת השערות ברמת מובהקות של 1% האם ניתן להגיע למסקנה ע"י רוחח הסמך? נמקו.
- 2)** חוקר מעוניין לבדוק השפעת דיאטה חדשה על רמת הסוכר בدم. ידוע כי מספר מיליגרים הסוכר בסמ"ק דם הוא משתנה מקרי שמתפלג נורמלית עם סטיית תקן 10.4 מ"ג. נלקח מדגם של 60 נבדקים שניזונו מדיאטה זו. נמצא כי ממוצע מספר המיליגרים סוכר היה 115.5 מ"ג לסמ"ק.
- א. בנה רוחח סמך ברמת סמך 95% לתוכלת רמת הסוכר בדם אצל הניזונים מדיאטה זו.
- ב. ידוע שתוחלת רמת הסוכר בדם באוכלוסייה היא 90 מ"ג לסמ"ק. האם לדעתך ניתן להסיק על סמך תוצאת סעיף א' שהדיאטה משפיעה על רמת הסוכר בדם? הסבירו.
- 3)** יצרן אנטיביוטיקה רושם על גבי התרופות שכמות הפנצליין היא 200 מ"ג لكפסולה. משרד הבריאות ביצע מדגם של 8 קפסולות אקרראיות מקו הייצור ומצא שבממוצע יש 196 מ"ג פנצליין لكפסולה עם סטיית תקן מדגמית של 5 מ"ג. בהנחה וכמות הפנצליין בקפסולה מתפלגת נורמלית.
- א. בנו רוחח סמך ברמת סמך של 95% למומוצע כמות הפנצליין لكפסולה המיוצרת על ידי יצרן האנטיביוטיקה.
- ב. בדקו ברמת מובהקות של 5% האם יש אמת באינפורמציה המسطופקת על ידי הייצן.

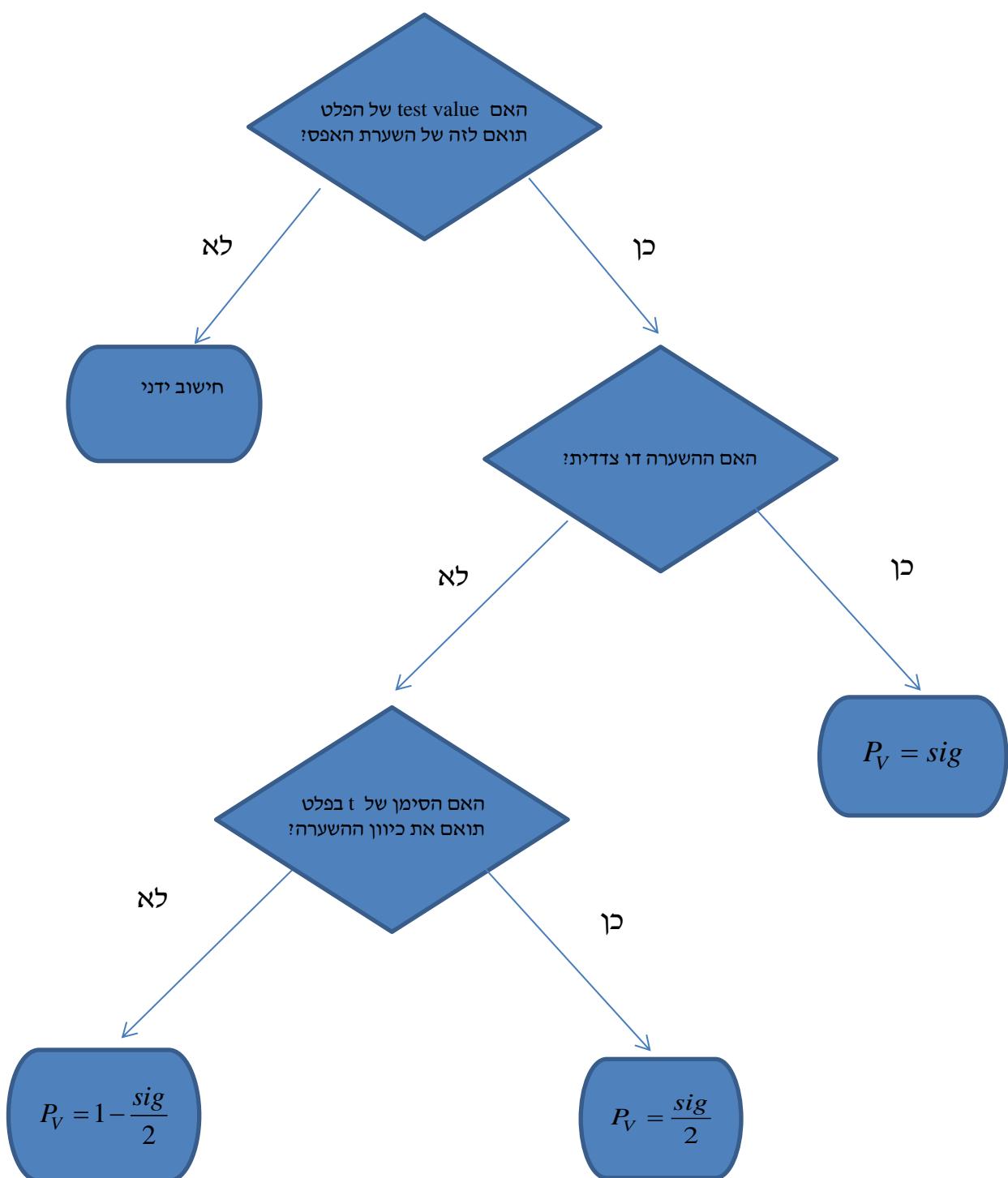
**תשובות סופיות:**

- 1)** קיבל השערת.
- 2)** א.  $112.87 \leq \mu \leq 118.13$   
ב. נזכיר שהדיאטה משפיעה על תוחלת רמת הסוכר בדם.
- 3)** א.  $200.2 \leq \mu \leq 191.8$ . ב. נזכיר שיש אמת בפרסום.

## ניתוח פלטים:

רעיון:

חישוב מובהקות התוצאה באמצעות פלט תוכנת SPSS :



**דוגמה (פתרון בהקלטה):****One-Sample Statistics**

	N	Mean	Std. Deviation	Std. Error Mean
X	25	87.6400	64.90434	12.98087

**One-Sample Test**

X	Test Value = 60					
	t	df	Sig. (2-tailed)	Mean	95% Confidence Interval of the Difference	
				Difference	Lower	Upper
X	2.129	24	.044	27.64000	.8488	54.4312

ממוצע הציונים ב מבחן המיצב בחשבון הוא 60. הוחלט לדגם כייתה אקראית של 25 תלמידים ולמד אותם בשיטת לימוד חדשה.

- א. מהו רוח הסמק לVERAGE הציונים בחשבון אם יוחלט ליישם את שיטת הלימוד החדש?
- ב. מהו  $P_{\text{value}}$  לבדיקהיעילותה של שיטת הלימוד החדש?
- ג. מה יוכרע ברמת מובהקות של 5% לגבי יעילותה של שיטת הלימוד החדש?

**שאלות:**

- 1) באוניברסיטה גודלה גיל הסטודנטים לתואר ראשון מתפלג נורמלית. בעבר פורסם שהגיל הממוצע של הסטודנטים הינו 23. להלן פلت תוכנת SPSS על מדגם של 16 סטודנטים אקראים מתואר ראשון:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
age	16	23.4375	2.50250	.62562

One-Sample Test

	Test Value = 23					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
age	.699	15	.495	.43750	-.8960	1.7710

- א. מהו המבחן הסטטיסטי שנעשה כאן?  
 ב. מה ערכו של הפרמטר לפי השערת האפס?  
 ג. רשום רוח סמך ברמת סמך של 95% לתוחלת גיל הסטודנטים באוניברסיטה לתואר ראשון.  
 ד. בדוק ברמת מובהקות של 5% האם גיל הממוצע ביום שונה מבעבר?
- 2) קבוצת ילדים בגיל 6 קיבלה משימה לביצוע. עברו כל ילד בדקו כמה זמן לוקח לו לסיים את המשימה בדקות. להלן תוצאות הניטוח הסטטיסטי:

One-Sample Test

	Test Value = 4.5					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
time	-1.853	24	.076	-.09200	-.1944	.0104

- א. כמה ילדים השתתפו במחקר?  
 ב. מצא רוח סמך ברמת סמך של 95% לתוחלת זמן ביצוע המשימה עברו ילדים בני 6.  
 ג. מה יש להניח כדי שרוח הסמך מסעיף א' יהיה תקף?  
 ד. בדוק ברמת מובהקות של 5% שזמן ביצוע המשימה הממוצע נמור מ-4.5 דקות.

(3) להלן פלט מחשב עבור ניתוח סטטיסטי שנעשה בתוכנת SPSS. הניתוח הוא עבור מבחן אקראי של קבוצת נבחנים בבריות באנגלית.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
grade	???	???	19.62787	2.95901

One-Sample Test

	Test Value = 75					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
grade	???	43	.017	-7.34091	-13.3083	????

- א. השלימו את הגדרים החסרים המסומנים בסמני שאלה בפלט.
- ב. מהי מובאות התוצאות לבדיקת ההשערה שהתוחלת של הציונים שונה מ-75?
- ג. מהי מובאות התוצאות לבדיקת ההשערה שהתוחלת של הציונים קטנה מ-75?
- ד. מהי מובאות התוצאות לבדיקת ההשערה שהתוחלת של הציונים גדולה מ-75?

(4) יצרו סיגריות מפרסם כי תוחלת הניקוטין בסיגריות שהוא מיצר קטנה מ-27 מ"ג. בבדיקה מקרית של 5 סיגריות מתוצרתו נמצאו כמותות הניקוטין הבאות : 21, 20, 24, 22 מ"ג. הניחו כי כמות הניקוטין בסיגריות מפולג נורמלי.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
nicotine	5	21.6000	1.51658	.67823

One-Sample Test

	Test Value = 27					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
nicotine	-7.962	4	.001	-5.40000	-7.2831	-3.5169

- א. האם ברמת מובהקות של 5% ניתן להסיק שיש אמת בפרסום?
- ב. אם היינו מוסיפים עוד תצפית שערך 20. כיצד הדבר היה משנה על ערך Sig ועל המסקנה?
- ג. בדקו האם ניתן להגיד שתוחלת רמת הניקוטין שונה מ-26 ברמת מובהקות של 5%.

**תשובות סופיות:**

- (1) א. הסקה של תוחלת אחת. ב. 23. ג. (22.104, 24.771).
- ד. נקלט  $H_0$ .
- (2) א. 25. ב. (4.3056, 4.5104). ג. המשטנה מתפלג נורמלית.
- ד. נדחה  $H_0$ .
- (3) א.  $\bar{X} = 67.66$ ,  $n = 44$ . ב. 0.017. ג. 0.0085. ד. 0.9915.
- (4) א. נכרייע שיש אמת בפרסום. ב. המסקנה לא תשתנה.  
ג. נכרייע שהתוחלת שונה מ-26.

# bijustisjka le-refoat shinim

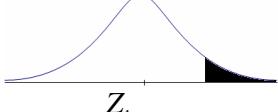
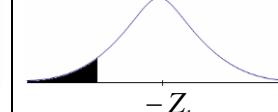
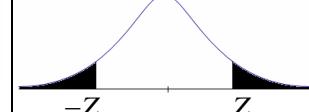
## פרק 37 - בדיקת השערות על פרופורציה

### תוכן העניינים

197	1. התהליין
200	2. סיכון לטעויות ועוצמה
204	3. קביעת גודל מוגן
206	4. מובהקותות התוצאה - אלף מינימלית

## התהילה:

רקע:

$H_0 : p = p_0$	$H_0 : p = p_0$	$H_0 : p = p_0$	השערת האפס: השערת אלטרנטיבית:		
$H_1 : p > p_0$	$H_1 : p < p_0$	$H_1 : p \neq p_0$	<b>תנאים:</b> $np_0 \geq 5 \text{ \& } n(1-p_0) \geq 5$		
$Z_{\hat{p}} > Z_{1-\alpha}$  $Z_{1-\alpha}$ $H_0$ -דוחים את 	$Z_{\hat{p}} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ $H_0$ -דוחים את 	$Z_{\hat{p}} < -Z_{\frac{1-\alpha}{2}}$ או $Z_{\hat{p}} > Z_{\frac{\alpha}{2}}$  $-Z_{\frac{1-\alpha}{2}}$ $Z_{\frac{1-\alpha}{2}}$ $H_0$ -דוחים את 	כלל הבדיקה: אזור הדחיה: של $H_0$		

$$\text{סטטיסטי המבחן: } Z_{\hat{p}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

חלופה אחרת לכלל הבדיקה:

כלל הבדיקה – אזור הדחיה של $H_0$ :		
$\hat{p} > p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} < p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} > p_0 + Z_{\frac{1-\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ $\hat{p} < p_0 - Z_{\frac{1-\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$

דוגמה (פתרון בהקלטה):

בחודש ינואר השנה פורסם שאחוז האבטלה במשק הוא 8% במדגם עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים.  
בדקו ברמת מובהקות של 5% האם כיום האבטלה הוא כמו בתחילת השנה.

**שאלות:**

- 1)** במשך שנים אחזו המועמדים שהתקבל לפוקולטה מסוימת היה 25%. השנה מתוך מוגם של 120 מועמדים התקבלו 22. בرمת מובהקות של 5% האם השנה הקשו על תנאי הקבלה?
- 2)** במדגם של 300 אזרחים 57% מתנגדים להצעת חוק מסוימת. לאור נתונים אלה האם רוב האזרחים מתנגדים להצעת החוק? בדקו ברמת מובהקות של 10%.
- 3)** הטילו מטבח 50 פעמים וקיבלו 28 פעמים עז. האם המטבח הוגן ברמת מובהקות של 5%?
- 4)** קפיטריה במכלה מסוימת מעירica כי אחזו הסטודנטים שוכנים קפה בקפיטריה הינו 20%. נערך סקר אשר כלל 200 סטודנטים. התברר כי 33 מהם רוכשים קפה בקפיטריה. מטרת הסקר הייתה לבדוק את אמינותה הערכה של הקפיטריה.
- רשמו את ההשערות.
  - בדקו את ההשערות ברמת מובהקות של 10%.
  - מה תהיה המסקנה אם נקטין את רמת המובהקות?
- 5)** חבר כנסת רוצה להעביר חוק. לצורך כך הוא דוגם 400 אזרחים במטרה לבדוק האם רוב האזרחים תומכים בחוק. במדגם התקבל ש-276 אזרחים תומכים בחוק.
- מה מסקנתכם ברמת מובהקות של 5%?
  - אם ניתן לדעת מה תהיה המסקנה אם רמת המובהקות תהיה גדולה יותר? הסבירו.
- 6)** שני חוקרים בדקו את ההשערות הבאות:  $H_0: p = p_0$ ,  $H_1: p > p_0$ . חוקר א' השתמש ברמת מובהקות  $\alpha_1$  וחוקר ב' ברמת מובהקות  $\alpha_2$  החוקר הראשון דחה את  $H_0$  ואילו החוקר השני קיבל את  $H_0$ .
- שנייהם התבססו על אותן תוצאות של מוגם. בחר בתשובה הנכונה:
- $\alpha_1 = \alpha_2$ .
  - $\alpha_1 > \alpha_2$ .
  - $\alpha_1 < \alpha_2$ .
  - המצב המתואר לא אפשרי.

**תשובות סופיות:**

- (1) נדחה  $H_0$ .  
 (2) נדחה  $H_0$ .  
 (3) קיבל  $H_0$ .
- ב. נקבל  $H_0$ . ג. המסקנה לא תשתנה.
- ב. המסקנה לא תשתנה. א. נדחה  $H_0$ .  
 (4)  $H_0 : p = 0.2$  .  
 $H_1 : p \neq 0.2$  .  
 (5) א. נדחה  $H_0$ .  
 (6) ג' .

## סיכום לטיעויות ועוצמה:

**רקע:**

**הגדרת הסטבריות:**

הסיכוי לבצע טיעות מסוג 1 (רמת מובהקות) :  
 $(\text{לדוחות } H_0 = P_{H_0} (H_0 \text{ נכונה}) | \text{ לדוחות את } \alpha = P_{H_0} (H_0 \text{ לא נכון}))$

הסיכוי לבצע טיעות מסוג 2 :  
 $(\text{לקבל } H_0 = P_{H_1} (H_1 \text{ נכונה}) | \text{ לדוחות את } \beta = P_{H_1} (H_1 \text{ לא נכון}))$

רמת בטחון :  
 $(\text{לקבל } H_0 = P_{H_0} (H_0 \text{ נכונה}) | \text{ לדוחות את } 1 - \alpha = P_{H_0} (H_0 \text{ לא נכון}))$

עוצמה :  
 $(\text{לקבל } H_1 = P_{H_1} (H_1 \text{ נכונה}) | \text{ לדוחות את } \pi = 1 - \beta = P_{H_1} (H_1 \text{ לא נכון}))$

		הכרעה	
מציאות		$H_0$	$H_1$
	$H_0$	אין טיעות	טיעות מסוג 1
	$H_1$	טיעות מסוג 2	אין טיעות

**התהליך לחישוב סיכוי לטיעות מסוג שני:**

$H_0 : p = p_0$ $H_1 : p > p_0$	$H_0 : p = p_0$ $H_1 : p < p_0$	$H_0 : p = p_0$ $H_1 : p \neq p_0$	השערת האפס: השערת אלטרנטיבית:
$np_0 \geq 5 \& n(1-p_0) \geq 5$			תנאים:
$\hat{p} > p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} < p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} > p_0 + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ או $\hat{p} < p_0 - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	כל ההכרעה: אזור הדחיה של $H_0$

חישוב : $\beta$
$P_{H_1} \left( \hat{p} < p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} \right)$
$P_{H_1} \left( p_0 - Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} < \hat{p} < p_0 + Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} \right)$
$P_{H_1} \left( \hat{p} > p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} \right)$

כאשר :  $\hat{P} \sim N \left( p, \frac{p(1-p)}{n} \right)$

והתקנו :  $Z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

#### דוגמה (פתרון בהקלטה) :

רופא ShinNim טען כי שיעור ממחצית האוכלוסייה הבוגרת בארץ אינם מבקרים אצל רופא ShinNim באופן קבוע, כנדרש. כדי לבדוק טענה זו, נערך סקר בקרב 150 אנשים בוגרים.

- .א. רשמו את ההשערות וכלל הכרעה ברמת מובהקות של 10%.
- .ב. מהי עוצמת המבחן אם מסתבר ש 60% מהאוכלוסייה אינם מבקרים אצל רופא ShinNim באופן קבוע.

**שאלות:**

- 1)** משרד הבריאות פרסם ש-10% מתושבי המדינה סובלים ממחלה האסתמה. מחקר דורך לבדוק האם בחיפה, בגל זיהום האוויר, שיורט הסובלים מאסתמה גביה יותר. לצורך המחקר נבדקו 260 מתושבי חיפה.  
 א. רשמו את השערות המחקר, וצרו מבחן ברמת מובהקות של 5% לבדיקה.  
 ב. מהי עצמת המבחן של סעיף א' בהנחה ובחיפה 16% מהתושבים סובלים מאסתמה?  
 ג. כיצד תנסה התשובה לסעיף ב' אם מסתבר שבחיפה 18% סובלים מאסתמה?  
 ד. בהמשך לסעיף א' האם נכון לומר שבנסיבות של 5% ההשערה שבחיפה 10% מהתושבים סובלים מאסתמה אינה נכונה?
- 2)** אחוז הסובלים מתופעות הלוואי מתרופה מסוימת הוא 15%. חברת תרופות טוענת שפיתחה תרופה שאמורה לצמצם את אחוז הסובלים מתופעות לוואי. לצורך בדיקת הטענה הוחלט לבצע מחקר שיכלול 120 חולים שיקבלו את התרופה הנבדקת.  
 נניח שהתרופה נבדקה אכן מורידה את פרופורציות הסובלים מתופעות הלוואי ל-10%, מהי עצמת המבחן עבור רמת מובהקות של 5%?
- 3)** בעיר מסוימת היו 20% אקדמאים. בעקבות פтиחת מכלה בעיר לפני כמה שנים מעוניינים לבדוק האם אחוז האקדמאים גדול. מעוניינים שהמחקר יכלול 200 אנשים והוא יהיה ברמת מובהקות של 5%.  
 א. חשבו את הסיכוי לבצע טעות מסווג שני בהנחה והיום יש 28% אקדמאים.  
 ב. כיצד התשובה לסעיף הקודם תשתנה אם נגדיל את רמת המובהקות?
- 4)** מעוניינים לבדוק האם בפקולטה מסוימת ישנה העדפה לגברים. הוחלט לדגום 200 מתקובלים ועל סמך מספר הבנים לקבוע אם טענת המחקר מתק傍ת. חוקר אי קבע רמת מובהקות של 5% וחוקר ב' החליט לקבל את טענת המחקר אם במדגם יהיו לפחות 120 בניים. למי מבין החוקרים רמת מובהקות גדולה יותר?
- 5)** חוקר ביצע מחקר ובו עשה טעות מסווג שני בכך (בחרו בתשובה הנכונה):  
 א. השערת האפס נכונה.  
 ב. השערת האפס נדחתה.  
 ג. השערת האפס לא נדחתה.  
 ד. אף אחת מהטעויות לא נכונה בהכרח.
- 6)** קבעו אם הטענה הבאה נכונה: בבדיקה השערות לא ניתן לבצע בו זמני טעות מסווג ראשון וטעות מסווג שני.

**תשובות סופיות:**

- .0.9015      ב. גודל.      ג. טענה לא נכונה.  
 $H_0 : p = 0.1$       א.  $H_1 : p > 0.1$       (1)
- .0.4404      (2)
- ב. תקין.      0.1446      (3)
- (4) חוקר א'.  
 (5) ג'.  
 (6) נכונה.

## קביעת גודל מוגן:

**רקע:**

השערות המחקר הן:  $H_0: p = p_0$ ,  $H_1: p = p_1$  מעוניינים לבצע מחקר שרמת המובהקות לא תעלה על  $\alpha$  והסיכוי לטעות מסוג שני לא תעלה על  $\beta$ .

הנוסחה הבאה נותנת את גודל המוגן הרצוי:

$$n \geq \left( \frac{Z_{1-\alpha} \sqrt{p_0 q_0} + Z_{1-\beta} \sqrt{p_1 q_1}}{p_0 - p_1} \right)^2$$

**דוגמה (פתרון בהקלטה):**

רוצים לבדוק האם אחוז האנשים השווים בשמש ללא הגנה ירד בעקבות הפרסומת על נזקי השימוש. בעבר 60% מהאוכלוסייה שהתחבשה בשמש ללא הגנה. מה גודל המוגן המינימלי שיש לקחת כדי לבדוק שהאחוז הניליל ירד לפחות 48% אם מעוניינים שהסיכוי לטעות מסוג ראשון יהיה 5% והסיכוי לטעות מסוג שני יהיה 1%?

**שאלות:**

- 1)** משרד התמ"ת פרסם ש אחוז האבטלה במשק היום עומד על 8%. לעומת זאת, משרד הפנים טוען ש חלה עלייה בשיעור האבטלה עד לכדי 11%. כדי לבדוק מי מבניהם צודק, מה צריך להיות גודל המדגם שייננה על שני התנאים הבאים:
- אם משרד התמ"ת צודק, נדחה את טענתו בסיכוי של 10%.
  - אם משרד הפנים צודק, נדחה את טענתו בסיכוי של 4%.
- 2)** מפעיל קזינו מפרסם שהסיכוי לזכות במכונות מזל הינו 0.42. אדם טוען שהסיכויים לזכות במשחק נמוכים יותר. כמה פעמים יש לשחק את המשחק כדי שאם טענת מפעיל הקזינו נכונה נקבל את טענת האדם בסיכוי של 1% ואם למציאות הסיכוי לזכות במכונה הוא 0.3 נקבל את מפעיל הקזינו בסיכוי של 8%?

**תשובות סופיות:**

.891 (1)

.224 (2)

## МОבקות התוצאה – אלף מינימלית:

**רעיון:**

דרך נוספת להגעה להכרעות שלא דרך כלל הכרעה, היא דרך חישוב מובהקות התוצאה: באמצעות תוצאות המדגם מחשבים את מובהקות התוצאה שמסומן ב-  $p_v$ . את רמת המובהקות החוקר קובע מראש לעומת זאת, את מובהקות התוצאה החוקר יכול לחשב רק אחרי שייהו לו את התוצאות. המשקנה של המחקר תקבע לפי העיקרון הבא:

אם  $p_v \leq \alpha$  דוחים את  $H_0$ .

מובהקות התוצאה זה הסיכוי לקבלת תוצאות המדגם וקיוצוני מתוצאות אלה בהנחה השערת האפס.

לקבל את תוצאות המדגם וקיוצוני  $\cdot p_v = P_{H_0}$ .

אם ההשערה היא דו צדדית:

לקבל את תוצאות המדגם וקיוצוני  $\cdot p_v = 2P_{H_0}$

מובהקות התוצאה היא גם האלפא המינימלית לדחינת השערת האפס.

השערת האפס:	השערה האלטרנטיבית:
$H_0: p = p_0$	$H_0: p = p_0$
$H_1: p > p_0$	$H_1: p < p_0$
$np_0 \geq 5 \& n(1-p_0) \geq 5$	תנאים:
$P_{H_0}(\hat{P} \geq \hat{p})$	$2 \cdot P_{H_0}(\hat{P} \geq \hat{p}) \leftarrow \hat{p} > p_0$ אם $2 \cdot P_{H_0}(\hat{P} \leq \hat{p}) \leftarrow \hat{p} < p_0$ אם
$P_{H_0}(\hat{P} \leq \hat{p})$	<b>p-value</b>

כאשר בהנחה השערת האפס:  $\hat{P} \sim N\left(p_0, \frac{p_0(1-p_0)}{n}\right)$

התקנון:  $Z_{\hat{p}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$

**דוגמה (פתרון בהקלטה):**

ישנה טענה שיש הבדל בין אחוז הבנים ואחוז הבנות הפונאים ללימוד להנדסאי מחשבים. לשם כך נלקח מוגם מקרי של 200 תלמידים הלומדים מחשבים והתברר כי 112 מהם בנים.

א. מהי מובاهקות התוצאות?

ב. מה המסקנה ברמת מובاهקות של 5%?

**שאלות:**

- 1)** במשך שנים אוחז המועמדים שהתקבל לפקולטה מסוימת היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. רוצים לבדוק האם האם השנה הקשו על תנאי הקבלה.
- מהי מובהקות התוצאה?
  - מה תהיה המסקנה ברמת מובהקות של 1% וברמת מובהקות של 5%?
- 2)** נהוג לחשב ש-60% מהילדים בגיל שלוש קמים מהORITY במהלך הלילה לפחות פעם אחת. ישנה טענה שלאו שנות צהרים פחות מ-60% מהילדים בגיל זה יקומו לפחות פעם אחת במהלך הלילה. נדגו 80 ילדים בגיל 3 אשר אינם ישנים בצהרים מתוכם התקבל ש-41 קמו במהלך הלילה.
- מהי רמת המובהקות המינימלית עבורה שתתקבל הטענה במחקר?
  - מהי רמת המובהקות המקסימלית עבורה לא תתקבל טענת המחקר?
  - עבור אילו רמות מובהקות קיבל את טענת המחקר?
  - מה תהיה מסקנת המחקר ברמת מובהקות של 6%?
- 3)** במטרה לבדוק האם מטבח הוגן מטילים אותו 80 פעמים. התקבל ש-60 מההטלות הראו עצ. רשמו את השערות המחקר, חשבו את מובהקות התוצאה והסיקו מסקנה ברמת מובהקות של 5%.
- 4)** בבדיקה השערות על פרופורציה התקבל שה- $p-value = 0.02$ . מה תהיה מסקנת חוקר המשמש ברמת מובהקות 5%:  
(בחרו בתשובה הנכונה)
- קיבל את השערת האפס
  - ידחה את השערת האפס.
  - לא ניתן לדעת כי אין מספיק נתונים.
- 5)** קבעו אם הטענה הבאה נכונה:  
"בבדיקה השערות חד-צדדי התקבל ערך  $p-value$  של 3%, לכן אם היינו מבצעים מבחן דו-צדדי (כאשר יתר הנסיבות ללא שינוי), היינו מקבלים ערך  $p-value$  של 6%".
- 6)** במפעל 10% מהעובדים נפגעים לפחות פעם אחת בשנה מתאונות עבודה. לאור זאת, המפעל החליט לצאת בתוכנית לצמצום שיעור הנפגעים. תוכנית זו נוסתה על 100 עובדים. מתוכם 12 נפגעו בתאונות עבודה במשך השנה. מהי רמת המובהקות הקטנה ביותר עבורה יוחלט שהתוכנית יעילה?

**תשובות סופיות:**

(1) א. 0.0455

ב. ברמת מובהקות של 1% : לא דוחים את  $H_0$ .ברמת מובהקות של 5% : נדחה את  $H_0$ .

(2) א. 0.0548      ב. 0.0548      ג. מעל 0.0548

ד. נכרייע לטובת טענת המחקר.

(3)  $H_0$ , נדחה את  $p_v = 0$ 

(4) ב'.

(5) הטענה נכונה.

(6) 0.7486

# bijustisjka le-refoat shinim

פרק 38 - בדיקת השערות על הפרש תוחלות במדגים בלתי תלויים

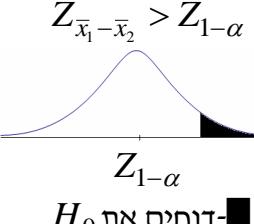
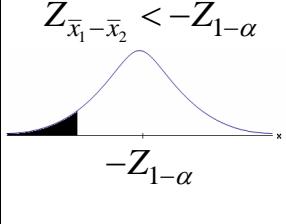
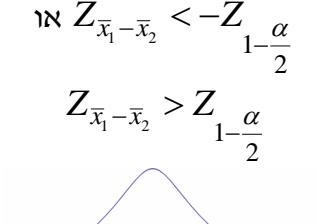
## תוכן העניינים

210 .....	1. כשלונות האוכלוסייה ידועות.....
214 .....	2. כשלונות האוכלוסייה לא ידועות ומণחים שהן שווות.....
218 .....	3. ניתוח פלטימ.....

## בדיקות השערות על הפרש תוחלות במדגמים בלתי תלויים

---

### כשהשונות של האוכלוסייה ידועות – רקע

$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	השערת האפס: השערת אלטרנטיבית:
$H_1 \quad \mu_1 - \mu_2 > c$	$H_1 \quad \mu_1 - \mu_2 < c$	$H_1 \quad \mu_1 - \mu_2 \neq c$	מדגמים בלתי תלויים $\sigma_1, \sigma_2$ $X_1, X_2 \sim N$ או מדגמים מספיק גדולים
			תנאים:  כלל ההכרעה: אזרור הדחיה של $H_0$
$Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\alpha}$  $Z_{1-\alpha}$ $H_0$ דוחים את ■	$Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ $H_0$ דוחים את ■	$Z_{\bar{x}_1 - \bar{x}_2} < -Z_{\frac{1-\alpha}{2}}$ או $Z_{\bar{x}_1 - \bar{x}_2} > Z_{\frac{1-\alpha}{2}}$  $-Z_{\frac{1-\alpha}{2}} \quad Z_{\frac{1-\alpha}{2}}$ $H_0$ דוחים את ■	$-Z_{\frac{1-\alpha}{2}} \quad Z_{\frac{1-\alpha}{2}}$

**סטטיסטי המבחן:**  $Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

**חלופה אחרת לכלל הכרעה:**

נחתה $H_0$ אם מתקיים :	
$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ <b>או</b> $\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	

**התפלגות הפרש המומוצעים:**  $\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

$$\text{התקנון: } Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

**דוגמה (פתרו בהקלטה) :**

בשנת 2004 הופיע בין השכר הממוצע של הגברים לנשים היה 3000₪ לטובת הגברים. מעוניינים לבדוק האם כיוון הצטמצם הופיע בין הגברים לנשים מבחינת השכר הממוצע. נדגומו 100 עובדים גברים. שכרם הממוצע היה 9,072 ₪. נדגומו 80 עובדים, שכרכו הממוצע היה 9,780 ₪. לצורך פתרון נניח שסטיות התקן של השכר ידועות ו שוות ל-2000₪ באוכלוסייה הנשים ו-3000₪ באוכלוסייה הגברים. מה המסקנה ברמת מובהקות של 5%?

## שאלות

- 1) מחקר טוען שאנשים החיים במרכז הארץ צופים בממוצע בטלוויזיה יותר מאשרים שלא חיים במרכז. נדגו 100 אנשים מהמרכז ו-107 אנשים לא מהמרכז. אנשים אלה נשאלו כמה שעות ביום הם נוהגים לצפות בטלוויזיה. במדגם של מרכז הארץ התקבל ממוצע 2.7 שעות. במדגם של מחוז למרכז הארץ התקבל ממוצע 1.8 שעות. לצורך פתרון הניחו שככל אזור, סטיית התקן היא שעה 1 ביום. בדקו את טענת המחקר ברמת מובהקות של 1%.
- 2) ציוני פסיכומטרי מתפלגים נורמלית עם סטיית התקן 100. מכון ללימוד פסיכומטרי טוען שהוא יכול לשפר את ממוצע הציונים ביותר מ-30 נקודות. במדגם של 20 נבחנים שניגשו לבחן ללא הינה במכון התקבל ממוצע 508. במדגם של 25 נבחנים שעברו הינה במכון התקבל ממוצע ציוניים 561. מה מסקנכם ברמת מובהקות של 5%.
- 3) במדגם אקראי של 20 ימים נבדקה התפוקה של מפעל ביום. התפוקה הממוצעת הייתה של 340 מוצרים ליום. במדגם אקראי של 20 ימים אחרים נבדקה התפוקה של המפעל בלילה וההתפוקה הממוצעת הייתה 295. לצורך פתרון נניח שסטיית התקן של התפוקה ביום היא 40 מוצרים ובלילה 30 מוצרים.  
א. מהי מובהקות התוצאה לבדיקה האם התפוקה הממוצעת היומית גבוהה מההתפוקה הממוצעת הלילית.  
ב. מה תהיה המסקנה ברמת מובהקות של 8%?
- 4) במחקר מקייף שנעשה באירופה נקבע שגברים גבוהים מנשים ב-8 ס"מ ממוצע. מחקר ישראלי מתעניין לבדוק האם בישראל הפער גדול יותר. לצורך המחקר נדגו 40 גברים ו 40 נשים באקראי. כמו כן, נניח שסטיות התקן של הגברים והנשים ידועות ושותת ל-6 ס"מ אצל הנשים ו-12 ס"מ אצל הגברים.  
א. מהן השערות המחקר ומהו כלל ההכרעה ברמת מובהקות של 10%?  
ב. אם בישראל הפער בין גברים לנשים מבחינת הגובה הממוצע הוא 11 ס"מ, מה ההסתברות שהמחקר לא יגלה זאת? איך קוראים להסתברות זאת?

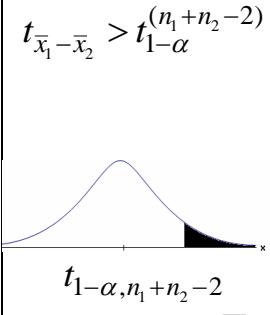
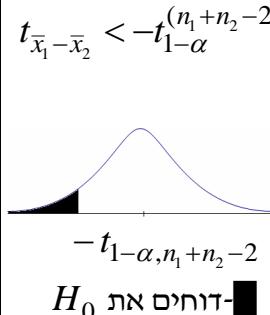
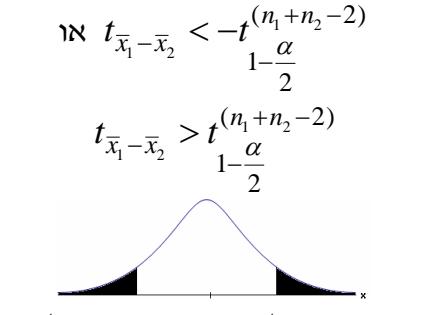
**תשובות סופיות**

- 1) נדחה  $H_0$ .
- 2) לא נדחה את  $H_0$ .
- 3) א. 0  
ב. נדחה את  $H_0$ .
- 4) א. נדחה את  $H_0$ , אם במדגם הגברים יהיו גבוהים בממוצע מהנשים ביוטרמו-10.72 ס"מ.  
ב. 0.6331

## בדיקות השערות על הפרש תוחלות במדגמים בלתי תלויים

---

### כששונוויות האוכלוסייה לא ידועות ומניחים שהן שווות – רקע

$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	השערת האפס: השערת אלטרנטיבית:
$H_1 \quad \mu_1 - \mu_2 > c$	$H_1 \quad \mu_1 - \mu_2 < c$	1. מדגמים בלתי תלויים 2. $\sigma_1, \sigma_2$ לא ידועות אך שווות 3. המשתנים בכל אוכלוסייה מתפלגים נורמלית	
$t_{\bar{x}_1 - \bar{x}_2} > t_{1-\alpha}^{(n_1+n_2-2)}$ 	$t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\alpha}^{(n_1+n_2-2)}$ 	$t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$ או $t_{\bar{x}_1 - \bar{x}_2} > t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$ 	אזור הדחיה של $H_0$

**סטטיסטי המבחן:**

$$t_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

**השונות המשוקלلت:**

$$S_p^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$$

**חלופה אחרת לכל הכרעה:**

נדחה $H_0$ אם מתקיים :	
$\bar{x}_1 - \bar{x}_2 < c - t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$ או $\bar{x}_1 - \bar{x}_2 < c - t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$
$\bar{x}_1 - \bar{x}_2 > c + t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$	

**דוגמה (פתרון בהקלטה) :**

חברה המייצרת מוצרי בנייה טוענת שפיתחה סגסוגת (תערובת מתכות) שטמפרטורת ההתחכה שלה גבוהה משמעותית מטמפרטורת ההתחכה של הסגסוגת לבנייה שימושים בה כיום לבניית בניינים. לצורך בדיקת טענתה המחקר נדגמו 10 יחידות של מתכוות מהסוג היין ו-12 יחידות של מתכוות מהסוג החדש. להלן תוצאות המדגם:

טמפרטורת ההתחכה הממוצעת במתכת הישנה 1170 מעלות עם אומד חסר הטיה לשונות  $S^2 = 200$ .

טמפרטורת ההתחכה הממוצעת במתכת החדשה 1317 מעלות עם אומד חסר הטיה לשונות  $S^2 = 260$ .  
 נניח לצורך פתרון שטמפרטורת ההתחכה מתפלגת נורמללית עם אותה שונות במתכוות השונות. בדקו ברמת מובהקות של 5%.

## שאלות

**1)** להלן נתונים של שטחי דירות מtower דירות שנבנו בשנת 2012 ובשנת 2013 (במ"ר) :

120	94	90	130	95	112	120	<b>2012</b>
69	74	105	91	82	100		<b>2013</b>

בדקו שבסנת 2013 הייתה ירידה משמעותית בשטחי הדירות לעומת שנת 2012  
 עבור רמת מובהקות של 5%.  
 הניתנו שטחי הדירות בכל שנה מתפלגים נורמלית עם אותה שוננות.

**2)** נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו ל מבחון IQ. להלן תוצאות

הדגם :	המדינה	ישראל	ארה"ב
	גודל המדגם	15	15
	סכום הציונים	1560	1470
	סכום ריבועי הציונים	165,390	147,560

בדקו ברמת מובהקות של 5% האם קיים הבדל של נקודה בין ישראלים  
 לאמריקאים מבחינת ממוצע הציונים ב מבחון-h-IQ ל佗בת ישראל.  
 רשמו את כל ההנחות הדרושים לצורך פתרון התרגיל.

**3)** להלן תוצאות מבחן אורך חיים של נורות מסוג W60 ומסוג W100.

אורך החיים ממיד ב שניות.

הקבוצה	60W	100W
$\bar{x}$	1007	956
S	80	72
n	13	15

- א. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקיות בממוצע יותר מאשר נורות מסוג W100. רשמו את כל ההנחות הדרושים לפתרון.
- ב. עבור איזו רמת מובהקות ניתן לקבוע שנורות מסוג W60 דולקיות בממוצע יותר מאשר נורות מסוג 100?
- ג. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקיות יותר מאשר נורות מסוג 1000 שניות. רשמו את כל ההנחות הדרושים.

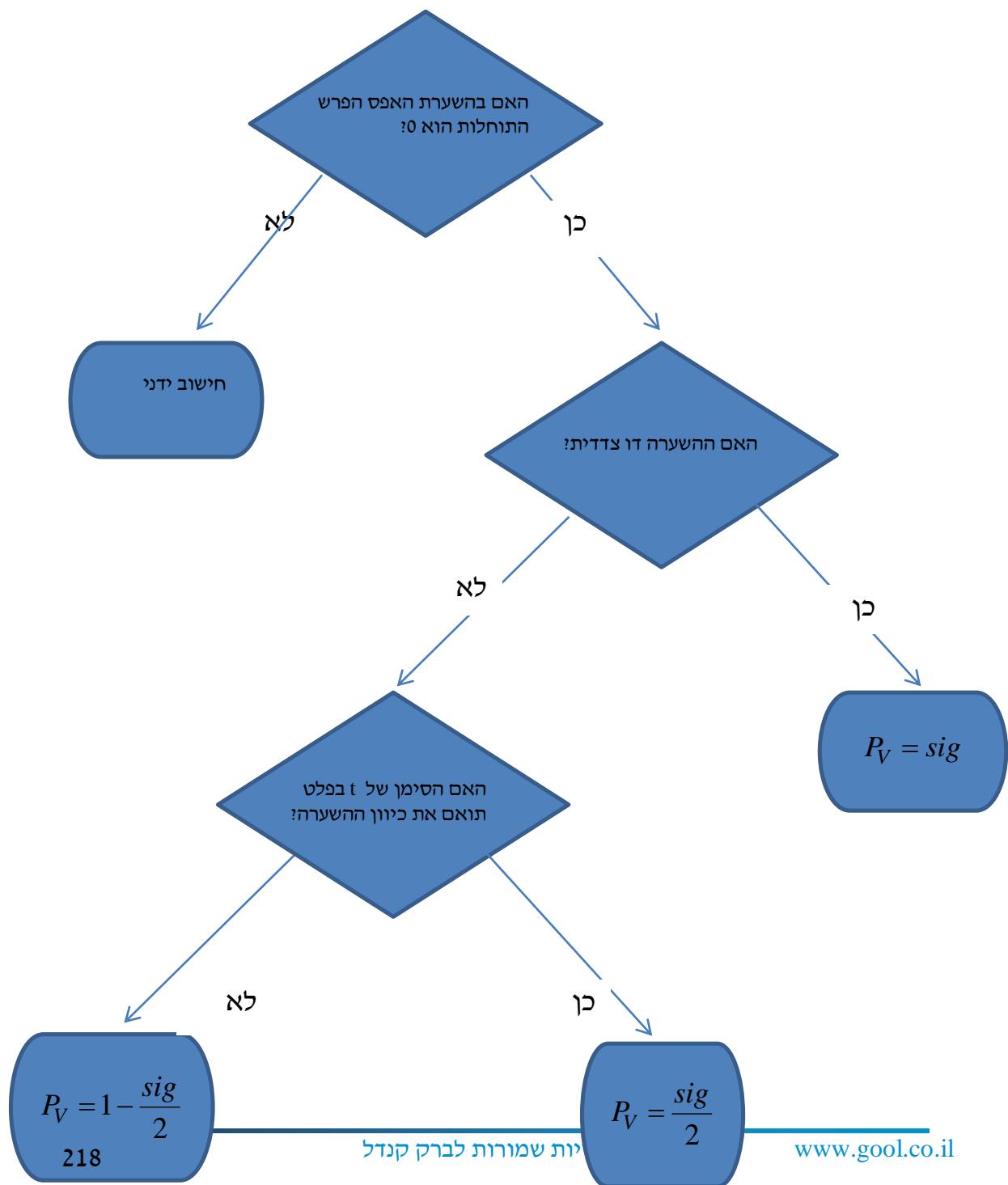
## תשובות סופיות

- 1) נדחה את  $H_0$ .
- 2) הנחות:  
1. סטיות התקן שוות.  
2. המשתנים מתפלגים נורמלית.
- נקבל את  $H_0$ .
- 3) א. נדחה את  $H_0$ .  
ב. רמת מובהקות של לפחות 5%.  
ג. לא נדחה את  $H_0$ .

## בדיקת השערות על הפרש תוחלות במדגים בלתי תלויים

### ניתוח פלטיטם – רקע

MOVBEKKOTHT HATOCZAH UL SMKHPFLT:



**דוגמה (פתרון בהקלטה) :**

בסקר שנערך בארה"ב בשנת 1993 נשאלו נסקרים שניים אזוריים שונים במדינה על מס' האחים והאחיות שלהם. להלן הפלט שהתקבל:

**Group Statistics**

	Region of the United States	N	Mean	Std. Deviation	Std. Error Mean
Number of Brothers and Sisters	North East	676	3.76	2.939	.113
	South East	410	4.05	2.993	.148

**Independent Samples Test**

	Levene's Test for Equality of Variances		t-test for Equality of Means						95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference			
							Lower	Upper		
Number of Brothers and Sisters	.173	.677	-1.583	1084	.114	-.293	.185	-.657	.070	
			-1.576	850.945	.115	-.293	.186	-.658	.072	

- מהו המבחן הסטטיסטי שנעשה כאן?
- בדוק ברמת מובהקות של 5% האם קיים שוויזון שוניות בין שני האזורים?
- בדוק האם קיים הבדל בין-"South East" ל-"North East" ברמת מובהקות של 5% מבחן מס' האחים והאחיות המוצע.
- מתי מובהקות התוצאה לבדיקת הטענה שההפרש המוצע בין-"South East" לבין-"North East" חיובי?

**שאלות**

- 1)** להלן פلت מתוכנת SPSS מתוך מחקר שבחן את רמת האופטימיות של גברים ונשים. רמת האופטימיות נמדזה בסולם ציונים של 1 עד 5.

**Group Statistics**

		GENDER	N	Mean	Std. Deviation	Std. Error Mean
optimizm	MALE		633	2.6053	.49781	.01979
	FEMALE		568	2.5503	.48483	.02034

**Independent Samples Test**

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
optimizm	Equal variances assumed	.383	.536	1.935	1199	.053	.05500	.02842	-.00076
				1.938	1190.977	.053	.05500	.02838	-.00068
	Equal variances not assumed								.11067

- א. האם ניתן להניח ששיעור האופטימיות של נשים וגברים שווה ברמת מובהקות של 5%?
- ב. ברמת מובהקות של 5% האם קיים הבדל בין הנשים לגברים ברמת האופטימיות הממוצעת שלהם?
- ג. מצא את הגבול העליון של רוח הסמק המשומן בסימן שאלה בפלט. דיקק עד 5 ספרות אחרי הקודעה.
- ד. בנה רוח סמק לתוחלת רמת האופטימיות של הגברים ברמת סמק של 95%.

2) פסיכולוגים טוענים שאנשים שניגשים לבחן אינטלקגנציה יותר מפעם אחת נוטים לקבל ציונים גבוהים יותר. להלן הפלט שהתקבל:

### Group Statistics

		N	Mean	Std. Deviation	Std. Error Mean
grade	A	9	96.8889	9.40006	3.13335
	B	11	108.4545	11.46616	3.45718

Independent Samples Test

	Levene's Test for Equality of Variances			t-test for Equality of Means							
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference			
								Lower	Upper		
grade	.206	.656	-2.428	18	.026	-11.56566	4.76333	-21.57304	-1.55828		
			-2.479	17.997	.023	-11.56566	4.66583	-21.36832	-1.76299		

### T-Test

מקרה :

A = נגשו פעם אחת.

B = נגשו יותר מפעם אחת.

א. רשמו את השערות המחקר והסבירו מהו המבחן המתאים כאן.

ב. כיצד הייתה משתנה התשובה לסעיף הקודם אם היה מדובר על אותם אנשים שציוונים נבדק פעמי אחדי המבחן הראשון שעשו ופעמי אחריה המבחן השני?

ג. האם ניתן לומר כי מידת הפיזור של ציוני האנשים הנבחנים בפעם הראשונה שונה ממידת הפיזור של ציוני האנשים אשר נבחנים בפעם השנייה. בדוק ברמת מובהקות של  $\alpha = 0.05$ .

ד. האם נכונה טענת הפסיכולוגים ברמת מובהקות של  $\alpha = 0.01$ .

(3) חלק מחקר בנושא הנישואין בישראל, אחד החוקרים העלה השערה שיש הבדל בממוצע גיל הנישואין (הראשונים), בין נשים הגרות בערים מרכזיות לבין נשים הגרות בערים מרוחקות מהמרכז.  
לשם כך נדגו 50 כלות מכל אחת משתי ערים עיר א'-מרכזית ועיר ב'-מרוחקת ונרשם גילן. תוצאות עיבוד הנתונים מופיעות בטבלאות שללון:

**T-Test****Group Statistics**

מקום המגורים	N	Mean	Std. Deviation	Std. Error Mean
גיל הנישואין	עיר א	24.8072	1.38978	.19654
	עיר ב	23.0131	1.62070	.22920

**Independent Samples Test**

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
גיל הנישואין	Equal variances assumed	.330	.567	5.942	98	.000	1.79415	.30193	1.19497 2.39332
	Equal variances not assumed			5.942	95.772	.000	1.79415	.30193	1.19480 2.39350

- א. מהו המבחן הסטטיסטי שנעשה כאן?
- ב. מצא רוח סמך ברמת סמך של 95% להפרש בין עיר א לעיר ב מבחינת גיל הנשים הממוצע בנישואין הראשונים.
- ג. האם ניתן לומר ברמת מובהקות של 1% שנשים בערים מרכזיות מתחנות בגיל מאוחר יותר מאשר נשים הגרות בערים מרוחקות?

(4) להלן פלט של תוכנת SPSS :

**T-Test**

	N	Mean	Std. Deviation	Std. Error Mean
X	26	36.3077	13.23259	2.59513
Y	24	46.4583	20.96369	4.27920

**Independent Samples Test**

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	4.446	.040	-2.164	???	.044	-10.15064	???	-20.03781	-.26347
Equal variances not assumed			-2.038	38.267	.048	???	5.00462	-20.27964	-.02164

- א. השלימו את סימני השאלה בטבלה.  
 ב. מהי מובוקות התוצאה לבדיקה הטענה שקיים הבדל בין השונות של  $X$  לעומת  $Y$ ?  
 ג. מהי מובוקות התוצאה לבדיקה הטענה שהתוחלת של  $X$  גדולה מהתוחלת של  $Y$ ?  
 ד. מהי מובוקות התוצאה לבדיקה הטענה שהתוחלת של  $X$  קטנה מהתוחלת של  $Y$ ?

**תשובות סופיות**

- 1)** א. קיבל את  $H_0$  ונכרייע שיש שוויון שוניות.  
 ב. נקבע שלא קיים הבדל בין נשים לגברים מבחינת האופטימיות המומוצעת.  
 ג.  $2.5665 \leq \mu \leq 2.6441$ .  
 ד. 0.11076.
- 2)** א. מבחן T להפרש ממוצעים במדגים בלתי תלויים.  
 ב. מבחן T למדגים מזוגיים.  
 ג. קיבל את  $H_0$ , נקבע לקיום שויפוי שוניות.  
 ד. קיבל את  $H_0$ , לא קיבל את טענת הפסיכולוגים.
- 3)** א. מבחן T להשוואת תוחלת במדגים בלתי תלויים.  
 ב.  $1.19497 \leq \mu_1 - \mu_2 \leq 2.39332$  ג. כן.  
 ד. 0.022.
- 4)** א. 10.15, 4.69,-48  
 ב. 0.04  
 ג. 0.978

# bijustisjka le-refoat shinim

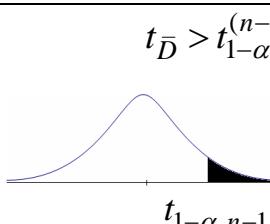
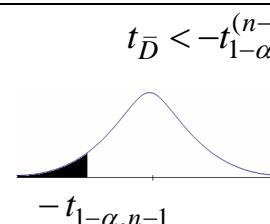
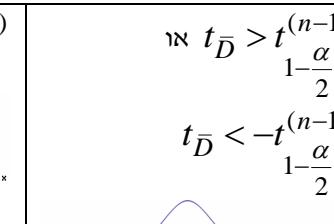
פרק 39 - בדיקת השערות לתוכלת ההפרש במדגמים מזוגים

## תוכן העניינים

225 .....	1. בדיקת השערות למדגמים מזוגים .....
229 .....	2. ניתוח פלטימ.....

## בדיקות השערות על תוחלת הרפישים במדגמים مزוגים (תלויים)

### בדיקות השערות למדגמים מזוגים – רקע

$H_0: \mu_D = C$	$H_0: \mu_D = C$	$H_0: \mu_D = C$	השערת האפס: השערת אלטרנטיבית:
$H_1: \mu_D > C$	$H_1: \mu_D < C$	$H_1: \mu_D \neq C$	תנאים:
		1. $\sigma_D$ אינה ידועה 2. או מדגם מספיק גדול $D \sim N$	כלל ההכרעה: ازור הדחיה של $H_0$
$t_{\bar{D}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha, n-1}$	$t_{\bar{D}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha, n-1}$	$t_{\bar{D}} > t_{\frac{1-\alpha}{2}, n-1}^{(n-1)}$ או $t_{\bar{D}} < -t_{\frac{1-\alpha}{2}, n-1}^{(n-1)}$  $-t_{\frac{1-\alpha}{2}, n-1} \quad t_{\frac{1-\alpha}{2}, n-1}$	$H_0$ - דוחים את ■ $H_0$ - דוחים את ■ $H_0$ - דוחים את ■
$\bar{D} > C + t_{1-\alpha}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} < C - t_{1-\alpha}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} > C + t_{\frac{1-\alpha}{2}}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$ ו $\bar{D} < C - t_{\frac{1-\alpha}{2}}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	חלופה לכלל הכרעה: נדחה $H_0$ אם מתתקיים:

$$S_D^2 = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n-1} = \frac{\sum_{i=1}^n D_i^2 - n\bar{D}^2}{n-1}, \quad t_{\bar{D}} = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}$$

**סטטיסטי המבחן:**

דוגמה (פתרו בהקלטה):

חברה שיווקית מעוניינת לבדוק את טענת רשות השיווק "מגה בעיר" הטענה שמחירים נמוכים מהמחירים מרשות השיווק "שופרסל". לצורך הבדיקה נבחרו באקראי 4 מוצרים שונים. המוצרים נבדקו בשתי הרשותות. להלן המוצרים:

ה מוצר / רשות	מגה בעיר	שופרסל
18	17	שמפו
57	48	gil כביסה
35	35	עוגת גבינה
10	12	לחם
47	49	קפה נמס
142	113	בקבוק יין
26	20	גבינה בולגרית

בහנחה והמחירים מתפלגים נורמלית, בדקו ברמת מובהקות של 5% את טענת רשות "מגה בעיר".

**שאלות**

- 1)** במטרה לבדוק האם קיימים הבדל בין חברת  $X$  לחברת  $Y$  מבחינת המחיר לשיחות בין-יל. נדגמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דלקת שיחה. להלן התוצאות:

חברה/ מדינה	ארה"ב	קנדה	הולנד	פולין	מצרים	סין	יפן
$X$	1.5	2.1	2.2	3	3.5	3.2	4.2
$Y$	1.4	2	1.9	3.1	3.2	3.2	4.2

בנהנה והמקרים מתפלגים נורמלית בכל חברת, בדקו ברמת מובהקות של 5% האם קיימים הבדל בין החברות מבחינת המחיר במומוץ?

- 2)** מכון המRAIN לפסיכוןטרא טוען שהוא מעלה את ממוצע הציונים ביוטר מ-30 נקודות. 8 נבחנים נבדקו לפני ואחרי שהם למדו במכון. להלן התוצאות שהתקבלו:

אחרי	לפני	590	500	390	670	640	420	470	506
580	520	510	680	610	430	540	570	570	506

מה מסקנתכם ברמת מובהקות 5%? הניחו שציוני פסיכוןטרא מתפלגים נורמלית.

- 3)** נדגמו 5 סטודנטים שישימנו את הקורס סטטיסטיקה ב'. להלן הציונים שלהם בסMASTER א' ו- ב':

סטטיסטיקה ב'	סטטיסטיקה א'	82	75	90	68	74
סטטיסטיקה ב'	סטטיסטיקה א'	100	76	87	84	80

פורסם שתלמידים שמשיכים את סטטיסטיקה ב' מושפרים ממוצע את הציונים ב-5 נקודות לעומת סטטיסטיקה א'. הניחו שהציונים מתפלגים נורמלית.

- א. מהי מובהקות התוצאה לבדיקת הטענה שהשיפור הוא יותר מ 5 נקודות?  
 ב. על סמך הסעיף הקודם, מהי רמת המובהקות המינימלית להכרעה שהשיפור הוא יותר מ- 5 נקודות?  
 ג. לאור זאת, מה המסקנה ברמת מובהקות של 10%?

- 4)** לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני לא היפנוזה. לאחר מכן נערך לשניהם מבחון באנגלית. נניח שציוני המבחן מתפלגים נורמלית ללא ידיעת השונות האמיתית. המבחן שיש לבצע כאן הוא:

- א. מבחן  $Z$  למדגם יחיד.  
 ב. מבחן  $T$  למדגם יחיד.  
 ג. מבחן  $T$  למדגמים בלתי תלויים.  
 ד. מבחן  $T$  למדגמים מזוגיים.

5) בتحقנת טיפת חלב מסויימת יש שני מכשירי שקילה. על מנת להשוות בין שני המשקלים נדגמו 4 תינוקות. כל תינוק בן חודשיים נשקל בכל אחד מהמשקלים.

להלן תוצאות השקליה (בק"ג) :

	משקל במכשיר 1	4.5	9.6	0.7	2.5
	משקל במכשיר 2	3.5	6.9	1.7	0.5

נניח שהמשקלים מתפלגים נורמלית, המבחן שיש לבצע כאן הוא :

- א. מבחן Z למדגם יחיד.
- ב. מבחן T למדגם יחיד.
- ג. מבחן T למדגמים בלתי תלויים.
- ד. מבחן T למדגמים מזוגים.

6) כדי להשוות בין שני אצנים נדגמו 5 תוצאות מריצת 100 מטר של כל אצן. זמני הריצה נרשמו ויש להניח שמתפלגים נורמלית. המטרה להשוות בין האצנים.

המבחן שיש לבצע כאן הוא :

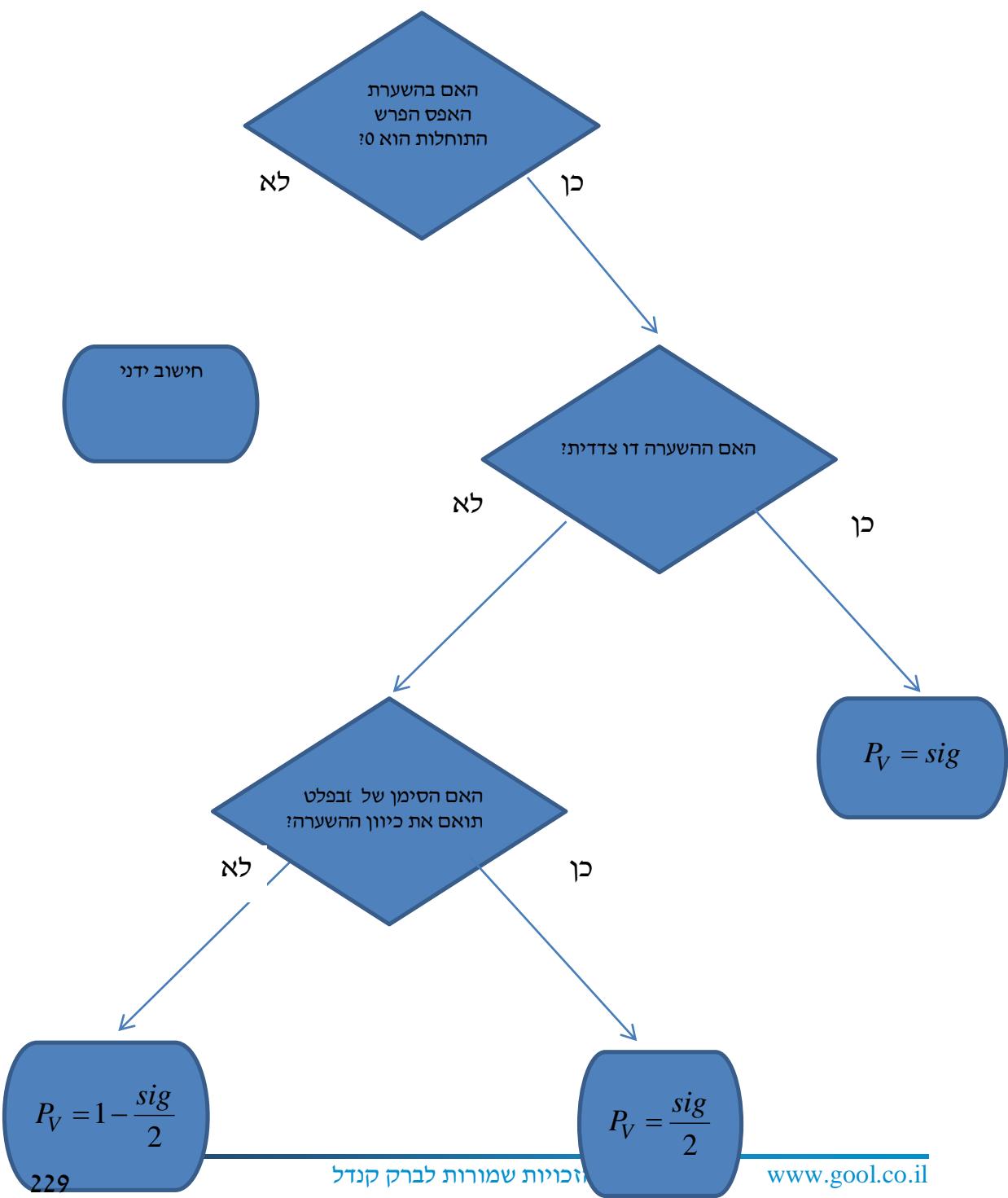
- א. מבחן Z למדגם יחיד.
- ב. מבחן T למדגם יחיד.
- ג. מבחן T למדגמים בלתי תלויים.
- ד. מבחן T למדגמים מזוגים.

## תשובות סופיות

- 1) לא נדחה  $H_0$ .
- 2) לא נדחה  $H_0$ .
- 3) א. לא נדחה  $H_0$ .      ב.  $0.5 \leq pv \leq 0.5$       ג. לא נדחה  $H_0$ .
- 4) ד'.
- 5) ד'.
- 6) ג'.

## בדיקות השערות על תוחלת הרפישים במדגמים مزוגים (תלוים)

### מדגמים מזוגים – ניתוח פלטיט – רקע



**דוגמה (פתרון בהקלטה) :**

כדי לבדוק את ההשפעה של קורס לגמilia מעישון נלקח מדגם מקרי של 5 נבדקים. עבור כל אחד מהם נמדדה צריכת הסיגריות היומיית לפני הקורס וחודשיים אחריו. הניבו שצריכת הסיגריות מתפלגת נורמלית. להלן התוצאות:

5	4	3	2	1	נבדק
30	28	25	22	40	לפני
12	10	13	24	30	אחרי

**Paired Samples Statistics**

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 BEFORE	29.0000	5	6.85565	3.06594
AFTER	17.8000	5	8.72926	3.90384

**Paired Samples Test**

	Paired Differences					t	df	Sig. (2-tailed)			
	Mean	Std. Deviation	Std. Error Mean	90% Confidence Interval of the Difference							
				Lower	Upper						
Pair 1 BEFORE - AFTER	11.20000	8.19756	3.66606	3.38452	19.01548	3.055	4	.038			

בדקו ברמת מובהקות של 5% האם הקורס יעיל.

**שאלות**

**1)** בסקר שנערך באראה"יב בשנת 1993 נשאלו נסקרים על השכלה הוריהם, להלן הפלט שהתקבל:

Paired Samples Test											
	Paired Differences					t	df	Sig. (2-tailed)			
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference							
				Lower	Upper						
Pair 1	Highest Year School Completed, Father - Highest Year School Completed, Mother	-.007	3.115	.100	-.203	.189	-.072	.973	.943		

- א. תננו אומדן להפרש הממוצעים.
- ב. תננו אומדן לטיעות התקן של הפרש הממוצעים.
- ג. האם קיים הבדל מובהק בין השכלה האבות להשכלה האימהות ברמת מובהקות של 5%?

**2)** בתחרויות קפיצה למים שופטים באופן קבוע שופט איטלקי ושופט דרום קורייני. להלן פלט המנתה את הציונים לשופטים אלה נתנו בתחרויות השונות:

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Italy South Korea	???	300	.86742	.05008
		8.9183	???	.81992	.04734

Paired Samples Test											
	Paired Differences					t	df	Sig. (2-tailed)			
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference							
				Lower	Upper						
Pair 1	Italy - South Korea	-.42233	.36153	.02087	-.46341	-.38126	-20.234	???	???		

- א. השלימו את החלקים החסרים בפלט (מסומנים בסימני שאלה).
- ב. בדקו את הטענה שהשופט הדרום קורייני נותן בממוצע 0.2 נקודות יותר מאשר השופט האיטלקי ברמת מובהקות של 5%.
- ג. מהו רוחח הסמך ברמת סמך של 95% ל佗חת פער הציונים בין השופטים?
- ד. בנו את הרוחח כת ערך ברמת סמך של 98% ל佗חת פער הציונים בין השופטים.

**3)** בדקו את ציוניהם של 44 נבדקים אקראים ב מבחן הפסיכומטרי. פעם אחת לפני הכנה (Before) ופעם אחת אחרי הכנה (After).

Paired Samples Test											
	Paired Differences					t	df	Sig. (2-tailed)			
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference							
				Lower	Upper						
Pair 1 Before - After	-7.45455	19.28303	2.90703	-13.31712	-1.59197	-2.564	43	.014			

- א. רשמו מהו המבחן הסטטיסטי ונשח את ההשערות אליהם מתייחס הפלט.
- ב. בדקו את ההשערה שממוצע ציונים משתפרים לאחר ההכנה ברמת מובהקות של .5%.
- ג. בדקו את ההשערה שממוצע ציונים משתפרים לאחר ההכנה ביותר מ-5% נקודות ברמת מובהקות של .5%.
- ד. מצאו רוח סמך ל佗חת שיפור ממוצע הציונים לאחר ההכנה ברמת ביטחון של .95%.

(4) להלן פלט של תכנת SPSS :

**T-Test****Paired Samples Statistics**

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1	x	54.0000	6	5.86515
	y	46.5000	6	10.72847
				4.37988

**Paired Samples Test**

	Paired Differences					t	df	Sig. (2-tailed)			
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference							
				Lower	Upper						
Pair 1	x - y	7.50000	??	4.72405	-4.64356	19.64356	??	5	.173		

- א. מלא את החלקים החסרים בטבלה.
- ב. מהי רמת המובהקות המינימלית לקבלת הטענה שיש הבדל בין  $X$  ל- $Y$  בממוצע?
- ג. האם התשובה לסעיף הקודם הייתה משתנה, ואם כן גדלתה או קטנה, אם היינו מושיעים עוד תצפית שההפרש בין  $X$  ל- $Y$  הוא 0.
- ד. מהי מובהקות התוצאה לבדיקת הטענה ש  $X$  גדול מ-  $Y$  בממוצע?
- ה. מהי מובהקות התוצאה לבדיקת הטענה ש  $X$  קטן מ-  $Y$  בממוצע?
- ו. בנו רוח סמן ל佗ולת של  $X$  ברמת סמן של 90%.

**תשובות סופיות**

- (1) א.  $-0.007$   
 ב. 0.1  
 ג. אין הבדל מובהק.
- (2) א.  $d.f = 299$   
 ב.  $n = 300$   
 ג.  $\bar{X} = 8.496$ . Sig = 0
- (3) א. ראה וידאו.  
 ב. נדחה את  $H_0$ .  
 ג. לא נדחה את  $H_0$ .
- (4) א. 11.5715 ,1.5876  
 ב. 0.173  
 ג. יגדל.  
 ד.  $0.0865 < \mu < 49.18$

## bijustisjka le-refoat shinim

פרק 40 - הקשר בין רוח סמרק לבדיקה השערות להפרש תוחלות

תוכן העניינים

1. הקשר בין רוח סמרק לבדיקה השערות להפרש תוחלות.....  
235 .....

## הקשר בין רוח סמך לבדיקה השערות על הפרש תוחלות

### רקע

ניתן לבצע בדיקת השערות דו צדדיות ברמת מובהקות  $\alpha$  על  $\mu_1 - \mu_2$  :

$$H_0: \mu_1 - \mu_2 = C, \quad H_1: \mu_1 - \mu_2 \neq C$$

על ידי בניית רוח סמך ברמת סמך של  $\alpha - 1$  על  $\mu_1 - \mu_2$  :

אם  $C$  נופל ברווח  $\leftarrow$  קיבל את  $H_0$ .

אם  $C$  לא נופל ברווח  $\leftarrow$  נדחה את  $H_0$ .

### דוגמה (פתרו החלטה) :

חוקר ביצע בדיקת השערות לתחולת ההפרש במדגים מזוווג.

להלן השערותיו :  $H_0: \mu_D = 80, \quad H_1: \mu_D \neq 80, \quad \alpha = 5\%$

החוקר בנה רוח סמך ברמה של  $90\% < \mu_D < 78$ .

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

## שאלות

**1)** נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'. להלן ציוניהם בסMASTER א' ו- ב':

סMASTER A	סMASTER B
80	74
84	68
87	90
76	75
100	82

א. בנו רוח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סMASTER A' לבין סMASTER B'.

ב. פורסם שתלמידים ממשיכים את סMASTER B' משפרים ממוצע את הציונים ב-5 נק' לעומת סMASTER A'. האם יש אמת בפרסום?

**2)** הוחלט להשוות הציונים אצל מרצה X ואצל מרצה Y. נבחרו באקראי 6 סטודנטים, 3 סטודנטים של מרצה X ו-3 סטודנטים של מרצה Y, עבורם התקבלו הציונים הבאים:

מרצה X	82	90	68
מרצה Y	68	81	68

א. חשבו רוח סמך ברמת סמך 90% להפרש בין התוחלות של הציונים אצל שני המרצים.

ב. האם ברמת מובהקות של 10% נכיר שיש הבדל בין תוחלות הציונים אצל שני המרצים?

### שאלות רב-ברירה:

**3)** סטטיסטיאי נתקש לאמוד את הפרש הממוצעים של שני טיפולים לפי שני מוגדים מקרים בלתי תלויים.

הוא חישב רוח סמך להפרש ברמת סמך 0.98, וקיים את הרווח  $\mu_2 - \mu_1 < 2$ . אילו יתבקש החוקר לבדוק לפי אותן נתונות את השערות:

$H_0: \mu_1 - \mu_2 = 0$ ;  $H_1: \mu_1 - \mu_2 > 0$ , מסקנתו תהיה:

א. לדחות את השערת האפס.

ב. לא לדחות את השערת האפס.

ג. שלא ניתן לדעת את המסקנה עבור רמת מובהקות 0.05.

ד. שלא נתנות בשאלת סטיות התקן של האוכלוסיות, ולכן לא ניתן להסיק דבר.

- (4) במטרה לבדוק האם קיימים הבדל בין קווי זהב לבזק מבחינות ממוצע המחרירים לשיחות ביןיל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקט שיחה. בהנחה והמחקרים מתפלים נורמלית בנו רוח סמך ממוצע ההפרשנים וקיים :  $\mu_D < 0.2145 - 0.0293$ , רוח הסמך הוא ברמת סמך של 95% .  
לכן מסקנת המחקר היא :
- א. ברמת מובהקות של 5% לא נוכל לקבוע שקיימים הבדל בין החברות.
  - ב. ברמת מובהקות של 5% נקבע שקיימים הבדל מובהק בין החברות.
  - ג. לא ניתן לדעת מה המסקנה ברמת מובהקות של 5% כיון שלא נאמר מה ההגדרה של  $D$  .

### תשובות סופיות

(1) א.  $-3.8 \leq \mu_D \leq 19$   
ב. נכרייע שיש אמת בפרסום.

(2) א.  $-8.5 \leq \mu_X - \mu_Y \leq 26.5$   
ב. נכרייע שאין הבדל.

(3) ג'.  
א. (4)

# bijustisjka le-refoat shinim

## פרק 41 - מבחני חיות רפואי

### תוכן העניינים

238	1. מבחן לאי תלות
243	2. מקדם המתאים של קרמר
245	3. ניתוח פלטינם בבדיקה לאי תלות

## מבחן חי בריבוע לאי תלות בין משתנים – רקע

מבחן לאי תלות מטרתו לבדוק האם קיים קשר בין שני משתנים. שני המשתנים שנבדקים צריכים להיות מחולקים למספר קטגוריות.

**מבנה המבחן:**

**השערות:**

אין תלות בין המשתנים  $H_0$ .

יש תלות בין המשתנים  $H_1$ .

**כלל הכרעה:**

הערך הקритי נקבע על סמך התפלגות חי בריבוע. התפלגות זו היא אסימטרית חיובית ותלויה בדרגות החופש  $(r-1)(c-1)$ .  
 כאשר:  $r$  - מספר הקטגוריות של המשתנה שبشורות.  
 $c$  - מספר הקטגוריות של המשתנה שבעמודות.

הערך הקритי הוא:  $\chi^2_{1-\alpha,(r-1)(c-1)}$ , קלומר האחזוון ה-  $1 - \alpha$  בתפלגות חי בריבוע שדרגות החופש הן  $(r-1)(c-1)$ . אם  $\chi^2 > \chi^2_{1-\alpha,(r-1)(c-1)}$  או דוחים את השערת האפס.

$$\text{סטטיסטי המבחן: } \chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

כאשר:



$O_i$  - השכיחות נצפית במדגם בתא  $i$ .

$E_i$  - שכיחות צפואה במדגם בתא  $i$  בהנחה השערת האפס.

$$E_i = \frac{f(x) \cdot f(y)}{n}$$

**הערה:**

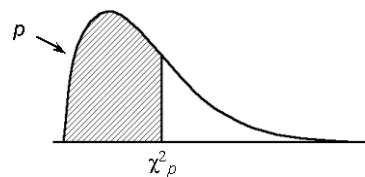
תנאי כדי לבצע את המבחן הוא  $5 \geq E_i$  לכל  $i$ . במידה ותנאי זה לא מתקיים יש אפשרות לאחד קטגוריות סמכות עד שהתנאי יתקיים.

תנאי חלופי: אין  $E$  קטן מ-1 וגם אין יותר מ 20% מהתאים  $E$  קטנים מ-5.

**דוגמה (הפתרון בהקלטה) :**

האם יש תלות בין המגדר לבין דעה מסוימת?  
יש לבדוק ברמת מובהקות של 5% על סמך תוצאות הסקר :

המגדר / דעתה	סה"כ	נמנע	נגד	بعد	גברים
נשים	20	60	20	50	10
סה"כ					

**טבלת התפלגות חי-בריבוע – ערכי החלוקה  $\chi^2_p$** 


df	$p$												
	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995
1	0.04393	0.03157	0.03982	0.0393	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

## שאלות

1) נבדקה ה תלות בין גודל הארגון לבין שביעות הרצון של העובדים.

להלן התוצאות:

סה"כ	גבואה	ביןונית	נמוכה	שביעות רצון	גודל המפעול
600	215	203	182		גודל
400	136	110	154		קטן
1000	351	313	336		סה"כ

מה המסקנה ברמת מובהקות של 2.5%?

2) מפעל עובד בשלוש משמרות. להלן מספר המוצרים הפוגמים והתקינים בכל

אחד מן המשמרות לפי דגם שנעשה:

פוגמים	לילה	יום	ערב
50	60	70	
600	700	800	

האם יש הבדל בין שיעורי הפוגמים במשמרות השונות?

הסיקו עבור רמת מובהקות של  $\alpha = 0.05$ .

3) נדגמו 50 מוצרים ממפעל מסוים מתוך 30 מוצרים שיוצרו ביום 17 נבחרו

לייצוא מתוך המוצרים שיוצרו בלילה 10 נבחרו לייצוא.

האם יש קשר בין להיות מוצר לייצוא למועד שבו הוא יוצר?

בדקו ברמת בטחון של 95%.

4) במטרה לבדוק האם השתנו דפוסי החכבה למפלגות השונות בשבוע שעבר

לשבוע נלקחו שני סקרים אחד מהשבוע שעבר והאחר מהשבוע.

להלן דפוסי החכבה שהתקבלו בסקרים אלה.

א. מהי רמת המובהקות המינימלית עבורה ניתן להחליט שהשתנו דפוסי החכבה בשבוע שעבר לשבוע באופן מובהק?

ב. כיצד הייתה התשובה לסעיף א' משתנה אם כל השכיחויות בטבלה של תוצאות הדגם היו מוכפלות פי 2?

ג. בנו רוח סמך לשיעור המצביעים למפלגה א' בשבוע ברמת סמך של 95%.

סה"כ	מפלגה א'	מפלגה ב'	מפלגות אחרות	שבוע שעבר
550	253		143	השבוע
1050		314	243	סה"כ

- 5) בוחנות בגדים A בדקו את התפלגות הצבעים של הבגדים הנמכרים ביום מסוים. כמו כן בדקו את התפלגות הצבעים בוחנות שכנה B :

מספר פריטים / צבע	כחול	אדום	לבן	שחור	אדום
חנות A	50	15	20	15	
חנות B	20	10	20	60	

א. בדקו ברמת מובהקות של 5% האם התפלגות הצבעים בוחנות A היא ביחס של 1:1:1:3 לטובת הכחול.

ב. בדוק ברמת מובהקות של 2.5% האם קיים הבדל בין החניות מבחינת התפלגות הצבעים של הפריטים הנמכרים.

- 6) סטודנט קיבל בבדיקה השערות ערך  $\chi^2$  (chi-square) השוו לאפס. הסטודנט הסיק כי לא קיימת תלות בין שני המשתנים שבדק, בכל רמת מובהקות. נכון / לא נכון? נמקו.

- 7) להלן טבלה O של שני משתנים שהתקבל במדגם כלשהו :

$f(x)$	$Y_4$	$Y_3$	$Y_2$	$Y_1$	
200					$X_1$
200					$X_2$
	160	120	60	60	$f(y)$

מה צרכות להיות השכיחויות בתוך הטבלה כדי שМОובהקות התוצאה (PV) תהיה 100%?

### תשובות סופיות

- 1) נסיק שיש קשר בין גודל הארגון לשביעות הרצון של העובדים.
- 2) נסיק שאין הבדל מובהק בין שיעור הפגומים במשמרות השונות.
- 3) נסיק שאין קשר בין היות מוצר ליצוא למועד שבו הוא יוצר.
- 4) א. 10%      ב. קטן      ג. (0.223,0.297)
- 5) א. נסיק שההתפלגות הצבעים בוחנות היא כמו שמצוין.  
ב. נסיק שיש הבדל בין החניות מבחינת התפלגות הצבעים.
- 6) נכון
- 7) להלן טבלה :

$f(x)$	$Y_4$	$Y_3$	$Y_2$	$Y_1$	
200	80	60	30	30	$X_1$
200	-8	60	30	30	$X_2$
400	160	120	60	60	$f(y)$

## מדד קשור-מדד הקשר של קרמר – רקע

מתי משתמשים במדד זהה? – כאשר אחד המשתנים הוא מסולםשמי 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 רוצחים לבדוק את הקשר בין מין לדעה בנושא מסוים, שאלו 100 גברים ו- 100 נשים האם הם בעד/נגד/נמנעים באיזשהו נושא. להלן טבלת השכיחיות המשותפת שהתקבלה.

בחזור של קרמר הטבלה נקראת טבלת O

(observed)

X - מין (גבר/אישה) – סולםשמי.

Y - דעה (בעד/נמנע/נגד) – סולםשמי/סדר.

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

שלבים בחישוב :  $r_c$

שלב א' :

בנייה את טבלת E (Expected)

$$E_i = \frac{f(x) \cdot f(y)}{n} \text{ וזו כל}$$

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

שלב ב' :

$$\text{נחשב } \chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

שלב ג' :

$$\text{נחשב : } r_c = \sqrt{\frac{1}{n(L-1)} \chi^2}$$

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

## שאלות

1) להלן תוצאות מחקר שבודק את הקשר בין מין להשכלה. לגבי כל נחקר נבדק

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

המין שלו והשכלתו.  
להלן התוצאות:

האם קיימים קשר בין מין להשכלה? נמקו!

2) נלקחו 200 אנשים שמתוכם 60 הוכיחו שהם עוסקים בפעילויות גופנית סדירה.

מתוך אלו שעוסקים בפעילויות גופנית סדירה 50 נמצאו במצב בריאותי תקין.

מתוך אלו שלא עוסקים בפעילויות גופנית סדירה 90 נמצאו במצב בריאותי תקין.

א. בנו טבלה שכיהות משותפת לנתחים שהוצעו בשאלת.

ב. האם קיימים קשר בין פעילות גופנית למצב בריאותי?  
חשבו לפי מzelf הקשר של קרמר.

## תשובות סופיות

1) קיימים קשר בעוצמה בינונית בין המין להשכלה מקדם המתאים של קרמר הוא 0.595.

2) א. להלן טבלה:  
ב. מzelf קרמר 0.19 מעיד על קשר בעוצמה נמוכה.

$f(x)$	לא נכון	נכון	$x/y$
60	10	50	כן
140	50	90	לא
200	60	140	$f(y)$

## פלטיטים על מבחן לאי תלות – רקע

מבחן לאי תלות מטרתו לבדוק האם קיים קשר בין שני משתנים. שני המשתנים שנבדקים צריכים להיות מחולקים למספר קטגוריות.

**מבנה המבחן:**

**השערות:**

אין תלות בין המשתנים :  $H_0$ .

יש תלות בין המשתנים :  $H_1$ .

**דרגות חופש :**  $d.f = (r-1)(c-1)$ .

כאשר :  $r$  - מספר הקטגוריות של המשתנה שبشורות.

$c$  - ספר הקטגוריות של המשתנה שבעמודות.

$$\text{סטטיסטי המבחן: } \chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$O_i$  - שכיחות נצפית במדגם בתא  $i$ .

$E_i$  - שכיחות צפואה במדגם בתא  $i$  בהנחה השערת האפס.

$$E_i = \frac{f(x) \cdot f(y)}{n}$$

**הערה :**

תנאי כדי לבצע את המבחן הוא  $E_i \geq 5$  לכל  $i$ . במידה ותנאי זה לא מתקיים יש

אפשרויות לאחד קטגוריות סמכות עד שהתנאי יתקיים.

**תנאי חלופי:** אין  $E$  קטן מ-1 וגם אין יותר מ 20% מהתאים  $E$  קטן מ-5.

**דוגמה (פתרון בהקלטה) :**

- במחקר רצוי לבדוק את הקשר בין צבע שיער לבין צבע עיניים של אנשים. הפלטים שהתקבלו מצורפים.
- א. מהו המבחן הסטטיסטי המתאים?
  - ב. כמה קטגוריות יש לכל משתנה?
  - ג. רשמו את השערות המחקר.
  - ד. מה מספר דרגות החופש?
  - ה. כמה אנשים במדגם נמצאו עם שיער חום?
  - ו. כמה אנשים הייתה מזכה במדגם שהיה להם שיער חום ועיניים יroxות בהנחה ואין קשר בין צבע שיער לצבע עיניים?
  - ז. מתוך הבלונדיינים מה אחוז בעלי עיניים כחולות במדגם?
  - ח. מתוך בעלי עיניים יroxות מה אחוז הבלונדיינים במדגם?
  - ט. מה ערכו של סטטיסטי המבחן ומהי מובاهקות התוצאות?
  - י. מה מסקנת המחקר?  $\alpha = 5\%$

להלן הפלטים שהתקבלו:

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
hair_color * eye_color	78	100.0%	0	0.0%	78	100.0%

### hair\_color \* eye\_color Crosstabulation

		eye_color			Total
		brown	green	Blue	
hair_color	black	Count	13	7	27
		Expected Count	10.7	8.3	27.0
		% within hair_color	48.1%	25.9%	25.9%
	brown	% within eye_color	41.9%	29.2%	30.4%
		Count	12	12	30
		Expected Count	11.9	9.2	30.0
	blond	% within hair_color	40.0%	40.0%	20.0%
		% within eye_color	38.7%	50.0%	26.1%
		Count	6	5	21
Total	blond	Expected Count	8.3	6.5	21.0
		% within hair_color	28.6%	23.8%	47.6%
		% within eye_color	19.4%	20.8%	43.5%
	Total	Count	31	24	78
		Expected Count	31.0	24.0	78.0
		% within hair_color	39.7%	30.8%	29.5%
		% within eye_color	100.0%	100.0%	100.0%

### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5.880 <sup>a</sup>	4	.208
Likelihood Ratio	5.641	4	.228
Linear-by-Linear Association	2.682	1	.101
N of Valid Cases	78		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.19.

## שאלות

- 1) בסקר שנעשה על ידי משרד ראש הממשלה נדגו 60 אזרחים. כל אזרח נשאל על מגדרו והאם הוא بعد הקמת מדינה פלסטינית.
- א. מה ההשערות הנבדקות ומהו סטטיסטי המבחן?
  - ב. אם סטטיסטי המבחן היה גדול כיצד הדבר היה משפייע על SIG שבפלט.
  - ג. האם קיים קשר בין מגדר ודעה ברמות מובהקות של 5%?
  - ד. מהו האומדן לאחוז התומכים במדינה פלסטינית מתוך הגברים?
  - ה. Aiיזה אחוז מהנשאלים שהיו بعد מדינה פלשתינית הם גברים?

להלן הפלטים:

### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.973 <sup>a</sup>	2	.373
Likelihood Ratio	1.987	2	.370
Linear-by-Linear Association	1.882	1	.170
N of Valid Cases	60		

a. 0 cells (.0%) have expected count less than 5.

b. The minimum expected count is 7.25.

### gender \* opinion Crosstabulation

		opinion			Total	
		yes	now	no opinion		
gender	Male	Count	10	10	9	29
		Expected Count	12.6	9.2	7.3	29.0
		% within gender	34.5%	34.5%	31.0%	100.0%
		% within opinion	38.5%	52.6%	60.0%	48.3%
		% of Total	16.7%	16.7%	15.0%	48.3%
	female	Count	16	9	6	31
		Expected Count	13.4	9.8	7.8	31.0
		% within gender	51.6%	29.0%	19.4%	100.0%
		% within opinion	61.5%	47.4%	40.0%	51.7%
		% of Total	26.7%	15.0%	10.0%	51.7%
	Total	Count	26	19	15	60
		Expected Count	26.0	19.0	15.0	60.0
		% within gender	43.3%	31.7%	25.0%	100.0%
		% within opinion	100.0%	100.0%	100.0%	100.0%
		% of Total	43.3%	31.7%	25.0%	100.0%

2) להלן פلت על סמך סקר שנעשה בקרב סטודנטים, בסקר נשאלו הסטודנטים על המוזיקה אותה הם מעדיפים וצורת הבילוי המועדף עליהם.

**Crosstab**

Count

		בילוי			Total
		קריאה	ספורט	睦עדיון	
מוזיקה	רוק	0	0	11	11
	פופ	1	6	8	15
קלאסי		5	6	9	20
Total		6	12	28	46

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	11.929 <sup>a</sup>	?	.018
N of Valid Cases	46		

a. 5 cells (55.6%) have expected count less than 5.

b. The minimum expected count is 1.43.

א. בין אלו משתנים נבדק הקשר? כמה קטגוריות לכל משתנה?

ב. האם התנאים של המודל מתקיים?

ג. מה מספר דרגות החופש בבדיקה הנ"ל?

ד. מה ההשערות של הבדיקה?

- (3) מחקר התענין לבדוק את הקשר בין רמת הכנסה של משפחה לבין צריכת עגבנייהות ארגניות. הפלטים מצורפים.
- השלימו את שלושת המספרים החסרים בטבלה (היכן שיש סימני שאלה).
  - מה ערכו של חי בריון הסטטיסטי.
  - תנו הערה למועדוקות התוצאות לבדוק הקשר בין רמת הכנסה של משפחה לבין צריכת עגבנייהות ארגניות.

### רמת\_הכנסה \* צרךן עגבנייהות Crosstabulation

		צרךן עגבנייהות		Total
		לא ארגני	ארוגני	
הרבה מתחת לממוצע	Count	17	42	59
הכנסה_רמתchin	% within	28.8%	?	100.0%
עצביות צרךןchin	% within	13.6%	33.6%	23.6%
מתחת לממוצע	Count	27	22	49
הכנסה_רמתchin	% within	55.1%	44.9%	100.0%
עצביות צרךןchin	% within	?	17.6%	19.6%
ממוצע	Count	31	29	60
הכנסה_רמתchin	% within	51.7%	48.3%	100.0%
עצביות צרךןchin	% within	24.8%	23.2%	24.0%
מעל הממוצע	Count	44	26	70
הכנסה_רמתchin	% within	62.9%	37.1%	100.0%
עצביות צרךןchin	% within	35.2%	20.8%	28.0%
הרבה מעל הממוצע	Count	?	6	12
הכנסה_רמתchin	% within	50.0%	50.0%	100.0%
עצביות צרךןchin	% within	4.8%	4.8%	4.8%
Total	Count	125	125	250

- (4) חוקר בדק את הקשר בין צבע השיער לבין צבע העיניים בעזרת מבחן Chi בריבוע בקרב 52 נבדקים. תוצאות המבחן מוצגות בטבלה. בנוסף ידוע כי סטטיסטי המבחן שהתקבל מעיבוד הנתונים הוא 8.08.
- מה תהיה מסקנת המבחן ברמת מובהקות של 1%?
  - מה ערכו של E עבור עיניים כחולות וצבע שיער כהה.
  - מה יהיה בקירוב ערכו של מקדם המתאים של קרמר?
  - מהי פרופורציה בעלי צבע שיער הבוחר מקרב בעלי העיניים הירוקות?

להלן הפלט:

#### צבע עיניים \* צבע שיער Crosstabulation

		צבע שיער		Total
		כהה	בהיר	
צבע עיניים	כחול	Count		
		% within		50.0% 100.0%
		% within		21.6% 30.8%
		% of Total		15.4% 30.8%
חום	חום	Count		
		% within		83.3% 100.0%
		% within		27.0% 23.1%
		% of Total		19.2% 23.1%
ירוק	ירוק	Count		
		% within		79.2% 100.0%
		% within		51.4% 46.2%
		% of Total		36.5% 46.2%
Total		Count		
		% within		71.2% 100.0%
		% within		100.0% 100.0%
		% of Total		71.2% 100.0%

- 5) במחקר מסוים רצוי לבדוק האם יש קשר בין המגדר להוצאה על לבוש במשך שנה. דגמו באופן מקרי גברים ונשים ובדקו את רמת ההוצאה שלהם על לבוש השנה האחרונות. חוקר א' בדק האם קיימים הבדלים תוחלתיים בהוצאה בין גברים לנשים. חוקר ב' קיבץ את ההוצאה לקטגוריות ובאופן זה בדק האם קיימים הבדלים בתפלגות ההוצאה בין גברים לנשים.
- הקטגוריות חולקו לשולש קבוצות הוצאה.
- איזה פלט מתאים לאיזה אחד מהחוקרים? נמקו.
  - מה מסקנתו של חוקר א'? בדקו ברמת מובהקות  $\alpha = 0.05$ . (רשמו השערות, נחו הנחות, ציינו כלל החלטה ותנו מסקנה במונחי המשתנים).
  - איזו טעות יכולה להיות במסקנתו של חוקר א'?  
נחו את הטעות במונחי השאלה.
  - מהי מסקנתו של חוקר ב'? בדקו ברמת מובהקות  $\alpha = 0.05$ . (רשמו השערות, נחו הנחות, ציינו כלל החלטה ורשמו מסקנה במונחי המשתנים).
  - איזו טעות יכולה להיות במסקנתו של חוקר ב'? נחו זאת במונחי השאלה.
  - כיצד ניתן ליישב את מסקנות שני החוקרים?

להלן פלט ראשון:

### T-Test

#### Group Statistics

gender	N	Mean	Std. Deviation	Std. Error Mean
female	40	2.9000	1.15025	.18187
male	40	2.6000	2.52982	.40000
dimension n1				

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means							95% Confidence Interval of the Difference			
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference						
								Lower	Upper				
expese	Equal variances assumed	16.805	.000	.683	.78	.497	.30000	.43941	-.57479	1.17479			
				.683	54.464	.498	.30000	.43941	-.58078	1.18078			

להלן פלט שני:

### Crosstabs

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
gender * category	80	100.0%	0	.0%	80	100.0%

#### gender \* category Crosstabulation

		category			Total	
		a	b	c		
gender	Female	Count	2	30	8	40
		Expected Count	11.0	21.0	8.0	40.0
		% within gender	5.0%	75.0%	20.0%	100.0%
		% within category	9.1%	71.4%	50.0%	50.0%
	Male	Count	20	12	8	40
		Expected Count	11.0	21.0	8.0	40.0
		% within gender	50.0%	30.0%	20.0%	100.0%
		% within category	90.9%	28.6%	50.0%	50.0%
	Total	Count	22	42	16	80
		Expected Count	22.0	42.0	16.0	80.0
		% within gender	27.5%	52.5%	20.0%	100.0%
		% within category	100.0%	100.0%	100.0%	100.0%

#### Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	22.442 <sup>a</sup>	2	.000
Likelihood Ratio	25.064	2	.000
N of Valid Cases	80		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.00.

### תשובות סופיות

- (1) א. 1.973      ב. קטן .  
                         ג. לא נדחה  $H_0$ .  
                         ד. 34.5%  
                         ה. 38.5%
- (2) א. בילוי מועדף ומוזיקה מועדף עם 3 קטגוריות לכל משתנה.  
                         ב. לא  
                         ג.  $H_0$  אין תלות בין בילוי מועדף למוזיקה מועדף.  
                         ד.  $H_1$  יש תלות בין בילוי מועדף למוזיקה מועדף.
- (3) א. 21.6%, 71.2%      ב. 15.8      ג. קטן מ-0.005.
- (4) א. לא נדחה  $H_0$ .  
                         ב. 11.4  
                         ג. 0.394  
                         ד. 20.8%.
- (5) א. פלא אי- חוקר א', פלא ב'- חוקר ב'.  
                         ב. קיבל את  $H_0$ .  
                         ג. טעות מסוג שני- הכרענו שאין הבדל בין גברים לנשים לмерות שיש במציאות הבדל.  
                         ד. קיבל את  $H_1$ .
- ה. טעות מסוג ראשון- הכרענו שיש קשר בין מין להוצאה לмерות שבמציאות אין קשר.  
 ו. כל חוקר פעל בשיטה סטטיסטית שונה ובמצב כזה יתכן מסקנות סותרות.

# bijustisjka le-refoat shinim

פרק 42 - ניתוח שונות חד כיוונית

תוכן העניינים

256 ..... 1. כללי .....

## ניתוח שונות חד כיוונית

### רקע תיאורטי

ניתוח שונות (חד כיווני) הוא מבחן להשוואת תוחלות ( $\mu_1, \dots, \mu_k$ ) של  $k$  אוכלוסיות שונות. לכן, בניתוח שונות, השערות המחקר הן:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k \quad (\text{התוחלות של כל האוכלוסיות שוות})$$

$$H_1: \quad \text{אחרת} \quad (\text{לפחות שתיים מהתוחלות שונות})$$

**הנחהות הדרושות לביצוע התהילה:**

2) בכל אוכלוסייה מתוך  $k$  האוכלוסיות ההתפלגות נורמלית.

3) כל האוכלוסיות הן עם אותה שונות  $\sigma^2$ .

4) המדגמים בלתי תלויים זה זהה.

ישנו משתנה המבדיל בין הקבוצות השונות, הוא המשתנה הבלתי תלוי הנקרא גורם (factor). משתנה זה הוא קטגוריאלי עם  $k$  רמות (levels).

כדי לבצע את התהילה יש לבצע מוגדים מכל אוכלוסייה:

נסמן ב-  $n_i$  את גודל המוגדים בקבוצה  $i$ .

$$\sum_{i=1}^k n_i = n - \text{מספר התצפויות סך הכל (בכל המדגמים)}.$$

$\bar{X}_i$  - ממוצע המוגדים הראשונים, ...,  $\bar{X}_k$  - ממוצע המוגדים ה-  $k$ -י.

$\bar{X}$  - ממוצע כללי (של כל המדגמים).

$$SS_B = \sum_{i=1}^k n_i [\bar{X}_i - \bar{X}]^2 \quad \text{סכום ריבועים בין הקבוצות:}$$

$$SS_W = \sum_{i=1}^k n_i [n_i - 1] \cdot \hat{s}_i^2 \quad \text{סכום ריבועים בתוך הקבוצות:}$$

$$SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} [X_{ij} - \bar{X}]^2 \quad \text{סכום ריבועים כללי:}$$

$$SST = SSB + SSW$$

יש למלא את טבלת ניתוח השונות הבאה:

מקור השונות	סכום הריבועים SS	דרגות חופש $df$	ממוצע הריבועים MS	F
B - בין הקבוצות	SSB	$k-1$	$\frac{SSB}{k-1}$	$\frac{MSB}{MSW}$
W - בתוך הקבוצות	SSW	$n-k$	$\frac{SSW}{n-k}$	
T - סה"כ	SST	$n-1$		

$$F = \frac{\frac{SSB}{k-1}}{\frac{SSW}{n-k}} \sim F(k-1, n-k)$$

אזור דחיה  $F > F_{(k-1, n-k)} : 1-\alpha$  :  $H_0$

## שאלות

**1)** מחקר מעוניין להשוות בין שלוש תרופות לשיכוך כאבים במטרה לבדוק האם קיים הבדל בין התרופות מבחינת הזמן בדיקות שלוקח עד שהתרופה משפיעה. לצורך הבדיקה נלקחו 15 אנשים שסובלים מכאב ראש. אנשים אלה חולקו באקראי לשולש: קבוצה 1 קיבלה "אקסמול" קבוצה 2 קיבלה "אופטיגין" קבוצה 3 קיבלה "נוורופן".

כל אדם במחקר מסר את מספר הדיקות עד שהתרופה השפיעה עליו.

א. מהו המשנה תלוי ומהו המשנה הבלתי תלוי במחקר?  
מהו ה"גורם" וכמה רמות יש לו?

ב. מהו המבחן הסטטיסטי המתאים כאן? רשמו את ההשערות.

ג. מה הן הנקודות הדרושות כדי לבצע את המבחן הסטטיסטי שהצעת בסעיף הקודם?

**2)** בעיר מסוימת שלושה בתים ספר תיכון. ראש העיר התעניין לבדוק האם קיים הבדל בהצלחה של בתים הספר במקצוע מתמטיקה. לצורך כך הוא דגם מספר תלמידים שנבחנו ב מבחון הבגרות במתמטיקה ברמה של 3 יחידות בעיר ובדק עברו כל תלמיד מה ציון הבגרות שלו במתמטיקה. להלן הציונים שהתקבלו:

"חס"	"רבינ'	"המתלמיד"
85	98	78
83	62	65
74	55	70
85	80	90
75		56

א. מהו המבחן הסטטיסטי המתאים?

רשמו את ההשערות ואת הנקודות של המבחן.

ב. מהו גודל המדגים? מהו המשנה הבלתי תלוי  
(factor) כמה רמות יש לו?

ג. חשבו את הממוצע ואת סטיית התקן של הציונים בכל אחד מהمدגמים.

ד. מלאו את טבלת ANOVA.

ה. רשמו את כלל ההכרעה לבחן שהוצע בסעיף א ברמת מובהקות של 5%.

ו. האם קיים הבדל בין בתים הספר בעיר מבחינה רמת הצלחת התלמידים במקצוע המתמטיקה? ענה על סמך הסעיפים הקודמים.

**3)** מעוניינים לבדוק האם יש הבדל בהשפעה של שיטות טיפול שונות על לחץ הדם הסיסטולי (SBP) באוכלוסייה של קשישים. נבדקו 4 שיטות שונות.

בטבלה המצורפת מוצגים ממצאי המחקר.

D	C	B	A	השיטה
12	8	14	12	גודל המדגים
182	180	172	178	הממוצע
3	5	8	4	סטיית התקן

א. רשמו את השערות המחקר והנקודות הדרושות כדי לבצע את המבחן המתאים.

ב. מה מסקנת המחקר ברמת מובהקות של 5%?

ג. האם יש צורך לבצע השוואות מרובות?

- 4) שלושה אופים נתקשו להכין עוגת שוקולד. לכל אופה בדקו את משך זמן ההכנה בדקות. כל אופה נדרש לאפות בכל יום 4 עוגות. האם קיים הבדל בין האופים מבחינת תוחלת זמן ההכנה של העוגות?

בדקו ברמת מובהקות של 5%.

שלוט	מוסס	ניר	האופה	סכום הזמנים	סכום ריבועי הזמן
182	212	206			
8982	11250	10644			

- 5) להלן טבלת ניתוח שונות חד כיוונית. במחקר בחנו 4 סוגי סוללות. רצוי לבדוק האם לסוג הסוללה השפעה על תוחלת אורץ החיים שלה. הפעילו את כל הסוללות על אותו מכשיר ובדקו את אורץ החיים של כל סוללה בשעות. מה המסקנה ברמת מובהקות של 10%? רשמו את ההשערות וההנחות הדרשיות.

### ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10.317	3	3.439	1.361	.279
Within Groups	60.648	24	2.527		
Total	70.964	27			

- 6) להלן טבלת ANOVA בטבלה הושמדו חלקים. השלימו את החלקים בטבלה שהושמדו ומסומנו באותיות.

### ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	357.450	ב	ג	ה	.000
Within Groups	א	17	ד		
Total	522.950	19			

7) חברת תרופות לκחה 15 אנשים ברמת בריאות דומה. החברה חילקה את האנשים לשולש קבוצות שוות בגודלן. לכל קבוצה ניתנה תרופה במינון שונה (dosage). המינונים שניתנו הם: 10 מ"ג, 20 מ"ג ו-30 מ"ג. לאחר שעה מזמן לκיקחת התרופה ניבדק קצב פעימות הלב של כל אדם (pulse). הנתונים הוזנו לתוכנית סטטיסטית והתקבלו התוצאות הבאות:

ANOVA

						pulse	
pulse						Tukey HSD <sup>a</sup>	
	Sum of Squares	df	Mean Square	F	Sig.	dosage	N
Between Groups	414.400	2	207.200	19.733	.000	30.00	5
Within Groups	126.000	12	10.500			20.00	5
Total	540.400	14				10.00	5

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

### Post Hoc Tests

Multiple Comparisons

		pulse		Tukey HSD		
(I) dosage	(J) dosage	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
10.00	20.00	3.20000	2.04939	.299	-2.2675	8.6675
	30.00	12.40000*	2.04939	.000	6.9325	17.8675
20.00	10.00	-3.20000	2.04939	.299	-8.6675	2.2675
	30.00	9.20000*	2.04939	.002	3.7325	14.6675
30.00	10.00	-12.40000*	2.04939	.000	-17.8675	-6.9325
	20.00	-9.20000*	2.04939	.002	-14.6675	-3.7325

\*. The mean difference is significant at the 0.05 level.

- א. בדקו ברמת מובהקות של 5% האם קיים הבדל בין המינונים השונים מבחינת תוחלת הדופק של האנשים? רשמו את ההשערות וההנחות הדרושים לצורך פתרון.
- ב. הסבירו ללא חישוב כיצד הייתה משתנה התשובה לסעיף הקודם אם היינו מעלים את הדופק של כל התכפיות במחקר ב-2.
- ג. האם יש צורך במחקר בהשואת מרובות. נמקו!
- ד. לטבלת ANOVA צורפו טבלאות של השוואות מרובות בשיטה הנקראית "טוקי". ברמת בטחון של 95% מה הם הממצאים לפי שיטה זו?

8) בעיר מסוימת רצוי לבדוק האם קיימים הבדל ברמה של התלמידים בין בתיה הספר השונים בעיר. ביצעו מוגם מכל בית ספר ונתנו מבחן זהה לכל הנדגמים. לאחר מכן ריכזו את הנתונים בתוכנה סטטיסטית והפיעלו ניתוח שונות. מצורפים הפלטים שהתקבלו. ענו על השעיפים הבאים:

- א. כמה בתיה ספר יש בעיר?
- ב. כמה תלמידים השתתפו בסך הכל במחקר?
- ג. האם קיימים הבדל בין בתיה הספר בעיר מבחינה רמת הציונים? בדקו ברמת מובהקות של 1%
- ד. בביטחון של 95% אילו בתיה ספר שונים זה מזה ברמת התלמידים? נמקו והסבירו.

### Oneway

ANOVA

grade	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7799.600	4	1949.900	13.586	.000
Within Groups	2870.400	20	143.520		
Total	10670.000	24			

## Post Hoc Tests

### Multiple Comparisons

grade

Scheffe

(I) school	(J) school	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	5.40000	7.57681	.971	-20.2543	31.0543
	3.00	36.80000*	7.57681	.003	11.1457	62.4543
	4.00	36.40000*	7.57681	.003	10.7457	62.0543
	5.00	-2.60000	7.57681	.998	-28.2543	23.0543
2.00	1.00	-5.40000	7.57681	.971	-31.0543	20.2543
	3.00	31.40000*	7.57681	.011	5.7457	57.0543
	4.00	31.00000*	7.57681	.013	5.3457	56.6543
	5.00	-8.00000	7.57681	.888	-33.6543	17.6543
3.00	1.00	-36.80000*	7.57681	.003	-62.4543	-11.1457
	2.00	-31.40000*	7.57681	.011	-57.0543	-5.7457
	4.00	-.40000	7.57681	1.000	-26.0543	25.2543
	5.00	-39.40000*	7.57681	.001	-65.0543	-13.7457
4.00	1.00	-36.40000*	7.57681	.003	-62.0543	-10.7457
	2.00	-31.00000*	7.57681	.013	-56.6543	-5.3457
	3.00	.40000	7.57681	1.000	-25.2543	26.0543
	5.00	-39.00000*	7.57681	.001	-64.6543	-13.3457
5.00	1.00	2.60000	7.57681	.998	-23.0543	28.2543
	2.00	8.00000	7.57681	.888	-17.6543	33.6543
	3.00	39.40000*	7.57681	.001	13.7457	65.0543
	4.00	39.00000*	7.57681	.001	13.3457	64.6543

\*. The mean difference is significant at the 0.05 level.

### Homogeneous Subsets

Scheffe <sup>a</sup>	school	N	grade	
			Subset for alpha = 0.05	
			1	2
	3.00	5	45.0000	
	4.00	5	45.4000	
	2.00	5		76.4000
	1.00	5		81.8000
	5.00	5		84.4000
	Sig.		1.000	.888

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

## תשובות סופיות

ב. ניתוח שונות חד כיווני.

**1)** א. משתנה בלתי תלוי : סוג התרופה.

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \text{otherwise}$$

משתנה תלוי : הזמן עד להשפעת התרופה בדקות.

ג. 1. מדגמים בלתי תלויים.

2. שווין שונות.

3. משתנים מתפלגים נורמלית.

**2)** א. המבחן לניתוח שונות חד כיוונית.

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \text{otherwise}$$

הנחות :

1. מדגמים בלתי תלויים.

2. משתנים מתפלגים נורמלית.

3. שווין שונות.

ב. גודל המדגם : 14. משתנה בית : בית הספר, בעל 3 רמות.

$$\bar{X} = 80.4, \hat{S} = 5.46, \bar{X} = 73.75, \hat{S} = 19.29, \bar{X} = 71.8, \hat{S} = 12.93$$

ג. להלן טבלה :

F	MS	df	SS	מקור השונות
	100.3	2	200.6	B
	173.2	11	1904.75	W
0.58		13	2105.35	סה"כ

ה.  $F > 3.98$ .

ו. קיבל את  $H_0$ .

ג. כן.

ב. נדחה את  $H_0$ .

**3)** א.  $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$   
 $H_1: \text{otherwise}$

הנחות :

1. מדגמים בלתי תלויים.

2. שווין שונות.

3. משתנים מתפלגים נורמלית.

(4) קיבל את  $H_0$  : נכريع שאין הבדל מובהק בין האופים מבחינת תוחלת זמן הכנה.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \quad (5)$$

$$H_1: \text{otherwise}$$

הנחות :

1. מדגמים בלתי תלויים.

2. שוויון שוניות.

3. משתנים מתפלגים נורמלית.

נקבל את  $H_0$  : לסוג סוללה אין השפעה של תוחלת החיים ברמת ביטחון של 10%.

(6) להלן טבלה :

## ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	357.450	2	178.725	ג 18.36	.000
Within Groups	165.5 א	17	9.735 ד		
Total	522.950	19			

$$H_0: \mu_1 = \mu_2 = \mu_3 \quad (7) \text{ א.}$$

$$H_1: \text{otherwise}$$

הנחות :

1. מדגמים בלתי תלויים.

2. משתנים מתפלגים נורמלית.

3. שוויון שוניות.

נדחה את  $H_0$  : ברמת ביטחון של 5% קיים הבדל במינונים השונים מבחינת תוחלת הדופק.

ב. ראה וידאו. ג. כן. ד.  $\mu_{30} > \mu_{10}$ .

(8) א. 5 ב. 25 ג. (

ג. נדחה את  $H_0$  : יש לפחות שני בתים ספר בעיר עם תוחלת רמת ציונים שונה.

ד.  $(\mu_3 = \mu_4) < (\mu_1 = \mu_2 = \mu_3)$ .

# bijustisjka le-refoat shinim

## פרק 43 - שאלות מסכמת בבדיקה השערות

### תוכן העניינים

265 .....	1. שאלות פתוחות מסכמות .....
268 .....	2. שאלות רב ברירה ( אמריקאיות) .....

## שאלות מסכמת בבדיקה השערות על פרמטרים

### שאלות

- 1) שני חוקרים נתקשו לבדוק את ההשערות הבאות:  $H_0: \mu = 520$ ,  $H_1: \mu > 520$ . כל חוקר בדק מדגם של 225 נחקרים. ידוע ש-  $\sigma = 20$ . חוקר א' קבע את כל הכרעה לפי  $\alpha = 0.05$ . חוקר ב' מחליט לדחות  $H_0$  אם  $\bar{X} > 522$ .
- למי מהחוקרים הסתברות לטעות מסוג ראשוני יותר?
  - מהי ההסתברות לטעות מסוג שני של חוקר ב' עבור  $\mu = 525$ ?
  - הסביר ללא חישוב נוספת, האם ההסתברות לטעות מסוג שני עבור  $\mu = 525$ , של חוקר א' שווה/קטנה/גדולה לו של חוקר ב'.
  - חוקר א' קיבל במדגם שלו  $\bar{X} = 523$ . מהי מסקנותו?
- 2) ידוע כי תוחלת מספר הליקים היומי של דנה היא 12 עם סטיית התקן 5. דני טוען שהוא יותר פופולארי ממנה בכך שהוא מקבל יותר ליקים ממנה ביום. על-מנת לבדוק זאת ספר דני כמה ליקים הוא קיבל בכל יום במשך 7 שבועות (כלומר, ב- 49 ימים) וקיבל סך-הcole 637 ליקים. נניח כי סטיית התקן של מספר הליקים שدني מקבל ביום זהה לסטיית התקן של דנה.
- מהי רמת המובהקות שכך לדריש, כדי שדנה תשתכנע בזכות טענתו (shedni פופולרי יותר בכך שהוא מקבל יותר ליקים ממנה ביום).
  - אם דני משער שתוחלת מספר הליקים שהוא מקבל ביום היא 14 וקובע רמת מובהקות 2.5%, מהי עוצמת המבחן של דני?
- | B | A | רשות | מוצר / רשות |
|---|---|------|-------------|
| 5 | 5 |      | 1           |
| 5 | 4 |      | 2           |
| 3 | 5 |      | 3           |
| 4 | 7 |      | 4           |
- 3) ברצוננו להשוות בין רשותות אלBIN B. לשם כך בחרנו 4 מוצרים, ובדקנו את מחיריהם בשתי הרשותות. להלן התוצאות: הניחו כי המחיר מתפלגים נורמלית.
- אם יש הנחות נוספות כדי לבצע את המבחן הפרטורי רשמו אותן.
- בדקו האם קיים הבדל בין הרשותות מבחינה תוחלת המחיר. רמת מובהקות של 5%.
  - חזרו על הסעיף הקודם בהנחה ונבחרו בכל רשות מוצרים באקראי ולא בהכרח אותם מוצרים.

4) במטרה לבדוק האם סטודנטים הלומדים במכילות משקיעים יותר זמן ללימודים מאשר סטודנטים באוניברסיטה נציגו 12 סטודנטים ובדקו לכל סטודנט את הזמן שהוא משקיע ביום ללימודים. הזמנים נמדדו בדקות:

סטודנטים באוניברסיטה					
סטודנטים במכילות					
180	140	171	189	156	176
150	204	186	191	190	180

- א. נסחו את ההשערות ובדוק אותן ברמת מובהקות של 5%. רשום את כל ההחלטה ואת ההנחות הדרושים לביצוע המבחן הפרמטרי.
- ב. חשבו את p-value.
- ג. ישנה טענה שממוצע זמן ההשקה בלימודים במכילות הוא 3.5 שעות ביום. בדוק את הטענה כאשר רמת המובהקות הינה 5%.
- 5) בשנת 2000 ל-60% היה מדיח כלים בבית. מחקר רוצה לבדוק האם כיום פרופורציית המשפחות עם מדיח כלים עלה. הוחלט לבצע מדגם אקראי של 150 משפחות.

ברמת בmadgoem	גברים					מה היא מסקנת המחקר 證明ות של 5% אם ל-102 משפחות היה מדיח כלים.
	1.1	1.2	0.7	0.9	2	
	1.2	1.8	1.9	1.1	1.4	

א. רשמו את השערות המחקר.

ב. מה היא מסקנת המחקר

ג. מהי הטעות האפשרית במסקנה מהすべיף הקודם.

האם ניתן לדעת את הסתירותה?

- 6) נערך מחקר על הקשר בין עישון ויתר לחץ דם. נבדק מדגם מקרי של 200 מעשנים ונמצא כי 30 סבלו מיתר לחץ דם.  
ידעו שבאוכולוסייה 18% סובלים מיתר לחץ דם.
- א. בדקו ברמת מובהקות 0.1 את ההשערה כי אחוז הסובלים מיתר לחץ דם בקרב המעשנים גדול מאשר כלל האוכלוסייה.
- ב. מהי רמת המובהקות המינימלית לקבלת הטענה שאחוז הסובלים מיתר לחץ דם בקרב המעשנים גדול מאשר כלל האוכלוסייה.
- ג. מהי עצמת המבחן, אם אחוז הסובלים מיתר לחץ דם בקרב אוכלוסייה המעשנים היא בפועל 25%.

- 7) להלן התפלגות מספר הנסיעות לחופשה השנתית במדגם של משפחות ישראליות. בדקו ברמת מובהקות של 5% :

מספר הנסיעות					
מספר המשפחות					
4	3	2	1	0	
12	20	26	102	84	

- א. באיטליה משפחות נסעות ב ממוצע פעמיים בשנה לחופשה. האם בישראל משפחות נסעות פחות מאשר באיטליה?
- ב. בהולנד 80% מהמשפחות נסעות לפחות פעם אחת בשנה לחופשה, האם בישראל אחוז המשפחות שננסעות לפחות פעם אחת בשנה לחופשה נמוך מאשר בהולנד?

8) נתון כי:  $X \sim N(\mu, \sigma^2 = 10^2)$

.  $H_0: \mu = 40$ ,  $H_1: \mu > 40$  מעוניינים לבדוק את ההשערות:

.  $\bar{X} = 45$  דגמו 25 תצפיות מהאוכלוסייה והתקבל

א. חשבו את p-value (МОובחיקות התוצאה).

ב. חזרו על סעיף א אם ההשערה האלטרנטיבית הייתה:  $\mu < 40$

ג. חזרו על סעיף א אם ההשערה האלטרנטיבית הייתה:  $\mu \neq 40$

### תשובות סופיות

1)	א. חוקר אי'	
2)	א. לפחות 0.0808	
3)	א. לא נדחה $H_0$	
4)	א. לא נדחה $H_0$	
5)	א. $H_0: p = 0.6$ $H_1: p > 0.6$	
6)	א. לא נדחה $H_0$	
7)	א. נדחה $H_0$ .	
8)	א. 0.0062	

ד. נדחה $H_0$	ג. גדלתה.	ב. 0.0122	ב. חוקר אי'
		ב. 0.7995	א. לפחות 0.0808
	ב. לא נדחה $H_0$	ב. לא נדחה $H_0$	א. לא נדחה $H_0$
	ג. נדחה $H_0$	ב. בין 5% ל-10%	א. לא נדחה $H_0$
		ג. טעות מסוג ראשון בסיכון של 0.05	א. נדחה $H_0$
		ב. נדחה $H_0$	א. נדחה $H_0$
		ב. 0.8643	א. לא נדחה $H_0$
		ב. נדחה $H_0$	א. נדחה $H_0$ .
		ב. 0.9938	א. 0.0062
		ג. 0.0124	

## שאלות סיוכם – שאלות רב ברירה על בדיקת השערות

(1) בבדיקה השערה חד-צדדית ימנית ברמת מובהקות  $\alpha = 0.01$ , נדחתה השערת האפס. מה הייתה המסקנה לו נבדקה אותה ההשערה באמצעות נתונים ברמת מובהקות  $\alpha = 0.05$  ?

- א. השערת האפס הייתה נדחתה.
- ב. השערת האפס לא הייתה נדחתה.
- ג. ההשערה המחקרית הייתה נדחתה.
- ד. בהעדר נתונים נוספים, לא ניתן לדעת.

(2) על מנת לבדוק האם ההסתברות לילדת בן הינה חצי, נבחר מוגם מקרי של 200 ילדים, ונמצא שישנם 120 בניים.  
מהו ההשערה האלטרנטיבית להשערת האפס?

- א.  $H_1: p = 0.5$
- ב.  $H_1: p = 0.6$
- ג.  $H_1: p > 0.5$
- ד.  $H_1: p \neq 0.5$

(3) לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני לא היפנוזה. לאחר מכן נערך לשניהם מבחן באנגלית. נניח שצינוי המבחן מתפלגים נורמלית ללא ידיעת השונות האמיתית. המבחן שיש לבצע כאן הוא :

- א. מבחן Z למוגם יחיד.
- ב. מבחן Z למדוגמים יחיד.
- ג. מבחן T למוגדים בלתי תלויים.
- ד. מבחן T למוגדים מזוגניים.

(4) כדי לבדוק את הטענה שגברים רווקים שוקלים פחות מגברים נשואים ללח חוקר מוגם מקרי של 4 גברים ומדד את משקלם לפני נישואיהם ולאחר נישואיהם. הנה התוצאות :

מהו ההשערות הנבדקות? (ההפרש חושב  $Y - X$ )

68	82	93	69	X לפני הנישואין -
71	84	88	80	Y לאחר הנישואין -

- א.  $H_1: \mu_d < 0, H_0: \mu_d = 0$
- ב.  $H_1: \mu_X - \mu_Y < 0, H_0: \mu_X - \mu_Y = 0$
- ג.  $H_1: \mu_X - \mu_Y < 0, H_0: \mu_X - \mu_Y = 0$
- ד.  $H_1: \mu_d > 0, H_0: \mu_d = 0$

5) חוקר ביצע מחקר ובו עשה טעות מסווג שני לכך :

- השערת האפס נcona.
- השערת האפס נדחתה.
- השערת האפס לא נדחתה.
- אף אחת מהתובשות לא נcona בהכרח.

6) ידוע כי ילד בגיל שנתיים ישן בממוצע 9 שעות בלילה. במדגם של 20 תינוקות

בני שנתיים המתגוררים בצפון נמצא, כי ממוצע שעות השינה בלילה הינו 10 עם סטיית תקן של 1.1. במדגם של 10 תינוקות בדרום נמצא, כי ממוצע שעות השינה בלילה הינו 7.9 עם סטיית תקן של 1.1. על מנת להשוות בין ממוצע שעות השינה של ילדים אלה לבין ממוצע השינה של ילדים מזרחיים, ועל מנת להשוות בין ממוצע השינה של ילדים לזו של ילדים המתגוררים בצפון יש לערוך \_\_\_\_\_.

יש להניח שההנחה הדורשota מתקיים.

- מבחן Z למדגם יחיד ; מבחן T למדגם יחיד.
- מבחן T למדגם יחיד ; מבחן T למדגמים תלויים.
- מבחן T למדגם יחיד ; מבחן T למדגמים בלתי תלויים.
- מבחן T למדגמים תלויים ; מבחן T לממוצע יחיד.

7) מובהקות התוצאה (PV) היא גם :

- רמת המובהקות המינימאלית לדחות השערת האפס.
- רמת המובהקות המקסימאלית לדוחית השערת האפס.
- רמת המובהקות שנקבעה מראש על ידי החוקר טרם קיבל את תוצאות המחקר.
- רמת המובהקות המינימאלית לאי דוחית השערת האפס.

כדי לבדוק את הטענה שגברים רווקים שוקלים פחות מגברים נשואים לעומת חוקר מדגם מקורי של 4 גברים ומדד את משקלם לפני נישואיהם ולאחר נישואיהם. הנה התוצאות:

	לפני הנישואין	לאחר הנישואין	
68	82	93	69
71	84	88	80

בדיקות

באיזה התפלגות משתמשים

ההשערות, ובכמה דרגות חופש :

- התפלגות Z ללא דרגות חופש.
- התפלגות T ו-3 דרגות חופש.
- התפלגות T ו-6 דרגות חופש.
- התפלגות  $\chi^2$  ו-3 דרגות חופש.

- 9) שני סטטיסטיקים בודקים השערות ברמת מובהקות  $\alpha = 0.05$  על סמך אותו מבחן. סטטיסטיקי א' בודק את ההשערה:  $H_0: \mu = 20$  נגד האלטרנטיבית  $H_1: \mu \neq 20$  ומחליט לא לדוח את השערת האפס. סטטיסטייקי ב' בודק את ההשערה  $H_0: \mu \leq 20$  נגד האלטרנטיבית  $H_1: \mu > 20$ . מה יחליט סטטיסטייקי ב'?
- לדוח את השערת האפס.
  - לא לדוח את השערת האפס.
  - לא נתוניים נוספים אי אפשר לדעת מה יחליט.
- 10) חוקר בדק השערה מסוימת והחליט לדוח את השערת האפס ברמת מובהקות 5%. מה נכון לוומר?
- הוא בודאות ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.
  - הוא בודאות לא ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.
  - הוא בודאות ידחה את השערת האפס ברמת מובהקות 9% וברמת מובהקות 2%.
  - הוא בודאות לא ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.
- 11) רמת הcolesterol בדם של אנשים מתפלג נורמלית עם תוחלת של 180 מ"ג (ל 100 סמ"ק דם). וטיפות התקן של 10 מ"ג. מעוניינים לבדוק את הטענה שצמחיים הם בעלי רמתコレsterol נמוכה יותר. נניח שטיפות התקן אצל צמחוניים זהה לטרופית התקן של כלל האנשים. במקרה של 20 צמחוניים התקבל ממוצע רמתコレsterol 174.5 מ"ג. אם הוחלט לקבל את הטענה שצמחיים הם בעלי רמתコレsterol נמוכה יותר איזה סוג טעות אפשרית במסקנה?
- טעות מסוג ראשון.
  - טעות מסוג שני.
  - טעות מסוג שלישי.
  - לא ניתן לדעת כיון שאין לנו לידעים מה התוחלת האמיתית אצל הצמחוניים.

**12)** בסקר שנערך התקבל ש 60% מתוך 220 נשאלים מבקרים אצל השיננית לפחות פעם אחת בשנה. עבור אילו רמות מובהקות ניתן יהיה לקבוע שרוב האוכלוסייה מבקרת אצל השיננית לפחות פעם אחת בשנה?

- א. רמת מובהקות הגדולה מ-5%.
- ב. רמת מובהקות הקטנה מ-5%.
- ג. רמת מובהקות הגדלה מ-0.0015.
- ד. רמת מובהקות הקטנה מ-0.0015.

**13)** שני חוקרים העוסקים בתחום מחקרי משותף החליטו להסתמך על נתונים של מדגם שפורסם על ידי הלשכה המרכזית לסטטיסטיקה.

חוקר א' ניסח השערה זו צדדית ואילו חוקר ב' ניסח השערה חד צדדית. מסקנתו של איזה מבין המשפטים הבאים הוא הנכון בנוגע למסקנות החוקרים?

- א. אם חוקר א' ידחה את השערת האפס לא ניתן לדעת מה יחליט חוקר ב' באוטה רמת מובהקות.
- ב. אם חוקר א' קיבל את השערת האפס גם חוקר ב' קיבל את השערת האפס באותה רמת מובהקות.
- ג. אם חוקר ב' ידחה את השערת האפס גם חוקר א' ידחה את השערת האפס באותה רמת מובהקות.
- ד. אם חוקר א' ידחה את השערת האפס גם חוקר ב' ידחה את השערת האפס בתנאי שרמת המובהקות כפולה בגודלה.

**14)** ידוע מנתוני העבר כי תוחלת הציונים בבחינה בפסיכולוגיה היא 79. הועלתה השערה כי תוחלת הציונים בקרב העולים החדשניים נמוכה יותר. לצורך בדיקת הטענה נלקח מדגם מקרי של 47 סטודנטים עולים ונמצא ממוצע של 75. מה משמעות הפרמטר בניסוח ההשערות?

- א. תוחלת ציוני העולים באוכלוסייה.
- ב. ממוצע ציוני העולים במדגם.
- ג. תוחלת ציוני האוכלוסייה מנתוני העבר.
- ד. ממוצע ציוני שאר האוכלוסייה במדגם.

**15)** חוקר ביצע מחקר וידוע כי עשה טעות מסווג 1. מה מהබאים נכון?

- א. החוקר דחה את השערת  $H_0$  כאשר היא הייתה נכון.
- ב. החוקר דחה את השערת  $H_1$  כאשר היא הייתה נכון.
- ג. החוקר לא דחה את השערת  $H_0$  כאשר היא הייתה לא נכון.
- ד. המדגם של החוקר שייך בפועל להתפלגות הדגימה של  $H_1$ .

**16)** חוקר ביקש לבחון האם תאומים זהים אשר הופרדו בילדותם שונים מתאומים זהים אשר גדלו יחדיו מבחינות מידת הפער בין התאומים בלחץ הדם. הוא דגם 20 זוגות תאומים מכל אוכלוסייה ומדד את הפרש בין לחץ הדם בכל זוג תאומים. מהו המבחן הסטטיסטי המתאים?

- מבחן D למדגים בלתי תלויים עם 38 דרגות חופש.
- מבחן D למדגים מזוגיים, עם 39 דרגות חופש.
- מבחן D למדגים בלתי תלויים עם 39 דרגות חופש.
- מבחן D למדגים מזוגיים עם 38 דרגות חופש.

**17)** בינוואר השנה פורסם שהשכר הממוצע במשק הוא 900,9₪. במדגם שנעשה בחודש יוני על 60 עובדים רשום עבר כל עובד במדגם האם השכר שלו נמוך או לא נמוך מהשכר הממוצע שפורסם בחודש ינואר. מהו המבחן המתאים כדי לבדוק שרוב העובדים בחודש יוני קיבלו שכר הנמוך מהשכר הממוצע שפורסם בחודש ינואר?

- מבחן Z על פרופורציה.
- מבחן Z על תוחלת אחת.
- מבחן Z על שתי תוחלות במדגים בלתי תלויים.
- מבחן Z על שתי תוחלות במדגים תלויים.

**18)** שלושה חוקרים רצו לבדוק את השפעתו של שידור פרסומות נגד תאונות דרכים על מהירות הנהיגה של נהגים בישראל (השינויים של מהירות הנהיגה בישראל אינה ידועה). עידו השווה את מהירות הנהיגה של קבוצת נהגים אחת, חודש לפני שידור הפרסומות וחודש לאחר שידור הפרסומות. רון השווה את מהירות הנהיגה של קבוצת נהגים, שראו את הפרסומות, ל מהירות הנהיגה של קבוצת נהגים, שלא ראו את הפרסומות. יואב השווה את מהירות הנהיגה הממוצעת בישראל על פי נתוני משרד התחבורה. הפרסומות, ומהירות הנהיגה הממוצעת בישראל להשתמש הם :

- שלושתם במבחן T למדגים בלתי תלויים.
- עידו במבחן T למדגים מזוגיים, רון ויואב במבחן T למדגים בלתי תלויים.
- עידו במבחן T למדגים מזוגיים, רון במבחן T למדגים בלתי תלויים ויואב במבחן T למדגם יחיד.
- עידו במבחן T למדגים מזוגיים, רון ויואב במבחן T למדגם יחיד.

19) במחקר נמצא שתוצאה היא מובהקת ברמת מובהקות של 5%. מה תמיד נכון?

- א. הגדלת רמת המובהקות לא תשנה את מסקנת המחקר.
- ב. הגדלת רמת המובהקות תנסה את מסקנת המחקר.
- ג. הקטנת רמת המובהקות לא תשנה את מסקנת המחקר.
- ד. הקטנת רמת המובהקות תנסה את מסקנת המחקר.

20) חוקר ערך מבחן דו צדי ברמת מובהקות של  $\alpha$  והחליט לדוח את השערת האפס. אם החוקר היה עורך מבחן חד צדי ברמת מובהקות של  $\frac{\alpha}{2}$  איזי בהכרח:

- א. השערת האפס הייתה נדחתה.
- ב. השערת האפס הייתה לא נדחתה.
- ג. לא ניתן לדעת מה תהיה מסקנתו במקרה זה.

21) ליאור ורוני העלו את אותן השערות על ממוצע האוכלוסייה. כמו כן הם התבססו על אותן תוצאות של מדגמים.

ליאור השתמש בטבלה של התפלגות Z.

רוני השתמשה בטבלה של התפלגות T.

מה יוכל לומר בנוגע להחלטת המחקר שלהם?

- א. אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.
- ב. אם רוני תדחה את השערת האפס אז גם בהכרח ליאור.
- ג. שני החוקרים בהכרח הגיעו לאותה מסקנה.
- ד. לא ניתן לדעת על היחס בין דחינת השערת האפס של שני החוקרים.

22) נתון ש  $(\sigma^2, \mu) \sim N - X$  כמו כן נתונים ההשערות הבאות:  $H_0: \mu = \mu_0$ ,  $H_1: \mu < \mu_0$ .

חוקר בדק את ההשערות הללו על סמך מדגם שככל 10 תוצאות.  $\sigma^2$  לא הייתה ידועה לחוקר. החוקר החליט לדוח את השערת האפס ברמת מובהקות של 5% לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תוצאות וiscalל את תוצאות אלה גם למדגם כך שככל עכשו 15 תוצאות.

- א. כתע בברור הוא ידחה את השערת האפס.
- ב. כתע הוא דוקoa קיבל את השערת האפס.
- ג. כתע לא ניתן לדעת מה תהיה מסקנתו.

23) אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו איזי:

- א. הסיכוי לטעות מסוג ראשוני גדול.
- ב. העוצמה של המבחן גבוהה.
- ג. הסיכוי לטעות מסוג שני��大.
- ד. תשובות או-וב נכונות.

**24)** חוקר ביצע מחקר ובו עשה טעות מסווג שני לכך :

- השערת האפס נכונה.
- השערת האפס נדחתה.
- השערת האפס לא נדחתה.
- אף אחת מהתובשות לא נכון בהכרח.

**25)** מה המצב הרצוי לחוקר המבצע בבדיקה השערה :

- |          |             |
|----------|-------------|
| $\alpha$ | $1 - \beta$ |
| א. גדולה | גדולה       |
| ב. גדולה | קטנה        |
| ג. קטנה  | גדולה       |
| ד. קטנה  | קטנה        |

**26)** נערך שינוי בכלל החלטה של בדיקת השערה מסוימת ובקבוקתו אוזור דחיה  $H_0$  קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה מכך :

- הן  $\alpha$ , והן  $(\beta - 1)$ , יקטנו.
- $\alpha$  יישאר ללא שינוי ואילו  $(\beta - 1)$  יגדל.
- $\alpha$  יגדל ואילו  $(\beta - 1)$  יקטנו.
- הן  $\alpha$  והן  $(\beta - 1)$  יגדלו.

**27)** ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניה שלחץ הדם בקרוב עיתונאים גבוה יותר מה ממוצע באוכלוסייה. הואלקח מדגם של 60 עיתונאים וקיים ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק שלחץ הדם בקרוב העיתונאים אינו גבוה יותר. מה הטעות האפשרית שהרופא עושה?

- טעות מסווג ראשון.
- טעות מסווג שני.
- טעות מסווג שלישי.
- אין טעות במסקنته.

**28)** בבדיקה השערות התקבל שה-  $p-value = 0.02$ . מה תהיה מסקנת חוקר המשמש ברמת מובהקות 1%? בחר בתשובה הנכונה :

- יקבל את השערת האפס בכל מקרה.
- ידחה את השערת האפס מקרה.
- ידחה את השערת האפס רק אם המבחן הינו דו צדדי.
- לא ניתן לדעת כי אין מספיק נתונים.

**29)** מובಹקות התוצאות (PV) היא גם :

- רמת המובಹקות המינימאלית לדוחות השערת האפס.
- רמת המובಹקות המקסימאלית לדוחית השערת האפס.
- רמת המובಹקות שנקבעת מראש על ידי החוקר טרם קיבל את תוצאות המחקר.
- רמת המובಹקות המינימאלית לאי דוחית השערת האפס.

**30)** בבדיקה השערות מסוימת התקבל  $p value = 0.0254$ , לכן :

- ברמת מובಹקות של 0.01 אך לא של 0.05 נדחה את  $H_0$ .
- ברמת מובಹקות של 0.01 ושל 0.05 לא נדחה את  $H_0$ .
- ברמת מובಹקות של 0.05 אך לא של 0.01 נדחה את  $H_0$ .
- ברמת מובಹקות של 0.01 ושל 0.05 נדחה את  $H_0$ .

**31)** רמת המובಹקות במחקר הייתה 2% לכן.

- בסיוכו של 2% נדחה את השערת האפס.
- בסיוכו של 2% לא נדחה את השערת האפס.
- בסיוכו של 2% השערת האפס לא נכונה.
- אף תשובה לא נכונה.

**32)** נתון ש:  $(\mu, \sigma^2) \sim N$ . כמו כן נתונות ההשערות הבאות:  $H_0: \mu = \mu_0$ ,  $H_1: \mu < \mu_0$ .

חוקר בדק את ההשערות הללו על סמך מדגם שכלל 10 תצפיות.

$\sigma^2$  לא הייתה ידועה לחוקר. החוקר החליט לדוחות את השערת האפס ברמת מובಹקות של 5%. אם הוא היה מגדיל את רמת המובಹקות ל-10% אזי:

- icut בברור הוא ידחה את השערת האפס.
- icut הוא דווקא קיבל את השערת האפס.
- icut לא ניתן לדעת מה תהיה מסקנתו.

**33)** לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני ללא היפנוזה.

לאחר מכן נערכן לשניהם מבחן באנגלית. ננית שצינוי המבחן מתפלגים נורמללית ללא ידיעת השונות האטומית. מספר דרגות החופש במבחן הוא:

- 9
- 19
- 18
- 8

(34) בתקנת טיפת חלב מסוימת יש שני מכשירי שקילה. על מנת להשוות בין שני המשקלים נדגו 4 תינוקות. כל תינוק בן חודשיים נשקל בכל אחד מהמשקלים. להלן תוצאות השקילה (בק"ג) :

משקל במיכיר 1	2.5	0.7	9.6	4.5
משקל במיכיר 2	0.5	1.7	6.9	3.5

נניח שהמשקלים מתפלגים נורמלית.

המבחן שיש לבצע כאן הוא :

- א. מבחן Z למדגם יחיד.
- ב. מבחן T למדגם יחיד.
- ג. מבחן T למדגמים בלתי תלויים.
- ד. מבחן T למדגמים מזוגים.

(35) כדי להשוות בין שני אצנים נדגו 5 תוצאות מריצת 100 מטר של כל אצן. זמני הריצה נרשמו ויש להניח שמתפלגים נורמלית. המטרה להשוות בין האצנים.

המבחן שיש לבצע כאן הוא :

- א. מבחן Z למדגם יחיד.
- ב. מבחן T למדגם יחיד.
- ג. מבחן T למדגמים בלתי תלויים.
- ד. מבחן T למדגמים מזוגים.

(36) סטטיסטיκאי ערך מבחן סטטיסטי. הוא חישב את עצמת המבחן וקיבל 0. המשמעות של תוצאה זו היא :

- א. לעולם לא לדוחות את השערת האפס כאשר היא לא נכונה.
- ב. תמיד לדוחות את השערת האפס כאשר היא נכונה.
- ג. לעולם לא לדוחות את השערת האפס כאשר היא נכונה.
- ד. תמיד לדוחות את השערת האפס כאשר היא לא נכונה.

(37) סטטיסטיκאי נתקבש לאמוד את הפרש הממוצעים של שני טיפולים לפי שני מדגמים מקרים בלתי תלויים. הוא חישב רוחסן סמך להפרש ברמת סמך 0.98 וקיבל את הרוחסן  $\mu_2 - \mu_1 < 4.5$ . אילו יתבקש החוקר לבדוק לפי אותן

נתונים את השערות :  $H_1: \mu_1 - \mu_2 \neq 0$  ;  $H_0: \mu_1 - \mu_2 = 0$ ,

ברמת מובהקות 0.05 מסקנתנו תהיה :

- א. לדוחות את השערת האפס.
- ב. לא לדוחות את השערת האפס.
- ג. שלא ניתן לדעת את המסקנה עבור רמת מובהקות 0.05.
- ד. שלא נתנות בשאלת סטיות התקן של האוכלוסיות, ולכן לא ניתן להסיק דבר.

**38)** במטרה לבדוק האם קיימים הבדל בין קוווי זהב לבזק מבחינת ממוצע המחרירים לשיחות בינייל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקט שיחה. בהנחה והמחרירים מתפלים נורמלית בנו רוח סמך לממוצע ההפרשים וקיבלו:  $\bar{x} = 0.2145$ ,  $s_d = 0.0293$ . רוח הסמך הוא ברמת סמך של 95%. לכן מסקנת המחקר היא:

- ברמת מובהקות של 5% לא נוכל לקבוע שקיים הבדל בין החברות.
- ברמת מובהקות של 5% נקבע שקיים הבדל מובהק בין החברות.
- לא ניתן לדעת מה המסקנה ברמת מובהקות של 5% כיון שלא נאמר מה ההגדרה של  $D$ .

**39)** אם רמת מובהקות של מבחן סטטיסטי הינה 0, הכוונה היא:

- תמיד נדחה  $H_0$  כאשר היא נכונה, אך לא תמיד נדחה אותה כאשר היא לא נכונה.
- לא נדחה את  $H_0$  אף פעם.
- לא נדחה את  $H_0$  כאשר היא נכונה אך יתכן ונדחה אותה כאשר היא לא נכונה.
- כל התשובות לא נכונות.

**40)** חוקר ביצע ניסוי. הוא ניסח את ההשערות הבאות:  $H_0: \mu = 10$ ,  $H_1: \mu \neq 10$ . לצורך בדיקה הואלקח מוגרבי בגודל 5 מתוך אוכלוסייה המתפלגת נורמלית עם שונות לא ידועה. על סמך תוצאות המוגרם הוא חישב וקיבל:  $t_{\bar{x}} = -2.63$ . לכן המסקנה היא:

- הוא ידחה  $H_0$  ברמת מובהקות 0.1 אך לא כן ברמת מובהקות 0.05.
- הוא ידחה  $H_0$  ברמת מובהקות 0.05 אך לא כן ברמת מובהקות 0.025.
- הוא ידחה  $H_0$  ברמת מובהקות 0.025 אך לא כן ברמת מובהקות 0.01.
- הוא לא ידחה  $H_0$  ברמת מובהקות 0.1.

**41)** האיגוד האמריקני לרפואת ילדים מפרסם הנחיות חדשות הקובעות כי יש ליטול תוספת יוד במהלך תקופת ההריון וההנקה. מחסור במינרל זה עלול לגרום לפגיעה מוחית אצל העובר והתינוק. החלטה זו נקבעה על סמך מחקר בו השתתפו 1050 נשים שנטלו יוד במהלך תקופת ההריון וההנקה.

מתוך הנשים שהשתתפו במחקר, רק ל-21 נמצאו ילדים בעלי פגיעה מוחית לעומת 3% באוכלוסייה הכללית. בנוספ', פורסם שהאיגוד האמריקאי מגיע למסקנותיו על סמך רמת מובהקות של 0.5%. מה הסיכוי לבצע טעות מסווג ראשון במחקר?

- 0.005
- 0.03
- 0.0287
- 0.05

(42) חוקרת שיערה, כי משקלן של נשים כשנה לאחר החתונה גבוהה ממשקלן בעת החתונה. החוקרת דגמה 15 נשים, ובדקה את משקלן בשתי נקודות הזמן (בעת החתונה, ו שנה לאחריה), אך לא מצאה הבדל מובהק ברמת מובהקות 0.01. בהנחה, כי **במציאות** השערתה של החוקרת נכונה, סביר כי אם היא תגדיל את גודל המדגם, אז:

- א. יקטן הסיכוי לטעות מסווג שני ( $\beta$ ).
- ב. תגדל רמת הביטחון ( $\alpha - 1$ ).
- ג. אף תשובה לא נכונה.
- ד. כל התשובות נכונות.

(43) איזה מה המשפטים הבאים נכון תמיד?

- א.  $\text{POWER} + \alpha + \beta = 1$
- ב.  $\text{POWER} = 0.5 - \beta$
- ג.  $\text{POWER} + \alpha = 1$
- ד.  $\beta + \alpha = 1$
- ה. הכל לא נכון.

(44) מה נכון לומר לגבי הנחת שיוויון השווניות ב מבחן T למדגים בלתי תלויים?

- א. היא אומרת שהשווניות המדגמיות שווה.
- ב. בטעדייה אין שום דרך לבדוק השערה על הפרש בין תוחלות.
- ג. היא חשובה הן עבור מדגים מזוגיים והן עבור מדגים בלתי תלויים.
- ד. אף תשובה אינה נכונה.

(45) חוקר החליט לא לדוחות השערה ברמת מובהקות של  $\alpha$ . במידה וחוקר זה היה בודק השערה זו ברמת מובהקות של  $\alpha/2$  על סמך אותם נתונים, האם ההשערה תדחה?

- א. ההשערה תדחה.
- ב. ההשערה לא תדחה.
- ג. התשובה תלוי בעוצמת המבחן.
- ד. לא ניתן לדעת בוודאות אם ההשערה תדחה או לא.

(46) חוקרת שיערה, כי בגילאי הגן בנות יותר תקשורתית מבנים. אם החוקרת תדגום אקרואית 30 בנים ו-30 בנות, ובמדגם יתקבלו ממוצע של ציון תקשורת. סטטיסטי המבחן יהיה:

- א. אפס
- ב. חיובי
- ג. שלילי
- ד. לא ניתן לדעת

47) עצמה שווה ל-1 פרושה :

- לעולם לא לדחות את השערת האפס כאשר היא נכונה.
- תמיד לדחות את השערת האפס כאשר היא נכונה.
- לעולם לא לדחות את השערת האפס כאשר היא לא נכונה.

48) מה מהבאים נכון לגבי מבחן T מוגדים מזווגים?

- כל התצפיות במחקר אינן תלויות זו בזו.
- כל התצפיות במחקר תלויות זו בזו.
- כל הצמידים של תצפיות במחקר אינם תלויים זה זה.
- התצפיות בתוך כל צמד אינן תלויות זו בזו.

49) לבדיקת ההשערה חד צדדית על התוחלת של התפלגות נורמלית  $\mu \geq 10$ ,  $H_0: \mu = 10$ . נלקח מבחן ותקבלה רמת מובהקות מינימאלית לדחיה השערת האפס 0.058. לו רצינו לבדוק את ההשערה הדו צדדית  $\mu \neq 10$ ,  $H_1: \mu = 10$ , אז על סמך תוצאות אותו המבחן ברמת מובהקות 0.05:

- ניתן להכיריע בין ההשערות רק אם שונות האוכלוסייה נתונה.
- מקבלים את השערת האפס.
- זוחים את השערת האפס.
- לא ניתן להכיריע בין ההשערות שכן חסרים נתונים.

50) לבדיקת ההשערה חד צדדית ימנית  $\mu = 55$ ,  $H_0: \mu = 55$ ,  $H_1: \mu > 55$ . נלקח מבחן מקרי בגודל  $n$  מאוכלוסייה בעלת התפלגות נורמלית ושונות  $\sigma^2$ . רמת המובהקות היא 5%. נמצא שהעוצמה היא 0.9. להלן 3 טענות:

- עבור מבחן בגודל  $n$  ורמות מובהקות 5% לבדיקת ההשערות:  
 $H_0: \mu = 55$ ,  $H_1: \mu = 60$  העוצמה תהיה גדולה מ-0.9.
  - עבור מבחן בגודל  $n$  ורמות מובהקות 5% לבדיקת ההשערות:  
 $H_0: \mu = 55$ ,  $H_1: \mu = 65$  העוצמה תהיה גדולה מ-0.9.
  - עבור מבחן בגודל  $n$  ורמות מובהקות 10% לבדיקת ההשערות:  
 $H_0: \mu = 55$ ,  $H_1: \mu = 65$  העוצמה תהיה קטנה מ-0.9.
- שלושת הטענות אינן נכונות.
  - טענות 2 ו-3 אינן נכונות וטענה 1 נכונה.
  - טענת 1 ו-2 נכונות וטענה 3 אינה נכונה.
  - טענות 1 ו-3 אינן נכונות וטענה 2 נכונה.

**תשובות סופיות:**

שאלה	תשובה	שאלה	תשובה	שאלה
א	26	א	א	1
ב	27	ד	ד	2
א	28	ד	ד	3
א	29	א	א	4
ג	30	ג	ג	5
ד	31	ג	ג	6
א	32	א	א	7
א	33	ב	ב	8
ד	34	ג	ג	9
ג	35	א	א	10
א	36	א	א	11
ג	37	ג	ג	12
א	38	א	א	13
ג	39	א	א	14
א	40	א	א	15
א	41	א	א	16
א	42	א	א	17
ה	43	ג	ג	18
ד	44	א	א	19
ד	45	ג	ג	20
א	46	ב	ב	21
ד	47	ג	ג	22
ג	48	ד	ד	23
ב	49	ג	ג	24
ד	50	ג	ג	25