

# סטטיסטיקה והסתברות 01 88165



$$\{\sqrt{x}\}^2$$



## תוכן העניינים

1	סטטיסטיקה תיאורית-הקדמה
4	סטטיסטיקה תיאורית- הצגה של נתונים
15	סטטיסטיקה תיאורית-גבולות מדומים ואמיתיים
17	סטטיסטיקה תיאורית- סכימה
21	סטטיסטיקה תיאורית - מדדי מיקום מרכזי
30	סטטיסטיקה תיאורית - מדדי פיזור - הטווח, השונות וסטיית התקן
33	סטטיסטיקה תיאורית - מדדי פיזור - טווח בין רבעוני
35	סטטיסטיקה תיאורית- מדדי מיקום יחסי-ציון תקן
37	סטטיסטיקה תיאורית-מדדי מיקום יחסי-אחוזונים במחלקות
40	סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה
42	סטטיסטיקה תיאורית - טרנספורמציה לינארית
45	סטטיסטיקה תיאורית - שאלות מסכמות
52	סטטיסטיקה תיאורית -מקדם המתאם ( מדד קשר ) הלינארי
68	יסודות ההסתברות
72	פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים
81	קומבינטוריקה -כלל המכפלה
85	קומבינטוריקה- תמורה - סידור עצמים בשורה
88	קומבינטוריקה - תמורה עם עצמים זהים
90	קומבינטוריקה- סידור עצמים במעגל
93	קומבינטוריקה -דגימה סידורית ללא החזרה ועם החזרה
95	קומבינטוריקה - דגימה ללא סדר וללא החזרה
98	קומבינטוריקה - שאלות מסכמות
105	הסתברות מותנית-במרחב מדגם אחיד

# תוכן העניינים

108	24. הסתברות מותנית - מרחב לא אחיד
112	25. דיאגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה
117	26. תלות ואי תלות בין מאורעות
121	27. שאלות מסכמות בהסתברות
126	28. המשתנה המקרי הבדיד - פונקציית ההסתברות
130	29. המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן
134	30. המשתנה המקרי הבדיד -תוחלת של פונקציה של משתנה מקרי בדיד
137	31. המשתנה המקרי הבדיד- טרנספורמציה לינארית
140	32. תוחלת ושונות של סכום משתנים מקריים
143	33. התפלגויות בדידות מיוחדות -התפלגות בינומית
147	34. התפלגויות בדידות מיוחדות -התפלגות גיאומטרית
150	35. התפלגויות בדידות מיוחדות -התפלגות אחידה
153	36. התפלגויות בדידות מיוחדות- התפלגות פואסונית
156	37. התפלגויות בדידות מיוחדות-התפלגות היפרגאומטרית
159	38. התפלגויות בדידות מיוחדות -התפלגות בינומית שלילית
162	39. המשתנה המקרי הבדיד - שאלות מסכמות
169	40. המשתנה המקרי הרציף - התפלגויות כלליות ללא אינטגרלים
175	41. התפלגויות רציפות מיוחדות- התפלגות מעריכית
178	42. התפלגויות רציפות מיוחדות-התפלגות אחידה
181	43. התפלגויות רציפות מיוחדות - התפלגות נורמלית
189	44. טרנספורמציה על משתנה מקרי רציף
192	45. פונקציה יוצרת מומנטים
198	46. תכונות של פונקציית יוצרת מומנטים
203	47. משתנה דו-מימדי בדיד - פונקציית הסתברות משותפת
209	48. משתנה דו מימדי בדיד - מתאם בין משתנים
216	49. המשתנה המקרי הדו ממדי - קומבינציות ליניאריות
219	50. המשתנה המקרי הדו ממדי הבדיד - שאלות מסכמות
227	51. קומבינציות לינאריות על התפלגות נורמלית

230	המשתנה המקרי הדו ממדי הרציף
238	קונבולוציה - התפלגות סכום משתנים בלתי תלויים
241	חישוב תוחלת ושונות על ידי פירוק לאינדיקטורים
244	נוסחת התוחלת השלמה
247	נוסחת השונות השלמה ( שונות של משתנה מותנה )
249	הסקה סטטיסטית - הקדמה
252	התפלגות הדגימה ומשפט הגבול המרכזי
275	אי שוויונים בהסתברות
283	מושגי יסוד באמידה
288	אמידה נקודתית
303	רווח סמך לתוחלת (ממוצע)
314	רווח סמך להפרש תוחלות (ממוצעים) במדגמים בלתי תלויים
318	רווח סמך לשונות וסטיית תקן
323	בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן)
330	מבוא לבדיקת השערות על פרמטרים
336	בדיקת השערות על תוחלת (ממוצע)
363	בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים
366	בדיקת השערות על שונות

# סטטיסטיקה והסתברות 01 88165

פרק 1 - סטטיסטיקה תיאורית-הקדמה

תוכן העניינים

1. כללי ..... 1

## סטטיסטיקה תיאורית – הקדמה:

### רקע:

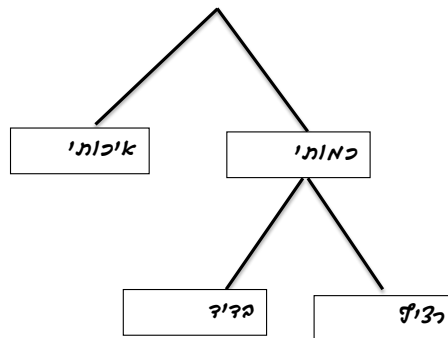
בסטטיסטיקה תיאורית אנו חוקרים קבוצה מסוימת, שיכולה להיות קבוצת ילדים בגן, קבוצת מניות בתיק, כלל התושבים בעיר מסוימת וכו'. בין ישות לישות בקבוצה ישנם גורמים היכולים לקבל מספר ערכים. גורמים אלה נקראים משתנים. למשל, בין מניה למניה בתיק משתנה התשואה היומית של המניה, הוותק של המניה, תחום המניה וכדומה. בסטטיסטיקה תיאורית אנחנו נתבונן בקבוצה מסוימת ובתוך הקבוצה הזו נאסוף נתונים לגבי משתנה מסוים ונלמד להציג את הנתונים ולנתח אותם מכל מיני אספקטים.

### דוגמה:

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

- 1) מי הקבוצה הנחקרת?
- 2) מה גודל הקבוצה?
- 3) מה המשתנה הנחקר?

### סוגי משתנים:



### משתנה איכותי

משתנה שלערכיו אין משמעות של יותר או פחות, אין עניין כמותי לערכים המתקבלים. כמו: מקום מגורים של אדם (רעננה, תל אביב, אשדוד...), מין האדם (זכר, נקבה) ומצב משפחתי (רווק, נשוי, גרוש, אלמן).

### משתנה כמותי

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

את המשתנה הכמותי נסווג לשני סוגים:

1. משתנה בדיד – משתנה שערכיו מתקבלים מתוך סידרה של ערכים אפשריים. כמו: מספר ילדים למשפחה (1,2,3...). וציון בבחינה (מ-0 ועד 100 בקפוצות של 1).
2. משתנה רציף – משתנה שערכיו מתקבלים מתוך אינסוף ערכים בתחום מסוים. הערכים מתקבלים ברצף וללא קפיצות של ערכים. כמו: גובה בס"מ – אם למשל, הגובה הנמוך ביותר הוא 150 ועד 190 ס"מ בקבוצה הגבהים הם ברצף. גם בין 160 ל-161 ס"מ יש רצף אינסופי של ערכים אפשריים לגובה (160.33 ס"מ הוא גם גובה אפשרי), משקל בק"ג, מהירות בקמ"ש וכולי.

## שאלות:

- (1) סווגו את המשתנים הבאים לפי: איכותי / כמותי בדיד / כמותי רציף:
- מספר הדירות בבניין.
  - גיל אדם בשנים.
  - אחוז האבטלה בעיר.
  - מקצוע לימוד מועדף.

- (2) להלן התפלגות מספר האיחורים לעבודה בחודש של העובדים בחברת "סטאר":

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

בחברה 200 עובדים.

- מהו המשתנה הנחקר כאן?
  - האם מדובר במשתנה איכותי או כמותי?
  - אם הוא כמותי האם הוא בדיד או רציף?
- (3) להלן רשימה של משתנים כמותיים, ציינו ליד כל אחד אם הוא רציף או בדיד:
- שכר עובד ב-ש.
  - ציון בחינת בגרות.
  - תוצאה בהטלת קובייה.
  - מהירות ריצה בתחרות.
  - שיעור התמיכה בממשלה.

## תשובות סופיות:

- א. כמותי בדיד.
  - ב. כמותי רציף.
  - ג. כמותי רציף.
  - ד. איכותי.
- א. מספר איחורים.
  - ב. כמותי בדיד.
- א. רציף.
  - ב. בדיד.
  - ג. בדיד.
  - ד. רציף.
  - ה. רציף.

# סטטיסטיקה והסתברות 01 88165

פרק 2 - סטטיסטיקה תיאורית- הצגה של נתונים

תוכן העניינים

1. כללי ..... 4

## סטטיסטיקה תיאורית – הצגה של נתונים:

### רקע:

דרכים להצגת נתונים שנאספו:

### רשימה של תצפיות:

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

### טבלת שכיחויות בדידה:

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

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

**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 תצפיות קטנות או שוות לערך.

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

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

**טבלת שכיחויות במחלקות:**

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

**דוגמה:**

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

מספר הילדים	זמן בדקות
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 איכותי. בדיאגרמת עוגה כל ערך במשתנה מקבל "נתח", שהוא פרופורציונלי לשכיחות היחסית של ערך המשתנה בנתונים.

### התפלגות המצב המשפחתי



### דיאגרמת מקלות:

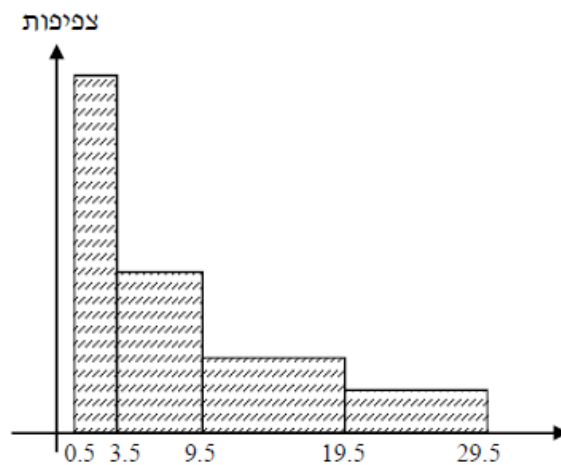
הציר האופקי הוא הציר של המשתנה והציר האנכי של השכיחות, כך שהגובה של המקל מעיד על השכיחות. רלבנטי למשתנה כמותי בדיד. לא נהוג להשתמש בתיאור למשתנה איכותי וכמו כן לא למשתנה כמותי רציף, וכן בסולמות מדידה עבור משתנה מסולם סדר.



### היסטוגרמה:

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

צפיפות	מצטברת	שכיחות	אמצע	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. נותן מראה חזותי לצורה של התפלגות המשתנה.

### צורות התפלגות נפוצות:

#### התפלגות סימטרית פעמונית

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

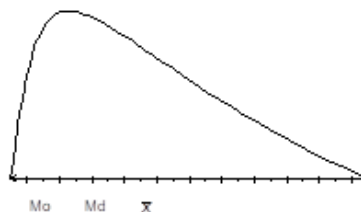


ישנן התפלגויות סימטריות שאינן פעמוניות, כגון:

#### התפלגות אסימטרית ימנית (חיובית)

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

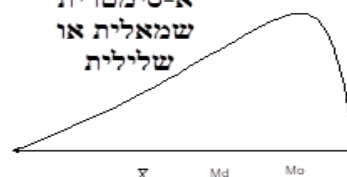
#### התפלגות א-סימטרית ימנית או חיובית



#### התפלגות אסימטרית שמאלית (שלילית)

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

#### התפלגות א-סימטרית שמאלית או שלילית



## שאלות:

- 1) בסקר צפייה בטלוויזיה התקבלו התוצאות הבאות: 25 צפו בערוץ הראשון, 25 צפו בערוץ 10, 75 צפו בערוץ השני, 50 צפו באחד מערוצי הכבלים ו-25 לא צפו בטלוויזיה בזמן הסקר.
- א. רשמו את טבלת השכיחות ואת השכיחות היחסית.
- ב. תארו את הנתונים באופן גרפי.

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

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

- א. מהו המשתנה הנחקר?
- ב. מהי פרופורציית התלמידים שמעדיפים תנ"ך?

- 3) להלן התפלגות ההשכלה במקום עבודה מסוים:

השכלה	מספר העובדים
נמוכה	60
תיכונית	120
אקדמאית	20

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

- 4) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
- 6, 5, 8, 7, 6, 7, 6, 8, 7, 8, 5, 4, 6, 10, 9, 8, 6, 7.
- א. מהו המשתנה? האם הוא בדיד או רציף?
- ב. תארו את הרשימה בטבלת שכיחויות.
- ג. הוסיפו שכיחויות יחסיות לטבלה.
- ד. תארו את הנתונים באופן גרפי.

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



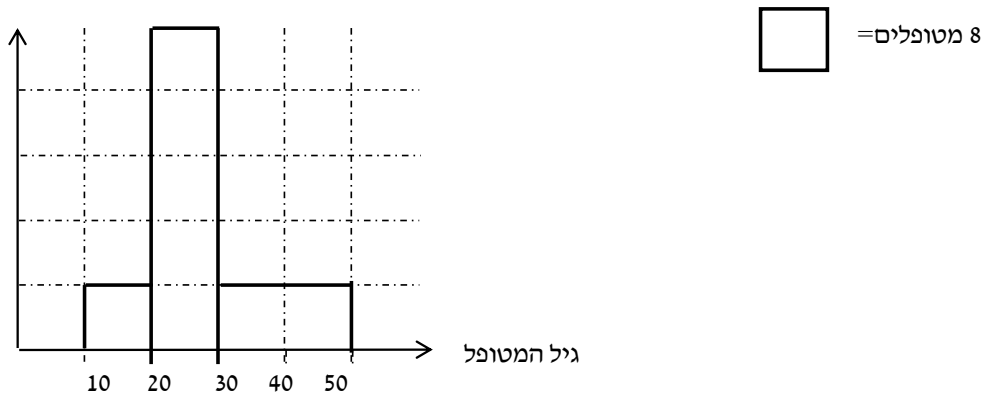
- מהו המשתנה הנחקר? האם הוא בדיד או רציף?
- תארו את הנתונים בטבלת שכיחויות במחלקות.
- הוסיפו שכיחות יחסית לטבלה.
- הוסיפו את הצפיפות של כל מחלקה לטבלה.
- מהי צורת ההתפלגות של הגבהים?

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

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

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

7) להלן גיל המטופלים של ד"ר שוורץ בשנים :  
 קנה מידה :



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

### 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 צפו
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.

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

ב+ג. להלן טבלה :

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

5) א. גובה בס"מ, רציף.

ב+ג+ד. להלן טבלה : ה. אסימטרית.

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

- 6) א. עיין גרף מלא בסרטון הוידאו.  
 ב. סימטרית.
- 7) א. המשתנה: גיל בשנים, משתנה רציף.  
 ב. המטופלים של ד"ר שוורץ.  
 ד. להלן טבלה:  
 ה. 62.5%.

$f(x)$	$x$
8	10-20
40	20-30
16	30-50

# סטטיסטיקה והסתברות 01 88165

פרק 3 - סטטיסטיקה תיאורית-גבולות מדומים ואמיתיים

תוכן העניינים

1. כללי ..... 15

## סטטיסטיקה תיאורית – גבולות מדומים וגבולות אמיתיים:

### רקע:

עבור משתנה רציף נהוג לתאר את הנתונים בטבלת שכיחויות במחלקות. הנתונים שנאספים הם ברמת דיוק מסוימת. לדוגמה: משקל של בני אדם ומשקל של יהלומים ישקלו ברמת דיוק שונה.

### גבולות מדומים:

כאשר גבול עליון של מחלקה אחת שונה מגבול תחתון של המחלקה הבאה אז הגבולות הם גבולות מדומים. כשהגבולות מדומים, ההפרש בין גבול תחתון של מחלקה לבין גבול עליון של המחלקה הקודמת יהיה רמת הדיוק.

**רמת הדיוק חייבת להיות קבועה** - אין אפשרות שחלק מהאנשים נדייק ברמה אחת ואת השאר ברמה אחרת. בגלל שהמשתנה הוא משתנה רציף, כשננתח את הנתונים נעבור מגבולות מדומים לגבולות אמיתיים. אם הנתונים יינתנו בגבולות מדומים נהפוך אותם תמיד לגבולות אמיתיים.

כיצד עוברים מגבולות מדומים לגבולות אמיתיים?

לוקחים את רמת הדיוק ומחלקים אותה ב-2, ואת התוצאה המתקבלת מוסיפים לגבולות העליונים ומפחיתים מהגבולות התחתונים. אם יתנו נתונים בגבולות מדומים אנחנו מוכרחים לעבור לגבולות אמיתיים על מנת להמשיך ולנתח, אך אם הנתונים כבר יינתנו בגבולות אמיתיים נשאיר אותם כמו שהם.

### דוגמה (פתרון בהקלטה):

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

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

## שאלות:

- (1) להלן התפלגות של משתנה בהצגה של מחלקות. יש להעביר את הנתונים לגבולות אמיתיים:

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

- (2) להלן התפלגות המשקלים בק"ג של קבוצת אנשים מסוימת. יש לרשום את הנתונים בגבולות אמיתיים:

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

## תשובות סופיות:

- (1) להלן טבלה:

$f(x)$	$x$
542	495-595
32	595-695
154	695-795
254	795-895

- (2) להלן טבלה:

$f(x)$	$x$
18	59.5-64.5
24	64.5-69.5
52	69.5-79.5
19	79.5-89.5

# סטטיסטיקה והסתברות 01 88165

פרק 4 - סטטיסטיקה תיאורית- סכימה

תוכן העניינים

1. כללי..... 17

## סטטיסטיקה תיאורית – סכימה:

רקע:

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

נסביר את צורת הרישום על ידי הדוגמה הבאה:

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

(הסבר מלא מופיע בסרטונים באתר).

## שאלות:

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

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

א.  $\sum_{i=1}^3 X_i$

ב.  $\sum_{i=1}^5 Y_i$

ג.  $\sum_{i=1}^4 X_i$

ד.  $\left(\sum_{i=1}^4 X_i\right)^2$

ה.  $\sum X_i$

ו.  $\sum X_i Y_i$

ז.  $\sum(X_i) \sum(Y_i)$

(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

$$\text{א. } \sum_{i=1}^4 y_i$$

$$\text{ב. } \sum_{i=1}^6 a$$

$$\text{ג. } \sum_{i=1}^6 x_i y_i$$

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

$$\text{ה. } \sum_{i=1}^6 x_i + a$$

(3) קבעו לכל זהות האם היא נכונה:

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

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

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

(4) נתון:  $\sum_{i=1}^{10} X_i = 80$ ,  $\sum_{i=1}^{10} X_i^2 = 1640$

$$\text{חשבו: } \sum_{i=1}^{10} (X_i - 4)^2$$

**תשובות סופיות:**

- |              |         |           |              |
|--------------|---------|-----------|--------------|
| ד. 121.      | ג. 11.  | ב. 9.     | א. 7. (1     |
|              | ז. 126. | ו. 27.    | ה. 14.       |
|              | ג. 7.   | ב. 12.    | א. 3. (2     |
|              |         | ה. 14.    | ד. 12.       |
| ג. לא נכונה. |         | ב. נכונה. | א. נכונה. (3 |
|              |         |           | .1160 (4     |

# סטטיסטיקה והסתברות 01 88165

פרק 5 - סטטיסטיקה תיאורית - מדדי מיקום מרכזי

תוכן העניינים

1. כללי ..... 21

## סטטיסטיקה תיאורית – מדדי מיקום מרכזי:

**רקע:**

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

**השכיח – Mode:**

השכיח הוא הערך הנפוץ ביותר בהתפלגות.

**ברשימה**

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

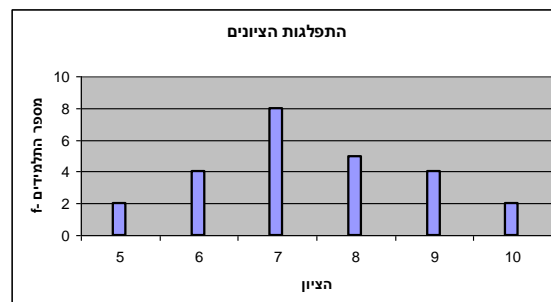
**בטבלת שכיחויות בדידה**

הערך שהשכיחות שלו היא הגבוהה ביותר.

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

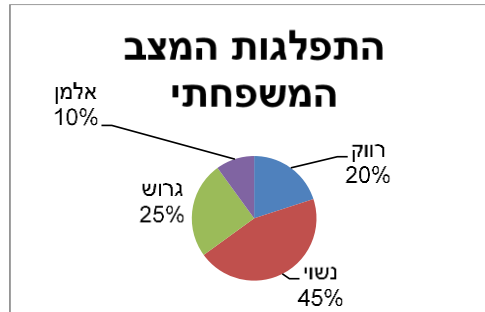
**בדיאגרמת מקלות**

שיעור ה-  $X$  של המקל הגבוה ביותר.



### בעוגה

הערך של הפלח הגדול ביותר.



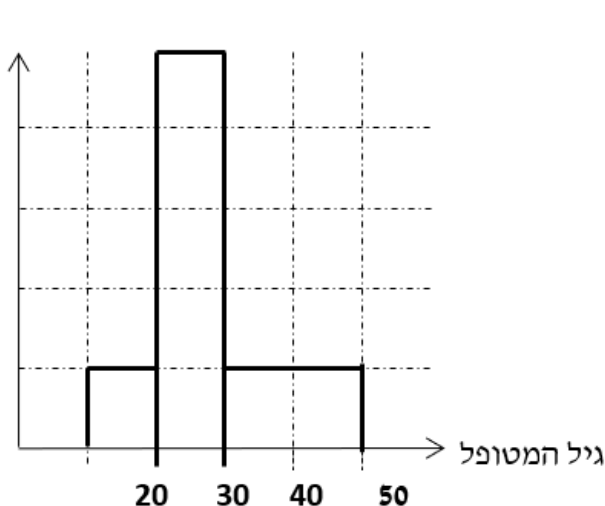
### בטבלת שכיחויות במחלקות

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

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

### בהיסטוגרמה

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



= 8

### כללי

יתכן שלהתפלגות יותר משכיח אחד.  
השכיח הוא מדד הרלבנטי לכל סוגי המשתנים.

### אמצע תחום (טווח) – Midrange:

הממוצע בין התצפית הגבוהה ביותר לתצפית הנמוכה ביותר:

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

### החציון – Median:

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

### ברשימה

נסדר את התצפיות בסדר עולה.

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

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

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

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

### בטבלת שכיחויות בדידה

נעשה תהליך דומה אך נעזר בשכיחות המצטברת.

### דיאגרמת מקלות

נמיר לטבלת שכיחויות בדידה במטרה למצוא את החציון.

### בטבלת שכיחויות במחלקות

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

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

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

$L_0$  - גבול התחתון של המחלקה.

$L_1$  - גבול העליון של המחלקה.

### היסטוגרמה

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

### כללי

החציון אינו רלבנטי למשתנה מסולם שמי ולא רלבנטי למשתנה איכותי.

### הממוצע – Average :

הממוצע הוא מרכז הכובד של ההתפלגות.

### ברשימה

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

### בטבלת שכיחויות

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

### במחלקות

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

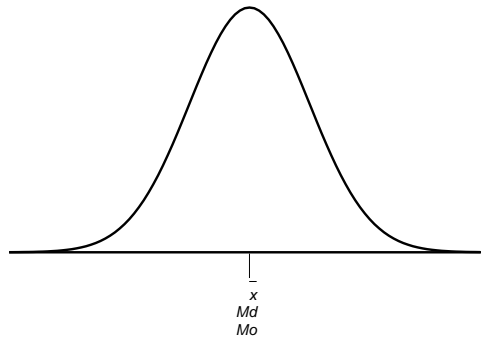
### כללי

הממוצע רלבנטי רק למשתנה כמותי.

### מדדי המיקום המרכזי בהתפלגויות המיוחדות:

בהתפלגות סימטרית פעמונית כל מדדי המרכז שווים זה לזה:

### התפלגות סימטרית

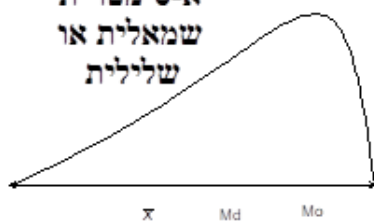


בהתפלגות סימטרית השכיח לא חייב להיות במרכז:

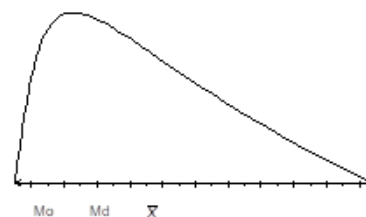
### התפלגות U



התפלגות  
א-סימטרית  
שמאלית או  
שלילית



התפלגות א-סימטרית  
ימנית או חיובית



## שאלות:

- (1) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:  
6, 5, 8, 7, 6, 6, 7, 8, 7, 5, 4, 6, 10, 9, 8, 6, 7.  
חשבו את החציון, השכיח, והממוצע של הציונים.
- (2) בדקו את מספר החדרים לדירה בבניין בן 5 דירות והתקבל ממוצע 3.8.  
לגבי 4 דירות נמצא מספר חדרים: 4, 3, 4, 5.  
א. כמה חדרים יש בדירה החמישית?  
ב. מהו השכיח ומהו החציון?
- (3) להלן התפלגות מספר מקלטי הטלוויזיה שנספרו עבור כל משפחה בישוב מסוים:

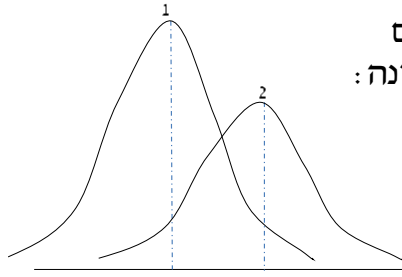
מספר משפחות	מספר מקלטים
22	0
28	1
18	2
22	3
10	4

- א. חשבו את הממוצע, החציון והשכיח של ההתפלגות.  
ב. הסבירו ללא חישוב כיצד כל מדד שחישבת בסעיף א' היה משתנה אם חלק מהמשפחות (לא כולן) שלא היה להם עד היום טלוויזיה היו רוכשים מקלט אחד.

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

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

- א. כמה משפחות יש בישוב?  
ב. מה אחוז המשפחות בישוב עם לכל היותר 2 מכוניות?  
ג. חשבו את הממוצע, החציון והשכיח.  
הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם!



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

- בכיתה 1 השכיח גבוה יותר מכיתה 2.
- בכיתה 2 השכיח גבוה יותר מכיתה 1.
- בשתי הכיתות אותו שכיח.
- לא ניתן לדעת באיזו כיתה השכיח גדול יותר.

6) ביישוב מסוים בדקו לכל משפחה את מספר הטלוויזיות שיש לה בבית. ביישוב גרות 200 משפחות. בממוצע יש למשפחה 1.5 טלוויזיות.

מספר טלוויזיות	מספר משפחות
0	28
1	62
2	
3	

- השלימו את הטבלה.
- מהו השכיח, אמצע טווח והחציון.
- חלק מהמשפחות להן הייתה טלוויזיה אחת בדיוק הוציאו את הטלוויזיה מביתם. כיצד כל מדד ישתנה (יגדל, יקטן או לא ישתנה). הסבירו ללא חישוב.

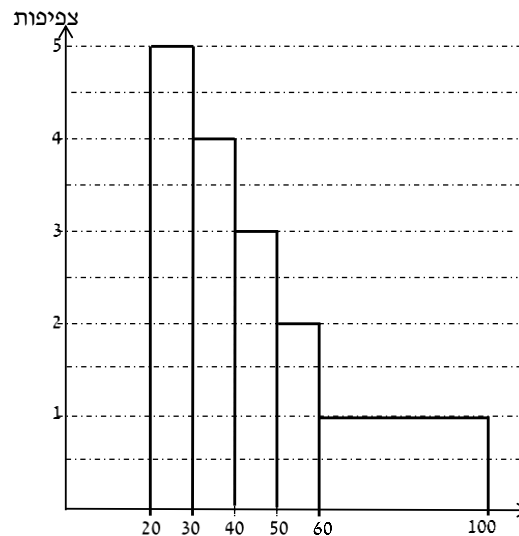
7) להלן התפלגות המשקל של קבוצה מסוימת בק"ג. מה הממוצע והחציון של ההתפלגות?

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

8) להלן התפלגות הגבהים בס"מ בקבוצה מסוימת. חשבו את הממוצע, החציון והשכיח של הגבהים בקבוצה זו.

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

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



- מצאו את השכיח בהתפלגות.
- מצאו את החציון בהתפלגות.
- הסבירו ללא חישוב האם הממוצע גדול/קטן/שווה לחציון.
- הסתבר שיש להוציא מספר תלמידים במחלקה בין 20-30 שקלים. כיצד הדבר ישפיע על הממוצע, החציון והשכיח? הסבירו ללא חישוב.

### תשובות סופיות:

- (1) חציון: 7, שכיח: 6, ממוצע: 6.9.
- (2) א. 3. ב. שכיח: 3.4, חציון: 4.
- (3) א. ממוצע: 1.7, חציון: 1.5, שכיח: 1.  
 ב. הממוצע יגדל ויתר המדדים לא ישתנו.
- (4) א. 630. ב. 34.13%. ג. שכיח וחציון: 3, ממוצע: 2.952.
- (5) ב'.
- (6) א. להלן טבלה: ב. חציון: 2, שכיח: 2, אמצע טווח: 1.5.

מספר משפחות	מספר טלויזיות
28	0
62	1
92	2
18	3

ג. שכיח: לא ישתנה, אמצע הטווח: לא ישתנה, חציון: לא ישתנה, ממוצע: יקטן.

- (7) חציון וממוצע: 55.
- (8) ממוצע: 172.6, חציון: 174.17, שכיח: 177.5.
- (9) א. 25. ב. 40. ג. גדול מהחציון.  
 ד. שכיח: לא ישתנה, חציון: יגדל, ממוצע: יגדל.

# סטטיסטיקה והסתברות 01 88165

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

תוכן העניינים

1. כללי..... 30

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

### רקע:

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

**הטווח / תחום (RANGE):**

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

### שונות וסטיית תקן:

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

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

### דוגמאות:

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

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

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

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

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

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

כשיש מחלקות נעזר באמצע המחלקה כדי לחשב את השונות.

## שאלות:

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

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

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

א. חשבו סטיית התקן.

ב. חשבו את הטווח של הנתונים.

הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם.

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

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

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

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

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

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

א. מה הממוצע של מספר החדרים לשכונה בדירה?

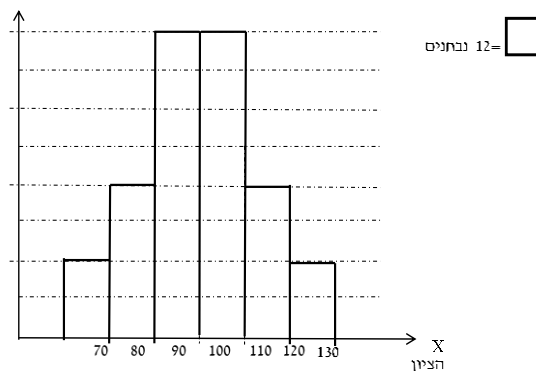
ב. חשבו את סטיית התקן של מספר החדרים לדירה.

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

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

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

7) להלן התפלגות הציונים במבחן אינטליגנציה:



- א. מה הממוצע ומה החציון של ההתפלגות?  
 ב. חשבו את סטיית התקן של הציונים.  
 ג. מסתבר שיש להוסיף 20 תצפיות לכל אחת משתי המחלקות 90-100 ו-100-110. כיצד הדבר ישתנה את כל אחד מהמדדים של הסעיפים הקודמים?

### תשובות סופיות:

- 1) שונות: 2.19, סטיית תקן: 1.48, טווח: 6.  
 2) א. סטיית תקן: 1.106. ב. טווח: 4.  
 3) א. ממוצע לא ישתנה, סטיית התקן תקטן.  
 ב. ממוצע לא ישתנה, סטיית התקן תגדל.  
 4) 10.8  
 5) א. 3.05. ב. 1.16. ג. ממוצע: יקטן, סטיית התקן: תגדל.  
 6) 7.73  
 7) א. 100. ב. 12.96. ג. ממוצע: לא ישתנה, סטיית תקן: תקטן.

# סטטיסטיקה והסתברות 01 88165

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

תוכן העניינים

1. כללי ..... 33

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

רקע:

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

שלבם במציאת טווח בין-רבעוני במחלקות:

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

שלב א:

נמצא את הרבעון התחתון (אחוזון 25) והרבעון העליון (האחוזון ה-75).

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

נוסחאות הרבעונים יהיו:

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

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

נציב:

$$Q_1 = 0.5 + \frac{\frac{200}{4} - 0}{56} \cdot 4 = 4.07 \text{ שניות}$$

$$Q_3 = 9.5 + \frac{\frac{3 \cdot 200}{4} - 106}{48} \cdot 2 = 11.33 \text{ שניות}$$

שלב ב:

$$IQR = Q_3 - Q_1 = 11.33 - 4.07 = 7.26 \text{ שניות}$$

נחסר את הרבעונים:

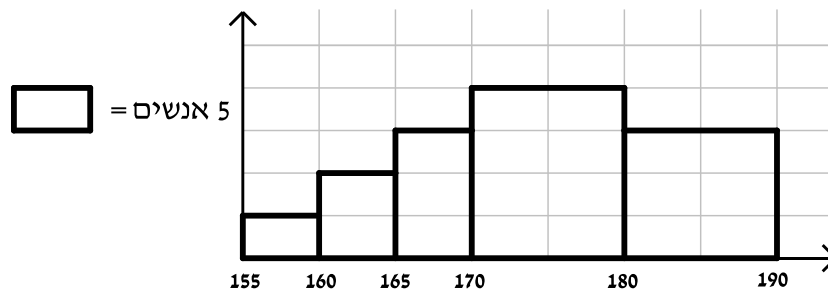
## שאלות:

(1) להלן התפלגות המשקל של קבוצה מסוימת בק"ג:

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

מצאו את הטווח הבין-רבעוני.

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



מצאו את הטווח הבין-רבעוני.

## תשובות סופיות:

(1) 13.75 ק"ג.

(2) 13.33 ק"ג.

# סטטיסטיקה והסתברות 01 88165

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

תוכן העניינים

1. כללי ..... 35

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

### רקע:

המטרה למדוד איך תצפית ממוקמת ביחס לשאר התצפיות בהתפלגות.

### ציון תקן:

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

- ציון התקן נותן כמה סטיות תקן סוטה התצפית מהממוצע. כלומר, ציון התקן מעיד על כמה סטיות תקן התצפית מעל או מתחת לממוצע:
- ציון תקן חיובי אומר שהתצפית מעל הממוצע.
  - ציון תקן שלילי אומר שהתצפית מתחת לממוצע.
  - ציון תקן אפס אומר שהתצפית בדיוק בממוצע.

### דוגמה (פתרון בהקלטה):

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

## שאלות:

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

המקצוע	ממוצע	סטיית תקן
לשון	74	12
מתמטיקה	80	16

עודד קיבל: 68 בלשון ו-70 במתמטיקה.

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

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

מספר פועלים	תפוקה	ממוצע
15	48	
2	10	סטיית תקן

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

- 3) הגובה הממוצע של המתגייסים לצבא הוא 175 סנטימטר עם סטיית תקן של 10 סנטימטר. המשקל הממוצע הוא 66 ק"ג עם סטיית תקן של 8 ק"ג. ערך התגייס כשגובהו 180 ס"מ ומשקלו 59 ק"ג.  
 א. במה ערך חריג יותר ביחס לשאר המתגייסים, גובהו או משקלו?  
 ב. כמה ערך אמור לשקול כדי שמשקלו יהיה שקול לגובהו?

## תשובות סופיות:

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

# סטטיסטיקה והסתברות 01 88165

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

תוכן העניינים

1. כללי ..... 37

## סטטיסטיקה תיאורית – מדדי מיקום יחסי – אחוזונים במחלקות:

### רקע:

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

### מציאת מאון במחלקות:

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

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

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

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

$L_0$  - גבול התחתון של המחלקה.

$L_1$  - גבול העליון של המחלקה.

אם נרצה לחלץ את אחוז התצפיות שמתחת לערך מסוים נשתמש בנוסחה

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

הבאה:

### דוגמה (פתרון בהקלטה):

להלן התפלגות השכר של עובדים בחברה מסוימת:

שכר ב-₪	
4000-6000	140
6000-10000	128
10000-15000	60
15000-20000	54
20000-40000	18

א. מצאו את המאון ה-40.

ב. מהו אחוז העובדים שמשתכרים מתחת ל-5,000 ₪?

## שאלות:

(1) להלן התפלגות השכר (באלפי שקלים) בחברה:

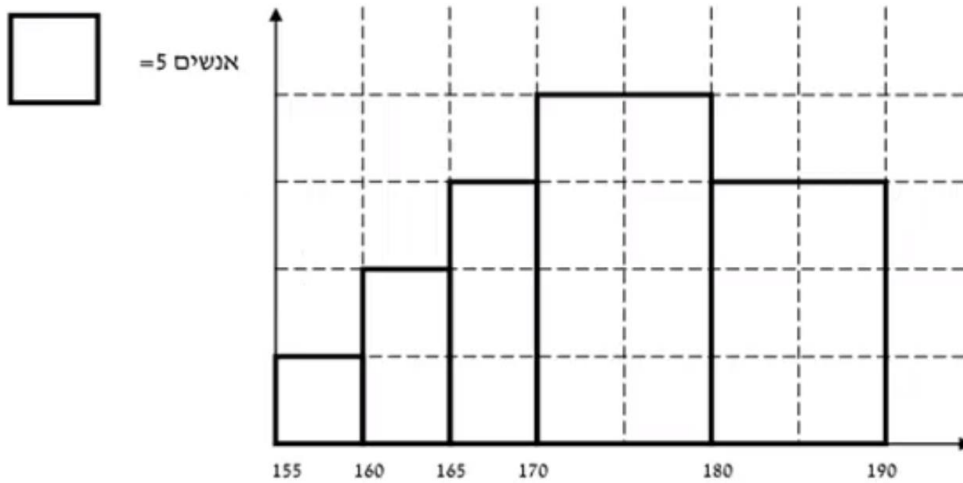
שכירות מצטברת	שכר - $X$
48	6-10
100	10-15
120	15-20
132	20-30
136	30-60

- א. חשבו את המאון ה-60.  
 ב. מהו העשירון העליון?  
 ג. 20% מהמשכורות הגבוהות ביותר הן משכורות של הבכירים, מהי המשכורת המינימאלית לבכיר?  
 ד. מה אחוז האנשים שמשכרם מתחת ל-7,000 ₪?  
 ה. איזה אחוז מהעובדים משכרם מעל ל-25,000 ₪?  
 ו. איזה אחוז מהעובדים משכרם בין 7,000 ₪ ל-25,000 ₪?

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

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

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



חשבו:

- העשירון התחתון.
- האחוזון ה-30.
- הגובה ש-20% מהתצפית גדולות ממנו.
- את אחוז התצפיות מתחת לגובה 158 ס"מ.
- את אחוז התצפיות מעל לגובה 185 ס"מ.
- את אחוז התצפיות בין גובה 170 ס"מ ל-185 ס"מ.

תשובות סופיות:

- 13.23
    - 22
    - 17.2
    - 8.82%
    - 7.36%
  - 83.82%
- להלן טבלה:

ציון- $X$	$f(x)$
50-60	40
60-70	60
70-80	200
80-90	60
90-100	40

- 162.5
  - 170
  - 183.33
  - 3%
  - 15%
- 55%

# סטטיסטיקה והסתברות 01 88165

פרק 10 - סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה

תוכן העניינים

1. כללי ..... 40

## סטטיסטיקה תיאורית – מדדי מיקום יחסי – אחוזונים בטבלה בדידה:

### רקע:

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

### חישוב האחוזון מתוך נתונים בטבלת שכיחויות בדידה:

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

### דוגמה (פתרון בהקלטה):

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

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

א. מצאו את האחוזון ה-25.

ב. מצאו את הערך ש-20% מהמקרים מעליו.

## שאלות:

(1) להלן התפלגות של משתנה כלשהו:

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

מצאו להתפלגות את:

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

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

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

חשבו את:

- א. העשירון התחתון.
- ב. האחוזון ה-30.
- ג. הערך ש-20% מהתצפית גדולות ממנו.
- ד. רבעון עליון.

## תשובות סופיות:

- (1) א. 2      ב. 1      ג. 3      ד. 1
- (2) א. 1      ב. 2      ג. 4      ד. 4

# סטטיסטיקה והסתברות 01 88165

פרק 11 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

תוכן העניינים

1. כללי ..... 42

## סטטיסטיקה תיאורית – טרנספורמציה לינארית:

### רקע:

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

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

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

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

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

**מדדי המרכז:**

$$R_y = |a| R_x$$

$$S_y = |a| S_x$$

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

**מדדי הפיזור:**

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

$$Z_y = \frac{a}{|a|} Z_x$$

**מדדי המיקום היחסי:**

### שלבי העבודה:

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

### דוגמה (פתרון בהקלטה):

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

## שאלות:

- (1) עבור סדרת נתונים התקבל:  $\bar{x} = 80, S = 15, MO = 70$ .  
הוחלט להכפיל את כל התצפיות ב-4 ולהחסיר מהתוצאה 5.  
חשבו את המדדים הללו לאחר השינוי.
- (2) בחברה מסוימת השכר הממוצע הוא 40 ש"ח לשעה עם סטיית תקן של 5 ש"ח לשעה.  
הוחלט להעלות את כל המשכורות ב-10%, אך זה לא סיפק את העובדים ולכן הם קיבלו לאחר מכן תוספת של 2 ש"ח לשעה.  
מה הממוצע ומהי השונות של השכר לשעה לאחר כל השינויים.
- (3) במבחן מסוים הציון החציוני היה 73, טווח הציונים היה 40 נקודות והעשירון העליון היה הציון 87. כיוון שהציונים בבחינה היו נמוכים, המורה החליט לתת פקטור של 4 נק' לכל התלמידים.  
חשבו את המדדים לאחר הפקטור.
- (4) דגמו מקו ייצור 50 קופסאות של גפרורים. בדקו בכל קופסא בה יש 40 גפרורים את כמות הגפרורים הפגומים. התקבל שבממוצע יש 3 גפרורים פגומים בקופסא, עם סטיית תקן של 1.5 גפרורים.  
מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקינים בקופסא?
- (5) חברת בזק הציעה את ההצעה הבאה: שלושים שקלים דמי מנוי חודשיים קבועים וכן 10 אגורות לכל דקה של שיחה יוצאת. אדם בדק במשך שנה את דקות השיחות היוצאות שלו, וקיבל שבממוצע חודשי יש לו 600 דקות שיחות יוצאות עם שונות של 2500 דקות רבועות, כמו כן בחודש ינואר ציון התקן היה 2.  
חשבו את המדדים הללו עבור חשבון הטלפון החודשי של אותו אדם בשקלים אם היה משתמש בחבילה המוצעת לו על ידי בזק.
- (6) הוכיחו שאם כל התצפיות בהתפלגות עברו טרנספורמציה לינארית:  $Y_i = a \cdot X_i + b$ , אזי הממוצע והשונות של כלל התצפיות לאחר הטרנספורמציה יהיו בהתאמה:  
$$\bar{y} = a \cdot \bar{x} + b, S_y^2 = a^2 S_x^2$$

**תשובות סופיות:**

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

# סטטיסטיקה והסתברות 01 88165

פרק 12 - סטטיסטיקה תיאורית - שאלות מסכמות

תוכן העניינים

1. כללי ..... 45

## סטטיסטיקה תיאורית – שאלות מסכמות:

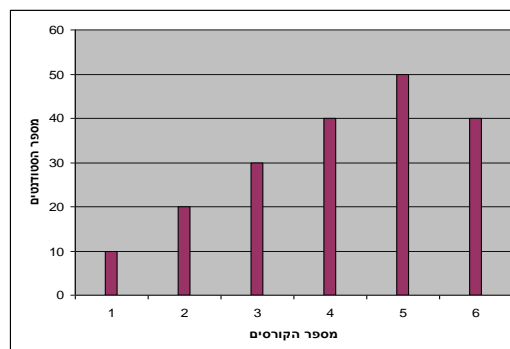
### שאלות:

1) בדקו עבור 5 תלמידים את המשקל שלהם:

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

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

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



- מה המשתנה הנחקר? בדיד או רציף?
- מהי צורת ההתפלגות?
- תארו את הנתונים בטבלת שכיחויות.
- חשבו את השכיח, החציון והטווח.

3) להלן התפלגות הציונים בבחינה בלשון שנעשתה עבור תלמידי כיתות ד'.

במחקר השתתפו 150 תלמידים. ממוצע הציונים שהתקבל:  $\bar{X} = 7\frac{1}{15}$ .

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

- השלימו את השכיחויות החסרות בטבלה.
- חשבו את הציון החציוני, השכיח.
- חשב שונות וסטיית תקן להתפלגות הציונים.

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

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

- מה אחוז האנשים עם רמת שביעות רצון נמוכה?
- מה המשתנה הנחקר ומאיזה סוג הוא?
- מהי הדרך הגרפית המתאימה ביותר לתיאור הנתונים?
  - היסטוגרמה.
  - דיאגרמת מקלות.
  - דיאגרמת עוגה.
- חשבו את המדדים הבאים:
  - טווח.
  - שכיח.
  - חציון.

5) להלן התפלגות מספר שעות העבודה לשבוע של העובדים (כ-200) בחברת "סטאר":

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

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

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

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

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

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

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

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

8) התפלגות ציוני מבחן אינטליגנציה היא סימטרית. נתון שהעשירון העליון הוא 130, הרבעון התחתון הוא 90, ושלמבחן נגשו 500 מועמדים.

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

- א. השלימו את הטבלה.
- ב. מהו הממוצע והחציון של ההתפלגות?
- ג. מהו הציון ש-40% מהתלמידים קיבלו מעליו? באיזה אחוזון מדובר?
- ד. הוחלט להעלות את כל הציונים ב-10 נקודות. כיצד הדבר ישפיע על הממוצע וסטיית התקן של הציונים?

- 9) להלן מספר טענות, עבור כל טענה ציינו אם היא נכונה או לא נכונה ונמקו.
- א. בסדרה שבה כל התצפיות שוות זו לזו השונות הינה 0.
  - ב. ציון התקן של החציון תמיד יהיה 0.
  - ג. ציון התקן של האחוזון ה-70 בהתפלגות אסימטרית ימנית (חיובית) תמיד יהיה חיובי.
  - ד. אם נוסיף תצפיות לסדרה של תצפיות, הדבר בהכרח יגדיל את הממוצע של הסדרה.
  - ה. בסדרה החציון הינו 80. הוספו שתי תצפיות אחת 79 ואחת 100 לכן החציון יגדל.
  - ו. אם נוסיף את הערך 4 לכל התצפיות אז סטיית התקן לא תשתנה.
  - ז. אם נחלק את כל התצפיות בהתפלגות ב-2 אז השונות תקטן פי 2.
  - ח. אם נגדיל את ממוצע המשכורות של עובדים בחברה אז גם השונות תגדל.

## תשובות סופיות:

- (1) א. המשתנה הנחקר: משקל תלמיד בק"ג, משתנה כמותי רציף.  
 ב.  $\bar{X} = 52$ ,  $Md = X_{\frac{n+1}{2}} = X_3 = 58$ , שכיח: 58.  
 ג.  $R = 28$ ,  $S = 10.12$ .  
 ד. הוא חריג יותר בגובה כי שם ציון התקן בערך מוחלט יותר גבוה.  
 ה. הממוצע לא ישתנה אך סטיית התקן תקטן.
- (2) א. מספר הקורסים, בדיד. ב. התפלגות אסימטרית שמאלית.  
 ג. להלן טבלה: ד. שכיח: 5, טווח: 5.

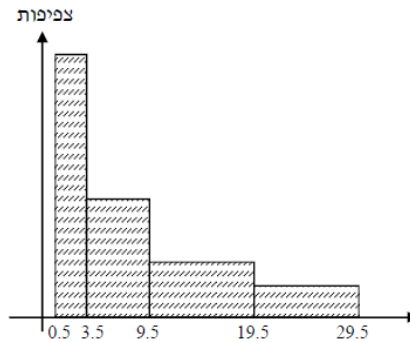
$f(x)$	$x$
10	1
20	2
30	3
40	4
50	5
40	6
190	סה"כ

- (3) א. 20 תלמידים קיבלו ציון 6 ו-40 תלמידים קיבלו ציון 8.  
 ב. חציון: 7, שכיח: 8. ג. שונות: 2.533, סטיית תקן: 1.592.
- (4) א. 20%. ב. שביעות רצון (סדר).  
 ג. ii. ד. טווח: 4, שכיח: 2, חציון: 2.
- (5) א. להלן טבלה: ב. חציון: 35, שכיח: 35, ממוצע: 35.

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

- ג. סטיית תקן: 12.65. ד. 53.333.  
 ה. 25%. ו. -0.395.  
 ז. חציון לא ישתנה, ממוצע יגדל, סטיית תקן תקטן.
- (6) א. 38%. ב. 40%. ג. 150. ד. חציון: 100.

7) א. שרטוט: ב. ההתפלגות היא א-סימטרית ימנית.



ג. שכיח: 2, חציון: 6.83.

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

8) א. ראו טבלה:

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

ב. 100. ג. 104.

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

9) א. נכון. ב. לא נכון. ג. לא נכון. ד. לא נכון. ה. לא נכון.  
ו. נכון. ז. לא נכון. ח. לא נכון.

# סטטיסטיקה והסתברות 01 88165

פרק 13 - סטטיסטיקה תיאורית - מקדם המתאם ( מדד קשר ) הלינארי

תוכן העניינים

1. מקדם המתאם הלינארי ( פירסון).....52
2. חישוב מקדם המתאם הלינארי (פירסון).....63

## מקדם המתאם (מדד קשר) הלינארי ומובהקותו

### מדד הקשר הלינארי (פירסון) – מבוא

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

### דוגמה (פתרון בהקלטה) :

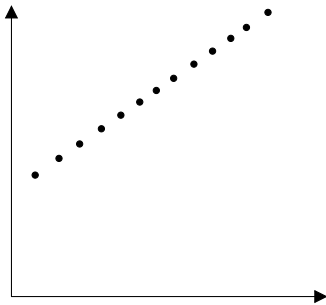
בבניין 8 דירות בדקו לכל דירה את מספר החדרים שלה וכמו כן את מספר הנפשות הגרות בדירה. להלן התוצאות שהתקבלו :

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

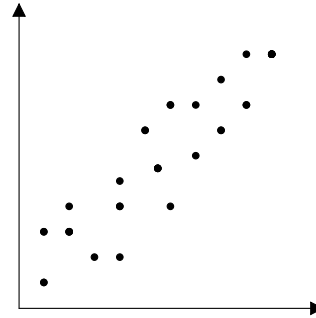
- (1) כמה תצפיות ישנן בדוגמה?
- (2) כמה משתנים ישנם בדוגמה, מי הם?
- (3) שרטטו לנתונים דיאגרמת פיזור.
- (4) מי המשתנה התלוי ומיהו המשתנה הבלתי תלוי?

## דיאגרמות פיזור לקשר בין משתנים וניתוחם

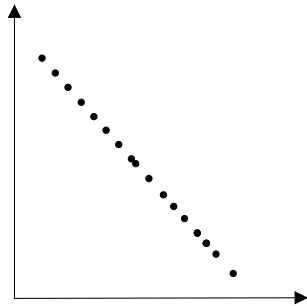
קשר לנארי חיובי מלא



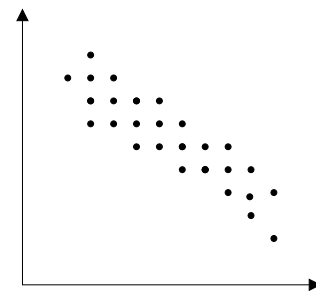
קשר לנארי חיובי חלקי



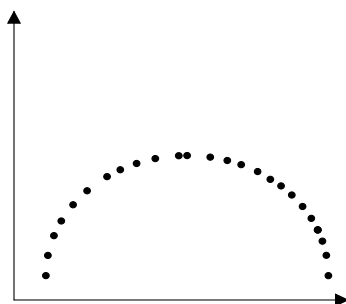
קשר לינארי שלילי מלא



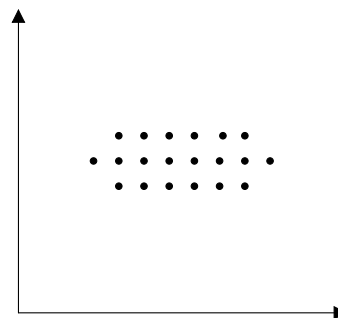
קשר לינארי שלילי חלקי



אין קשר לינארי



אין קשר



### משמעות מקדם המתאם:

כדי לבדוק עד כמה קיים קשר לינארי בין שני המשתנים ישנו מדד קשר שנקרא גם מקדם המתאם הלינארי הידוע גם בשם מקדם המתאם של פירסון. מקדם מתאם זה מקבל ערכים בין 1 ל-1.

-1

0

1

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

### מתאם חיובי מלא (מקדם מתאם 1):

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

### מתאם חיובי חלקי:

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

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

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

$r$  - מדד הקשר הלינארי במדגם שמהווה אומדן לפרמטר  $\rho$ .

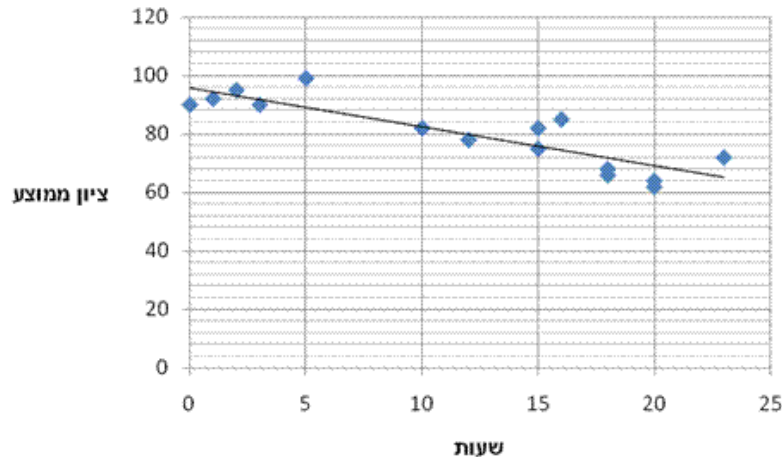
קיומו של מתאם בין שני משתנים אינו מצביע על סיבתיות בהכרח. למשל, אם נמצא מתאם חיובי בין כמות הסוכרזית שאדם אוכל לבין במשקל שלו אין זה אומר שהסיבה להשמנה היא הסוכרזית. מדד הקשר של פירסון הוא מדד קשר סימטרי, כלומר אם נחליף את  $X$  ב- $Y$  התוצאה תהיה זהה.

### דוגמה (פתרון בהקלטה):

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

## שאלות

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



- א. מיהו המשתנה הבלתי תלוי?  
 ב. מה ניתן לומר על כיוון הקשר בין מספר שעות הבילוי השבועיות לבין הציון הממוצע של הסמסטר? מה ניתן להגיד על עוצמת הקשר?
- 2) להלן טבלה המסכמת את מקדמי המתאם הלינארי בין ציוני מבחנים שונים שהתקבלו עבור תלמידים בכיתה מסוימת:

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

- א. השלימו את מקדמי המתאם שמסומנים בסימן שאלה בטבלה.  
 ב. בין אילו שני ציוני מקצועות שונים קיים מתאם בעל העוצמה החזקה ביותר?

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

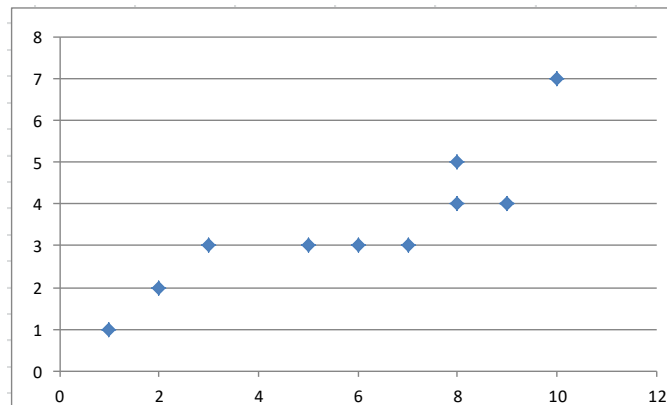
שעות תרגול	ציון סופי
20	90
25	90
30	95
15	60
30	90
20	85
10	50

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

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

- א. השלימו את המשפט:  
בחודש ינואר ככל שהטמפרטורה היומית נוטה לרדת, כך כמות המשקעים נוטה \_\_\_\_\_.
- ב. הוחלט להעביר את הטמפרטורה למעלות פרנהייט על מנת שיוכלו להשוות אותה לנתונים מארה"ב. נוסחת המעבר היא  $F^0 = 32 + \frac{9}{5} C^0$ .  
כיצד הדבר ישפיע על מקדם המתאם בין הטמפרטורה במעלות פרנהייט לכמות המשקעים במ"מ?

5) להלן דיאגרמת פיזור המראה קשר בין שני משנים:

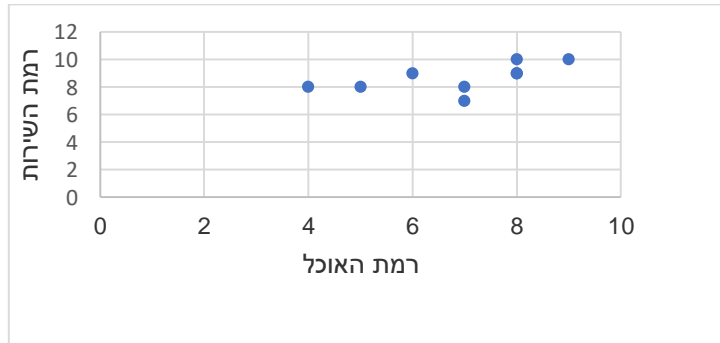


- א. השלימו: ניתן לראות שהקשר הוא לינארי \_\_\_\_\_ (מלאו חלקי) כיוון הקשר הוא (חיובי/שלילי).
- ב. השלימו: אם היינו מוסיפים תצפית שערך ה- $X$  שלה הוא 4 וערך ה- $Y$  שלה הוא 7, מקדם המתאם של פירסון היה \_\_\_\_\_ (גדלו קטן/לא משתנה).

**שאלות רב ברירה (יש לבחור את התשובה הנכונה):**

- 6) חוקר אקלים דגם כמה ימים בשנה ומדד את הטמפרטורה בטורונטו שבקנדה ואת הטמפרטורה בסידני שבאוסטרליה באותו היום. הוא חישב ומצא מקדם מתאם שלילי בין הטמפרטורה היומית בטורונטו לבין הטמפרטורה היומית בסידני. משמעות מקדם המתאם השלילי במדגם:
- א. אין קשר בין הטמפרטורה בטורונטו לבין הטמפרטורה בסידני בימים שנדגמו.  
ב. במדגם, רוב הטמפרטורות בטורונטו היו שליליות.  
ג. ההפרש בין הטמפרטורה בטורונטו לבין הטמפרטורה באוסטרליה, במדגם זה, הוא שלילי.  
ד. במדגם יש נטייה שהטמפרטורה יורדת בטורונטו לטמפרטורה לעלות בסידני.

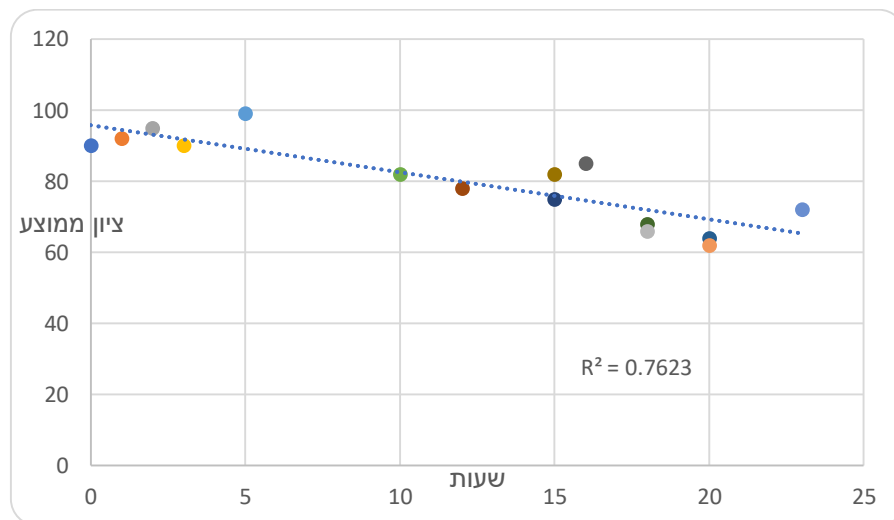
- 7) בסקר שביעות רצון שנערך בבית הקפה "פת לחם" התבקשו הלקוחות לדרג את מידת שביעות הרצון שלהם (בסולם 1-10) בשני נושאים: רמת האוכל ורמת השירות.



מה יהיה ערכו של מקדם המתאם ( $r$ )?

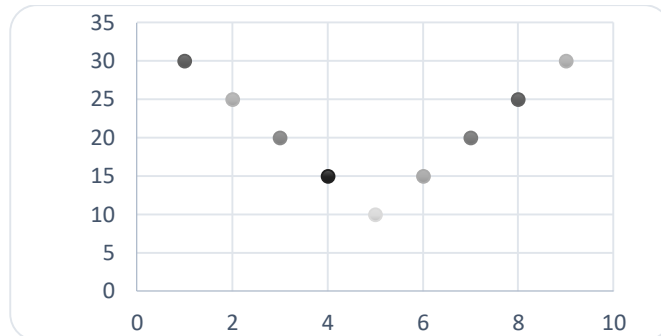
- א.  $r = -0.3$   
 ב.  $r = 0$   
 ג.  $r = 1.125$   
 ד.  $r = 0.593$

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



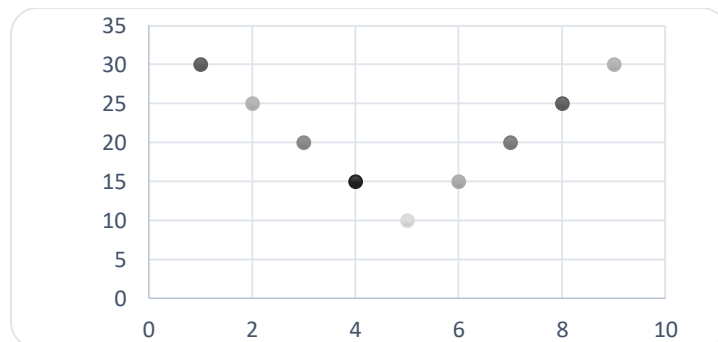
- מה ניתן לומר על כיוון הקשר במדגם בין מספר שעות הבילוי השבועיות לבין הציון הממוצע של הסמסטר?
- א. ככל שמבלים יותר הציון נוטה לרדת.  
 ב. אין קשר בין שעות הבילוי לציון.  
 ג. ככל שמבלים פחות הציון נוטה לרדת.  
 ד. ככל שהציון נוטה לרדת הסטודנט מבלה פחות.

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



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

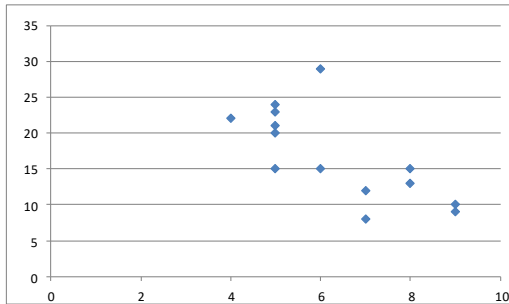
10) התרשים הבא מתאר דיאגרמת פיזור.



איזו טענה נכונה?

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

בגרף הבא מתוארת דיאגרמת פיזור של שני משתנים:



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

במדגם התקבל  $r^2 = 0.52$ .

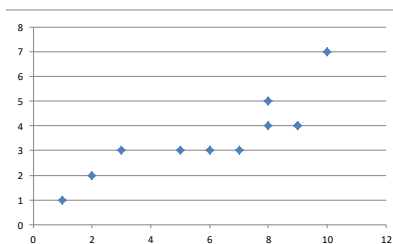
11) לאור הנתונים המופיעים בדיאגרמה, איזה מבין הערכים הבאים מתאים להיות התוצאה של  $r$ ?

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

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

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

13) להלן דיאגרמת פיזור:



מה יהיה מקדם המתאם בין שני המשתנים?

- א. 1
- ב. 0.85
- ג. 0.15
- ד. 0

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

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

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

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

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

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

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

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

- א. -1
- ב. 1
- ג. -0.5
- ד. 0.5

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

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

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

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

- 19) נמצא מתאם חיובי בעוצמה גבוהה בין  $X$  – ציון בבגרות בלשון ל  $Y$  – ציון בבגרות במתמטיקה. אילו מהמשפטים הבאים נכון?
- א. ניתן לומר שאחת מהסיבות להבדלים שיש לסטודנטים במתמטיקה נובעים מההבדלים שיש להם בלשון.
- ב. קיימת נוסחה של קו ישר שקושרת בין ציון בבגרות במתמטיקה לציון בבגרות בלשון.
- ג. ללא יוצא מן הכלל, ניתן להגיד שכל תלמיד שמצליח יותר מתלמיד אחר בלשון גם יצליח יותר מאותו תלמיד במתמטיקה.
- ד. אף אחד מהטענות שהוצגו אינה בהכרח נכונה.

- 20) עבור סדרה של תצפיות מדדו את  $X$  ואת  $Y$ . נמצא שעבור כל התצפיות שהערך של  $Y$  ירד הערך של  $X$  בהכרח ירד ללא יוצא מן הכלל. מקדם המתאם של פירסון יהיה בהכרח:
- א. 1
- ב. -1
- ג. 0
- ד. אף אחת מהתשובות.

## תשובות סופיות

- (1) א. שעות בילוי.  
 ב. הקשר חלקי, כיוון הקשר שלילי.  
 (2) א. להלן טבלה:

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

- (3) א. ב"ת- מס' שעות התרגול, תלוי- ציון.  
 ג. קשר לינארי חיובי חלקי.  
 (4) א. לעלות.  
 (5) א. חלקי, חיובי.  
 (6) ד'. (7) ד'. (8) א'. (9) ג'. (10) א'.  
 (11) ג'. (12) ד'. (13) ב'. (14) א'. (15) א'.  
 (16) א'. (17) ד'. (18) א'. (19) ד'. (20) ד'.

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

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

**דוגמה:**

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

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

**דוגמה:**

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

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

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

**שלב שני:** מחשבים את מקדם המתאם (מדד הקשר) שבודק עד כמה קיים קשר לינארי בין שני המשתנים. המדד (ניקרא גם מדד הקשר של פירסון) מכמת את מה שניראה בשלב הראשון רק בעין.

המדד בודק את כיוון הקשר (חיובי או שלילי) ואת עוצמת הקשר (חלש עד חזק). מקדם מתאם זה מקבל ערכים בין -1 ל-1.  
 מקדם מתאם -1 או 1 אומר שקיים קשר לינארי מוחלט ומלא בין המשתנים שניתן לבטאו על ידי הנוסחה:  $y = bx + a$ .

**מתאם חיובי מלא (מקדם מתאם 1):**

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

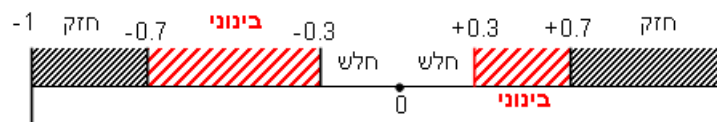
### מתאם חיובי חלקי:

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

### מתאם שלילי חלקי:

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

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



מקדם המתאם יסומן באות  $r$ .

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

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

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

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

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

## שאלות

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

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

- א. שרטטו דיאגרמת פיזור לנתונים. מה ניתן להסיק מהדיאגרמה על טיב הקשר בין מספר החיסורים של תלמיד לציונו? מיהו המשתנה הבלתי תלוי ומיהו המשתנה התלוי?
- ב. חשבו את מדד הקשר של פירסון. האם התוצאה מתיישבת עם תשובתך לסעיף א'?
- ג. הסבירו, ללא חישוב, כיצד מקדם המתאם היה משתנה אם היה מתווסף תלמיד שהחסיר 4 פעמים וקיבל ציון 80?

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

- 2) במחקר רפואי רצו לבדוק האם קיים קשר בין רמת ההורמון X בדם החולה לרמת ההורמון Y שלו. לצורך כך מדדו את רמת ההורמונים ההלו עבור חמישה חולים. להלן התוצאות שהתקבלו:
- א. מה הממוצע של כל רמת הורמון?
- ב. מהו מקדם המתאם בין ההורמונים? ומה משמעות התוצאה?

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

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

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

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

- א. חשב את מדד הקשר הלינארי בין X ל-Y. מיהו המשתנה התלוי?
- ב. מה המשמעות של התוצאה שקיבלת בסעיף א'?

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

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

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

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

חשבו את מדד הקשר הלינארי בין  $X$  ל- $Y$ .

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

- (6) להלן רשימת טענות, לגבי כל טענה קבעו נכון/לא נכון ונמקו.
- מתווך דירות המיר מחירי דירות מדולר לשקל. נניח שדולר אחד הוא 3.5₪. אם מתווך הדירות יחשב את מדד הקשר של פירסון בין מחיר הדירה בשקלים למחיר הדירה בדולרים הוא יקבל 1.
  - לסדרה של נתונים התקבל  $\bar{X} = \bar{Y} = 6$ ,  $S_x = S_y = 1$ . לכן, מדד הקשר של פירסון יהיה 1.
  - אם השונות המשותפת של  $X$  ושל  $Y$  הינה 0 אז בהכרח גם מקדם המתאם של פירסון יהיה 0.

### שאלות רב-ברירה:

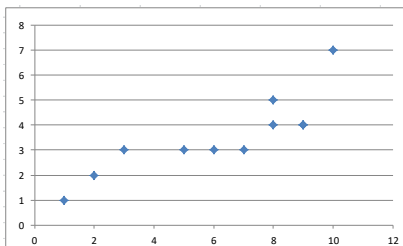
- (7) נמצא שקיים מקדם מתאם שלילי בין הציון בעברית לציון בחשבון בבחינה לכן:
- הדבר מעיד שהציונים בכיתה היו שליליים.
  - ככל שהציון של תלמיד יורד בחשבון יש לו נטייה לרדת בעברית.
  - ככל שהציון של תלמיד עולה בחשבון יש לו נטייה לרדת בעברית.
  - אף אחת מהתשובות לא נכונה.

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

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

9) להלן דיאגרמת פיזור:

מה יהיה מקדם המתאם בין שני המשתנים?



- א. 1  
 ב. 0.85  
 ג. 0.15  
 ד. 0

## תשובות סופיות

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

# סטטיסטיקה והסתברות 01 88165

פרק 14 - יסודות ההסתברות

תוכן העניינים

1. כללי ..... 68

## הגדרות יסודיות:

### רקע:

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

**מרחב מדגם:** כלל התוצאות האפשריות בניסוי המקרי. לדוגמה, בהטלת קובייה:  $\{1, 2, 3, 4, 5, 6\}$ , או: מזג האוויר בעוד שבועיים:  $\{\text{נאה, שרבי, מושלג, גשום, מעונן חלקית, אביד}\}$ .

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

**גודל מרחב המדגם:** מספר התוצאות האפשריות במרחב המדגם. בהטלת קובייה למשל נקבל:  $|\Omega| = 6$ .

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

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

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

**הסתברות במרחב מדגם אחיד:** במרחב מדגם אחיד הסיכוי למאורע יהיה:  $P(A) = \frac{|A|}{|\Omega|}$ .

דוגמה: מה הסיכוי בהטלת קובייה לקבל לפחות 5?  $P(A) = \frac{|A|}{|\Omega|} = \frac{2}{6}$

דוגמה: מה הסיכוי בהטלת 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:

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

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

מה ההסתברות שתלמיד אקראי שניבחר בכיתה יכשל?  $\frac{f}{n} = \frac{2}{25} = 0.08$

**הסתברות למאורע משלים:** הסתברות לקבוצת המשלים של המאורע ביחס למרחב המדגם:  $P(\bar{A}) = 1 - P(A)$ . למשל, בדוגמה הקודמת הסיכוי לעבור את הבחינה יכול

להיות מחושב לפי הסיכוי להיכשל:  $P(\bar{A}) = 1 - \frac{2}{25} = \frac{23}{25}$ .

## 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 לאירועים הבאים:  
 i. סכום התוצאות 7.  
 ii. מכפלת התוצאות 12.  
 ג. חשבו את הסיכויים לאירועים 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 מקלטי הטלוויזיה עבור כל משפחה בישוב מסוים:

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

- נבחרה משפחה באקראי מהישוב.  
 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":

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

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

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

**תשובות סופיות:**

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

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

ג.  $\bar{A} = \{FF, FG, GF, GG\}$

2) א.  $\Omega = \left\{ \begin{matrix} (1,1) & (2,1) & (3,1) & (5,1) & (4,1) & (6,1) \\ (1,2) & (2,2) & (3,2) & (4,2) & (5,2) & (6,2) \\ (1,3) & (2,3) & (3,3) & (4,3) & (5,3) & (6,3) \\ (1,4) & (2,4) & (3,4) & (4,4) & (5,4) & (6,4) \\ (1,5) & (2,5) & (3,5) & (4,5) & (5,5) & (6,5) \\ (1,6) & (2,6) & (3,6) & (4,6) & (5,6) & (6,6) \end{matrix} \right\}$

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

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

3) א. 0.4 ב. 0.4 ג. 0.5

4) א. 0.22 ב. 0.78 ג. 0.32

5) א. 0.1 ב. 0.2 ג. 0.8

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

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

ג.  $\bar{D} = \{EEE\}$

ד.  $\frac{1}{8}$

# סטטיסטיקה והסתברות 01 88165

פרק 15 - פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים

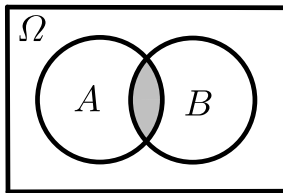
תוכן העניינים

1. כללי ..... 72

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

**רקע:**

**פעולת חיתוך:**

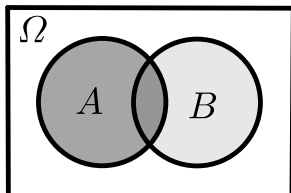


נותנת את המשותף בין המאורעות הנחתכים.  
 חיתוך בין המאורע  $A$  למאורע  $B$  יסומן כך:  $A \cap B$ .  
 מדובר בתוצאות שנמצאות ב- $A$  וגם ב- $B$ .

דוגמה:

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

**פעולת איחוד:**



נותנת את כל האפשרויות שנמצאות לפחות באחת מהמאורעות, ומסומנת:  $A \cup B$ .  
 הפעולה נותנת את אשר נמצא ב- $A$  או ב- $B$ .  
 כלומר, לפחות אחד מהמאורעות קורה.

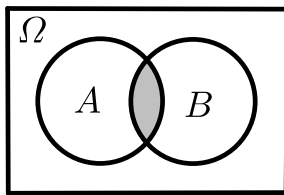
דוגמה:

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

דוגמה (הפתרון נמצא בהקלטה):

סטודנט ניגש בסמסטר לשני מבחנים. מבחן בסטטיסטיקה ומבחן בכלכלה. ההסתברות שלו לעבור את המבחן בסטטיסטיקה הוא 0.9, ההסתברות שלו לעבור את המבחן בכלכלה הוא 0.8 וההסתברות לעבור את המבחן בסטטיסטיקה ובכלכלה היא 0.75. מה ההסתברות שלו לעבור את המבחן בסטטיסטיקה בלבד? מה ההסתברות שלו להיכשל בשני המבחנים? מה ההסתברות לעבור 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:**

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

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

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

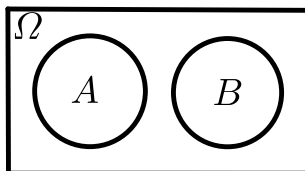
$$P(A \cup B) = 1 - P(\bar{A} \cap \bar{B})$$

**שיטת ריבוע הקסם:**

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

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

**מאורעות זרים:**



מאורעות זרים הם כאשר אין להם אף איבר משותף:  $A \cap B = \{ \}$ . כלומר, הם לא יכולים להתרחש בו זמנית.

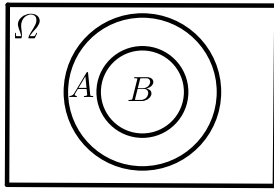
ההסתברות של חיתוך המאורעות היא אפס:  $P(A \cap B) = 0$ .

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

דוגמה:

בהטלת קובייה, האפשרויות לקבל לפחות 5 הן:  $A = \{5, 6\}$  והאפשרות לקבל 3

היא:  $B = \{3\}$ , ולכן החיתוך ביניהם הוא אפס, כלומר:  $A \cap B = \{ \}$ .

**מאורעות מוכלים:**


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 בתוך מאורע  $A$ .

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

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

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

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

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

למשל:

## שאלות:

- (1) מהאותיות  $E, F$  ו- $G$  יוצרים מילה בת 2 אותיות – לא בהכרח בת משמעות. נגדיר 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 ומבחן בסטטיסטיקה. נגדיר את המאורעות הבאים:
- $A$  - לעבור את המבחן בסטטיסטיקה.
  - $B$  - לעבור את המבחן בכלכלה.
- היעזרו בפעולות חיתוך, איחוד ומשלים בלבד כדי להגדיר את המאורעות הבאים וסמנום בדיאגרמת וון את השטח המתאים:
- א. התלמיד עבר רק את המבחן בכלכלה.
  - ב. התלמיד עבר רק את המבחן בסטטיסטיקה.
  - ג. התלמיד עבר את שני המבחנים.
  - ד. התלמיד עבר לפחות מבחן אחד.
  - ה. התלמיד נכשל בשני המבחנים.
  - ו. התלמיד נכשל בכלכלה.
- (3) נתבקשתם לבחור ספרה באקראי. נגדיר את  $A$  להיות הספרה שנבחרה היא זוגית. נגדיר את  $B$  להיות הספרה שנבחרה קטנה מ-5.
- א. רשמו את כל התוצאות למאורעות הבאים:  
 $A \cup B, A \cap B, \bar{B}, B, A$
  - ב. חשבו את ההסתברויות לכל המאורעות מהסעיף הקודם.
- (4) נסמן ב- $\Omega$  את מרחב המדגם וב- $\phi$  קבוצה ריקה. נתון כי  $A$  הינו מאורע בתוך מרחב המדגם. להלן מוגדרים מאורעות שפתרונם 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 :

א.  $A \cap B$

ב.  $A \cup B$

ג.  $\bar{A} \cap B$

ד.  $\bar{A} \cup \bar{B}$

ה.  $\bar{\bar{A}}$

(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 בדיוק (מהשפות הנ"ל).

(7) שתי מפלגות רצות לכנסת הבאה. מפלגת "גדר" תעבור את אחוז החסימה בהסתברות של 0.08 ומפלגת "עמיד" תעבור את אחוז החסימה בהסתברות של 0.20. בהסתברות של 76% שתי המפלגות לא תעבורנה את אחוז החסימה.

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

ב. מה ההסתברות ששתי המפלגות תעבורנה את אחוז החסימה?

ג. מה ההסתברות שרק מפלגות "עמיד" תעבור את אחוז החסימה?

(8) במקום עבודה מסוים 40% מהעובדים הם גברים. כמו כן, 20% מהעובדים הם אקדמאים. 10% מהעובדים הינן נשים אקדמאיות.

א. איזה אחוז מהעובדים הם גברים אקדמאיים?

ב. איזה אחוז מהעובדים הם גברים או אקדמאיים?

ג. איזה אחוז מהעובדים הם נשים לא אקדמאיות?

9) הסיכוי של מניה A לעלות הנו 0.5 ביום מסוים והסיכוי של מניה B לעלות ביום מסוים הנו 0.4. בסיכוי של 0.7 לפחות אחת מהמניות תעלה ביום מסוים. חשבו את ההסתברויות הבאות לגבי שתי המניות הללו ביום מסוים:

א. ששתי המניות תעלנה.

ב. שאף אחת מהמניות לא תעלנה.

ג. שמניה A בלבד תעלה.

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

A - בקובייה האדומה התקבלה התוצאה 4 ובשחורה 2.

B - סכום התוצאות משתי הקוביות הוא 6.

C - מכפלת התוצאות בשתי הקוביות 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) עבור 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?

ב. חשבו את  $P(\bar{A} \cap B)$ .

12) מטבע הוטל פעמיים. נגדיר את המאורעות הבאים:

A - קיבלנו עץ בהטלה הראשונה.

B - קיבלנו לפחות עץ אחד בשתי ההטלות.

איזו טענה נכונה?

א. A ו-B מאורעות זרים.

ב. A ו-B מאורעות משלימים.

ג. B מכיל את A.

ד. A מכיל את B.

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

שאר הכרטיסים ריקים. אדם קיבל כרטיס אקראי.

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

ב. מה ההסתברות לא לזכות בפרס?

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 שרק  $A$  יקרה או שרק  $B$  יקרה?

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

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

16 קבעו אילו מהטענות הבאות נכונות:

- א.  $A \cap B = B \cap A$ .  
 ב.  $\overline{A \cup B} = A \cap B$ .  
 ג.  $A \cap B \cap C = A \cap B \cap (C \cup B)$ .  
 ד.  $\overline{A \cap B \cap C} = \bar{A} \cup \bar{B} \cup \bar{C}$ .

17 נתון ש- $A$  ו- $B$  מאורעות במרחב מדגם. נתון ש- $P(A) = 0.3$ ,  $P(B) = 0.2$

- א. האם יתכן ש- $P(A \cup B) = 0.4$ ?  
 ב. האם יתכן ש- $P(A \cup B) = 0.6$ ?  
 ג. אם  $A$  ו- $B$  זרים מה הסיכוי  $P(A \cup B)$ ?  
 ד. אם  $A$  מכיל את  $B$  מה הסיכוי  $P(A \cup B)$ ?

18 מתוך אזרחי המדינה הבוגרים ל-30% חשבון בבנק הפועלים. ל-28% חשבון בבנק לאומי ול-15% חשבון בבנק מזרחי. כמו כן נתון כי 6% מחזיקים חשבון בבנק לאומי ובבנק הפועלים. ל-5% חשבון בבנק פועלים ומזרחי. ול-4% חשבון בבנק לאומי ומזרחי. כמו כן ל-1% מהאוכלוסייה הבוגרת חשבון בנק בשלושת הבנקים יחד.

- א. מה אחוז האזרחים להם חשבון בבנק לאומי בלבד?  
 ב. מה ההסתברות שאזרח כלשהו יחזיק חשבון בבנק פועלים ולאומי אבל לא בבנק מזרחי?  
 ג. מה ההסתברות שלאזרח יהיה חשבון בפועלים או במזרחי אבל לא בבנק לאומי?  
 ד. מה אחוז האזרחים שיש להם חשבון בנק אחד בלבד?  
 ה. מה אחוז האזרחים שיש להם בדיוק חשבון בשני בנקים בלבד?  
 ו. מה ההסתברות שלאזרח בוגר אין חשבון בנק באף אחד מהבנקים הללו?  
 ז. לאיזה אחוז מהאזרחים יש חשבון בנק בלפחות אחד מהבנקים הללו?

- 19** חברה מסוימת פרסמה את הנתונים הבאים לגבי האזרחים מעל גיל 21. הנתונים שהתקבלו היו: 40% מהאנשים מחזיקים כרטיס "ויזה", 52% מחזיקים כרטיס "ישראלכרט", 20% מחזיקים כרטיס "אמריקן אקספרס", 15% מחזיקים כרטיס ויזה וגם ישראלכרט, 8% מחזיקים כרטיס ישראלכרט וגם אמריקן אקספרס ו-7% מחזיקים כרטיס ויזה וגם אמריקן אקספרס. כמו כן, 13% לא מחזיקים באף אחד משלושת הכרטיסים הנ"ל.
- א. מה אחוז מחזיקי שלושת כרטיס האשראי גם יחד?
- ב. מה אחוז מחזיקי ישראלכרט וויזה אך לא את אמריקן אקספרס?
- ג. מה אחוז מחזיקי כרטיס אחד בלבד?

**20** הוכיחו:  $P(\bar{A} \cap \bar{B}) = 1 - P(A) - P(B) + P(A \cap B)$ .

- 21**  $A$  ו- $B$  מאורעות במרחב המדגם. האם נכון לומר שהסיכוי שיתרחש בדיוק מאורע אחד הוא:  $P(A) + P(B) - 2P(A \cap B)$ ?

## תשובות סופיות:

- (1) א.  $A \cap B = \{EG, EF, FE, GE\}$   
 ב.  $A \cup B = \{EG, EF, EE, FE, GE, EG, GF\}$
- (2) א.  $B \cap \bar{A}$  ב.  $A \cap \bar{B}$  ג.  $A \cap B$  ד.  $A \cup B$  ה.  $\bar{A} \cap \bar{B}$  ו.  $\bar{B}$
- (3) א.  $A = \{0, 2, 4, 6, 8\}$ ,  $B = \{0, 1, 2, 3, 4\}$ ,  $\bar{B} = \{5, 6, 7, 8, 9\}$   
 $A \cap B = \{0, 2, 4\}$ ,  $A \cup B = \{0, 2, 4, 6, 8, 1, 3\}$
- ב.  $P(A \cup B) = 0.7$ ,  $P(A \cap B) = 0.3$ ,  $P(\bar{B}) = 0.5$ ,  $P(B) = 0.5$ ,  $P(A) = 0.5$
- (4)  $\bar{\bar{A}} = A$ ,  $A \cap \phi = \phi$ ,  $A \cup \phi = A$ ,  $A \cap \Omega = A$ ,  $A \cup \Omega = \Omega$   
 $A \cap \bar{A} = \phi$ ,  $\bar{\phi} = \Omega$ ,  $A \cup \bar{A} = \Omega$
- (5) א.  $A \cap B$ : גובה בין 1.7 ל-1.8.  
 ב.  $A \cup B$ : כל גובה אפשרי.  
 ג.  $\bar{A} = \bar{A} \cap B$ : גובה לכלל היותר 1.7.  
 ד.  $\bar{A} \cup \bar{B}$ : לכלל היותר 1.7 או לפחות 1.8.  
 ה.  $A = \bar{\bar{A}}$ : גובה מעל 1.7.
- (6) א.  $A \cap B \cap C$  ב.  $A \cap \bar{B} \cap \bar{C}$  ג.  $A \cup B \cup C$   
 ד.  $\bar{C}$  ה.  $(A \cap B \cap \bar{C}) \cup (B \cap C \cap \bar{A}) \cup (A \cap C \cap \bar{B})$
- (7) א.  $P(A \cup B) = 0.24$  ב.  $P(A \cap B) = 0.04$  ג.  $P(B \cap \bar{A}) = 0.16$
- (8) א.  $P(A \cap B) = 10\%$  ב.  $P(A \cup B) = 50\%$  ג.  $P(\bar{A} \cap \bar{B}) = 50\%$
- (9) א.  $P(A \cap B) = 0.2$  ב.  $P(\bar{A} \cap \bar{B}) = 0.3$  ג.  $P(A \cup \bar{B}) = 0.3$
- (10) א. לא. ב. כן. ג. כן. ד. לא.
- (11) א. כן. ב.  $P(\bar{A} \cap B) = 0.3$
- (12) הטענה הנכונה היא ג'.
- (13) א. 0.05. ב. 0.95.
- (14) א.  $P(A \cap B) = 0.06$  ב. לא. ג.  $P((A \cap \bar{B}) \cup (B \cap \bar{A})) = 0.43$
- (15)  $P(B) = \frac{1}{5}$ ,  $P(A) = \frac{2}{5}$
- (16) א. נכון. ב. לא נכון. ג. לא נכון. ד. נכון.
- (17) א. כן. ב. לא. ג.  $P(A \cup B) = 0.5$  ד.  $P(A \cup B) = 0.3$
- (18) א. 19%. ב. 0.05. ג. 0.31. ד. 46%. ה. 12%. ו. 0.41.
- (19) א. 5%. ב. 10%. ג. 67%.
- (20) שאלת הוכחה.
- (21) נכון.

# סטטיסטיקה והסתברות 01 88165

פרק 16 - קומבינטוריקה - כלל המכפלה

תוכן העניינים

1. כללי ..... 81

## קומבינטוריקה – כלל המכפלה:

**רקע:**

**כלל המכפלה:**

כלל המכפלה הוא כלל שבאמצעותו אפשר לחשב את גודל המאורע או גודל מרחב המדגם.

אם לתהליך יש  $k$  שלבים:  $n_1$  אפשרויות לשלב הראשון,  $n_2$  אפשרויות לשלב השני...  $n_k$

אפשרויות לשלב  $k$ :

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

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

$$n_1 = 6, n_2 = 2$$

$$n_1 \cdot n_2 = 6 \cdot 2 = 12$$

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

$$n_1 = 26, n_2 = 10, n_3 = 10, n_4 = 10, n_5 = 10$$

$$n_1 \cdot n_2 \cdot n_3 \cdot n_4 \cdot n_5 = 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 260,000$$

## שאלות:

- (1) חשבו את מספר האפשרויות לתהליכים הבאים:
- הטלת קובייה פעמים.
  - מספר תלת ספרתי.
  - בחירת בן ובת מכתה שיש בה שבעה בנים ועשר בנות.
  - חלוקת שני פרסים שונים לעשרה אנשים שונים כאשר אדם לא יכול לקבל יותר מפרס אחד.
- (2) במסעדה מציעים ארוחה עסקית.
- בארוחה עסקית יש לבחור מנה ראשונה, מנה עיקרית ושתייה. האופציות למנה ראשונה הן: סלט ירקות, סלט אנטיפסטי ומרק היום. האופציות למנה עיקרית הן: סטייק אנטריקוט, חזה עוף בגריל, לזניה בשרית ולזניה צמחונית. האופציות לשתייה הן: קפה, תה ולימונדה.
- כמה ארוחות שונות ניתן להרכיב בעזרת התפריט הזה?
  - אדם מזמין ארוחה אקראית. חשב את ההסתברויות הבאות:
    - בארוחה סלט ירקות, לזניה בשרית ולימונדה.
    - בארוחה סלט, לזניה ותה.
- (3) בוחרים באקראי מספר בין חמש ספרות. חשבו את ההסתברויות הבאות:
- המספר הוא זוגי.
  - במספר כל הספרות שונות.
  - במספר כל הספרות זהות.
  - במספר לפחות שתי ספרות שונות.
  - במספר לפחות שתי ספות זהות.
  - המספר הוא פלינדרום (מספר הנקרא מימין ומשמאל באות הצורה).
- (4) חישה אנשים אקראיים נכנסו למעלית בבניין בן 8 קומות. חשבו את ההסתברויות הבאות:
- כולם ירו בקומה החמישית.
  - כולם ירדו באותה קומה.
  - כולם ירדו בקומה אחרת.
  - ערן ודני ירדו בקומה השישית והיתר בשאר הקומות.

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

## תשובות סופיות:

- (1) א. 0.36    ב. 0.900    ג. 0.70    ד. 0.90
- (2) א. 0.36    ב. i.  $\frac{1}{36}$     ב. ii.  $\frac{1}{9}$
- (3) א. 0.5    ב. 0.3024    ג. 0.0001    ד. 0.9999    ה. 0.6976    ו. 0.01
- (4) א.  $\frac{1}{8^5}$     ב.  $\frac{1}{8^4}$     ג. 0.205    ד.  $\frac{1 \cdot 1 \cdot 7^3}{8^5}$
- (5) א. 0.3375    ב. 0.2730
- (6) א.  $\frac{1}{216}$     ב.  $\frac{5}{18}$     ג.  $\frac{13}{18}$     ד.  $\frac{215}{216}$
- (7) א.  $\frac{23^5}{26^5}$     ב.  $\frac{1}{26^4}$     ג.  $1 - \frac{1}{26^4}$     ד.  $\frac{1}{26^2}$
- (8) א.  $0.9^a$     ב.  $1 - 0.9^a$     ג.  $0.5^a$
- (9)  $2^n$

# סטטיסטיקה והסתברות 01 88165

פרק 17 - קומבינטוריקה - תמורה - סידור עצמים בשורה

תוכן העניינים

1. כללי ..... 85

## קומבינטוריקה – תמורה – סידור עצמים בשורה:

**רקע:**

**תמורה:**

מספר האפשרויות לסדר  $n$  עצמים שונים בשורה:  $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$ .

הערה:  $0! = 1$ .

דוגמאות (הפתרונות בהקלטה):

- בכמה דרכים שונות ניתן לסדר את האותיות:  $a, b, c, d$ ?
- בכמה דרכים שונות ניתן לסדר את האותיות:  $a, b, c, d$ , כך שהאותיות:  $a, b$  יהיו ברצף?
- בכמה דרכים שונות ניתן לסדר את האותיות:  $a, b, c, d$ , כך שהאותיות:  $a, b$  יופיעו בתור הרצף  $ba$ ?

## שאלות:

- (1) חשבו : בכמה אופנים  
 א. אפשר לסדר 4 ספרים שונים על מדף?  
 ב. אפשר לסדר חמישה חיילים בטור?
- (2) סידרו באקראי 10 דיסקים שונים על מדף שמתוכם שניים בשפה העברית.  
 א. מה ההסתברות שהדיסקים בעברית יהיו צמודים זה לזה?  
 ב. מה ההסתברות שהדיסקים בעברית לא יהיו צמודים זה לזה?  
 ג. מה ההסתברות ששני הדיסקים בעברית יהיו כל אחד בקצה השני של המדף?
- (3) בוחנים 5 בנים ו-4 בנות בכיתה ומדרגים אותם לפי הציון שלהם בבחינה. נניח שאין תלמידים בעלי אותו ציון.  
 א. מהו מספר הדירוגים האפשריים?  
 ב. מהו מספר הדירוגים האפשריים אם מדרגים בנים ובנות בנפרד?
- (4) מסדרים 10 ספרים שונים על מדף.  
 א. בכמה אופנים ניתן לסדר את הספרים על המדף?  
 ב. שני ספרים מתוך ה-10 הם בסטיסטיקה.  
 א. מה ההסתברות שאם נסדר את הספרים באקראי, הספרים בסטיסטיקה יהיו צמודים זה לזה?  
 ב. מה ההסתברות שהספרים בסטיסטיקה לא יהיו צמודים זה לזה?  
 ג. מה ההסתברות שהספרים בסטיסטיקה יהיו בקצות המדף (כל ספר בקצה אחר)?
- (5) אדם יצר בנגן שלו פלייליסט (רשימת השמעה) של 12 שירים שונים. 4 בשפה העברית, 5 באנגלית ו-3 בצרפתית. האדם הריץ את הפלייליסט באקראי.  
 א. מה ההסתברות שכל השירים באנגלית יופיעו כשירים הראשונים כמקשה אחת?  
 ב. מה ההסתברות שכל השירים באנגלית יופיעו ברצף (לא חובה ראשונים)?  
 ג. מה ההסתברות ששירים באותה השפה יופיעו ברצף (כלומר כל השירים באנגלית ברצף, כל השירים בעברית ברצף וכך גם השירים בצרפתית)?

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

### תשובות סופיות:

- 1) א. 0.24      ב. 0.120
- 2) א. 0.2      ב. 0.8      ג. 0.022
- 3) א. 0.362880      ב. 0.2880
- 4) א. 0.3628800      ב. 0.2      ג. 0.8      ד.  $\frac{1}{45}$
- 5) א.  $\frac{1}{792}$       ב.  $\frac{1}{99}$       ג.  $\frac{1}{4620}$
- 6) א. 0.75      ב. 0.014      ג.  $\frac{1}{14}$       ד.  $\frac{1}{35}$

# סטטיסטיקה והסתברות 01 88165

פרק 18 - קומבינטוריקה - תמורה עם עצמים זהים

תוכן העניינים

1. כללי ..... 88

## קומבינטוריקה – תמורה עם עצמים זהים:

**רקע:**

**תמורה עם חזרות:**

אם יש בין העצמים שיש לסדר עצמים זהים, יש לבטל את הסידור הפנימי שלהם על ידי חלוקה בסידורים הפנימיים שלהם.

מספר האופנים לסדר  $n$  עצמים בשורה, ש- $n_1$  מהם זהים מסוג 1,  $n_2$  זהים מסוג 2

$$r\text{-}n_1 \text{ זהים מסוג } r : \frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_r!}$$

דוגמה (תשובה בהקלטה):

כמה מילים ניתן ליצור מכל האותיות הבאות: W, W, T, T, K, K

## שאלות:

(1) במשחק יש לצבוע שתי משבצות מתוך המשבצות הבאות:

--	--	--	--	--

בכמה דרכים שונות ניתן לבצע את הצביעה?

(2) בכמה אופנים שונים אפשר לסדר בשורה את האותיות: ב, ע, ב, ע, ג?

(3) בבית נורות מקום ל-6 נורות. בחרו שתי נורות אדומות, שתי נורות צהובות ושתי נורות כחולות. כמה דרכים שונות יש לסדר את הנורות?

(4) נרצה ליצור מספר מכל הספרות הבאות: 1, 2, 2, 2, 6. כמה מספרים כאלה אפשר ליצור?

(5) במשחק בול פגיעה יש 10 משבצות, אדם צובע 4 משבצות מתוך ה-10. המשתתף השני צריך לנחש אילו 4 משבצות נצבעו. מה ההסתברות שבניחוש אחד יהיה בול פגיעה?

(6) כמה אותות שונים, שכל אחד מורכב מ-10 דגלים שונים, ניתן ליצור, אם 4 דגלים הם לבנים, 3 כחולים, 2 אדומים ואחד שחור. דגלים שווי צבע זהים זה לזה לחלוטין.

## תשובות סופיות:

(1) 10.

(2) 60.

(3) 90.

(4) 20.

(5)  $\frac{1}{210}$ .

(6) 12600.

# סטטיסטיקה והסתברות 01 88165

פרק 19 - קומבינטוריקה - סידור עצמים במעגל

תוכן העניינים

1. סידור עצמים במעגל.....90

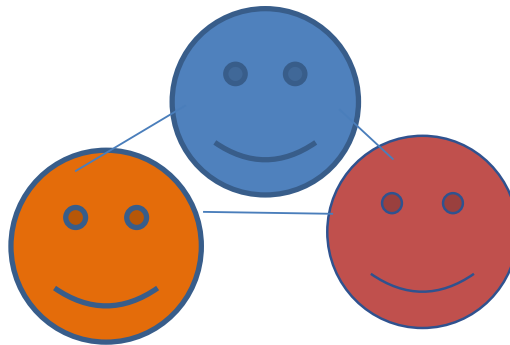
## קומבינטוריקה – סידור עצמים במעגל:

**רקע:**

מספר האפשרויות לסדר  $n$  עצמים שונים במעגל בו אין מקומות מסומנים הוא:  $(n-1)!$ .

דוגמה (פתרון בהקלטה):

דנה, רמה ושדה רוצות ליצור מעגל ריקוד. בכמה דרכים שונות הן יכולות להחזיק אחת לשנייה את הידיים, כדי ליצור את המעגל?



## שאלות:

- (1) מעצב פנים יצר ללקחותיו מניפת צבעים המוצגת במעגל. במניפה 12 צבעים שונים מתוכם 3 בגווני אפור, 3 בגווני לבן, 3 בגווני ירוק ו-3 בגווני צהוב. כמה מניפות שונות ניתן ליצור כאשר:
- גווני האפור צמודים זה לזה.
  - צבעים באותו גוון צמודים זה לזה.



- (2) דני יוצר שרשרת חרוזים הבנויה מעשרה חרוזים בצבעים שונים. הוא משחיל את עשרת החרוזים באקראי. חשבו את ההסתברויות הבאות:
- הסידור יהיה בדיוק כמוראה בציר.
  - החרוז הלבן והכתום יהיו בסמוך זה לזה.

- (3) אבא הכין עוגת יומולדת עגולה. הוא סידר 7 נרות כמוראה בשרטוט. הנרות זהים ונבדלים זה בזה בצבע: 2 כחולים זהים, 2 אדומים זהים, 2 צהובים זהים ו-1 כתום. סידור הנרות נעשה באקראי. חשבו את ההסתברויות הבאות:



- הנרות הצהובים סמוכים זה לזה.
- נרות באותו צבע סמוכים זה לזה.

- (4)  $n$  בנים ו- $n$  בנות הסתדרו במעגל באקראי.



- מה הסיכוי שכל הבנים יסתדרו זה לצד זה בלי להתפצל?
- מה הסיכוי שכל הבנים יסתדרו זה לצד זה בלי להתפצל וגם כל הבנות יסתדרו זו לצד זו בלי להתפצל?
- מה הסיכוי שהסידור יהיה שמימין ומשמאל לכל בן תהיה בת?

## תשובות סופיות:

$$(1) \quad \text{א. } .2177280 \quad \text{ב. } .7776$$

$$(2) \quad \text{א. } \frac{1}{9!} \quad \text{ב. } \frac{2}{9}$$

$$(3) \quad \text{א. } \frac{1}{3} \quad \text{ב. } \frac{1}{15}$$

$$(4) \quad \text{א. } \frac{(n!)^2}{(2n-1)!} \quad \text{ב. } \frac{(n!)^2}{(2n-1)!} \quad \text{ג. } \frac{(n-1)!(n!)}{(2n-1)!}$$

# סטטיסטיקה והסתברות 01 88165

פרק 20 - קומבינטוריקה - דגימה סידורית ללא החזרה ועם החזרה

תוכן העניינים

1. כללי ..... 93

## קומבינטוריקה – דגימה סידורית ללא החזרה ועם החזרה:

**רקע:**

**מדגם סידור בדגימה עם החזרה:**

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

**דוגמה:**

בוחרים שלושה תלמידים מתוך עשרה לייצג ועד בו תפקידים שונים, תלמיד יכול למלא יותר מתפקיד אחד.

כמה ועדים שונים ניתן להרכיב?  $10^3 = 1,000$ ,  $k = 3$ ,  $n = 10$ .

**מדגם סידור ללא החזרה:**

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

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

**דוגמה:**

שלושה תלמידים נבחרים מתוך 10 לייצג וועד בו תפקידים שונים.

תלמיד לא יכול למלא יותר מתפקיד אחד:  $\frac{10!}{7!} = 10 \cdot 9 \cdot 8 = 720$ .

## שאלות:

- (1) במפלגה 20 חברי כנסת, מעוניינים לבחור שלושה חברי כנסת לשלושה תפקידים שונים.
- א. חבר כנסת יכול למלא יותר מתפקיד אחד.  
כמה קומבינציות ישנן לחלוקת התפקידים?
- ב. חבר כנסת לא יכול למלא יותר מתפקיד אחד.  
כמה קומבינציות יש לחלוקת התפקידים?
- (2) במשחק מזל יש 4 משבצות ממוספרות מ-A-D (A עד D). בכל משבצת יש למלא סיפרה (0-9). הזוכה הוא זה שניחש נכונה את כל הספרות בכל המשבצות בהתאמה.
- א. מה ההסתברות לזכות במשחק?  
ב. מה ההסתברות שבאף משבצת לא תהיה את הספרה 3 במספר הזוכה?  
ג. מה ההסתברות שהתוצאה 4 תופיע לפחות פעם אחת במספר הזוכה?
- (3) קבוצה מונה 22 אנשים, מה ההסתברות שלפחות לשניים מהם יהיה יום הולדת באותו התאריך?
- (4) שלושה אנשים קבעו להיפגש במלון הילטון בסינגפור. הבעיה היא שבסינגפור ישנם 5 מלונות הילטון.
- א. מה ההסתברות שכל השלושה ייפגשו?  
ב. מה ההסתברות שכל אחד יגיע לבית מלון אחר?
- (5) בכיתה 40 תלמידים. מעוניינים לבחור חמישה מהם לוועד כיתה. בכמה דרכים ניתן להרכיב את הוועד אם:
- א. בוועד 5 תפקידים שונים ותלמיד יכול למלא יותר מתפקיד אחד.  
ב. בוועד 5 תפקידים שונים ותלמיד לא יכול למלא יותר מתפקיד אחד.

## תשובות סופיות:

- (1) א. 8000      ב. 6840
- (2) א. 0.0001      ב. 0.6561      ג. 0.3439
- (3) 0.476
- (4) א. 0.04      ב. 0.48
- (5) א.  $40^5$       ב. 78,960,960

# סטטיסטיקה והסתברות 01 88165

פרק 21 - קומבינטוריקה - דגימה ללא סדר וללא החזרה

תוכן העניינים

95 ..... 1. כללי

## קומבינטוריקה – דגימה ללא סדר וללא החזרה:

רקע:

מדגם לא סדור בדגימה ללא החזרה:

מספר האפשרויות לדגום  $k$  עצמים שונים מתוך  $n$  עצמים שונים כאשר אין

משמעות לסדר העצמים הנדגמים ואין החזרה:  $\frac{n!}{(n-k)!k!} = \binom{n}{k} = \frac{(n)_k}{k!}$

דוגמה:

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

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

הערות:

$$\binom{n}{k} = \binom{n}{n-k} \quad (1)$$

$$\binom{n}{n-1} = \binom{n}{1} = n \quad (2)$$

$$\binom{n}{n} = \binom{n}{0} = 1 \quad (3)$$

## שאלות:

- (1) בכיתה 15 בנות ו-10 בנים. יש לבחור 5 תלמידים שונים מהכיתה לנציגות הכיתה. בכמה דרכים אפשר להרכיב את הנציגות, אם:
- אין שום הגבלה לבחירה.
  - מעוניינים ש-3 בנות ו-2 בנים ירכיבו את המשלחת.
  - לא יהיו בנים במשלחת.
- (2) סטודנט מעוניין לבחור 5 קורסי בחירה בסמסטר זה. לפניו רשימה של 10 קורסים לבחירה: 5 במדעי הרוח, 3 במדעי החברה, 2 במתמטיקה.
- כמה בחירות שונות הוא יכול ליצור לעצמו?
  - כמה בחירות יש לו בהן 3 קורסים הם ממדעי הרוח?
  - כמה בחירות יש לו אם 2 מהן לא ממדעי הרוח?
  - כמה בחירות יש לו אם 2 ממדעי הרוח, 2 ממדעי החברה ו-1 ממתמטיקה?
- (3) בכיתה 30 תלמידים מתוכם 12 תלמידים ו-18 תלמידות.
- יש לבחור למשלחת 4 תלמידים מהכיתה. התלמידים נבחרים באקראי.
- מה ההסתברות שהמשלחת תורכב רק מבנות?
  - מה ההסתברות שבמשלחת תהיה רק בת אחת?
  - מה ההסתברות שבמשלחת תהיה לפחות בת אחת?
- (4) במשחק הלוטו יש לבחור 5 מספרים מתוך 45. המספרים הם 1-45.
- מה ההסתברות שבמשחק הזוכה כל המספרים הם זוגיים?
  - מה ההסתברות שבמספר הזוכה יש לכל היותר מספר זוגי אחד?
  - מה ההסתברות שבמספר הזוכה לפחות פעם אחת יש מספר זוגי?
  - מה ההסתברות שבמספר הזוכה כל המספרים גדולים מ-30?
- (5) בחפיסת קלפים ישנם 52 קלפים: 13 בצבע שחור בצורת עלה, 13 בצבע אדום בצורת לב, 13 בצבע אדום בצורת יהלום ו-13 בצבע שחור בצורת תלתן. מכל צורה (מתוך ה-4) יש 9 קלפים שמספרם 10-2, שאר הקלפים הם; נסיך, מלכה, מלך ואס (בעצם מדובר בקופסת קלפים רגילה ללא ג'וקר).
- שני אנשים משחקים פוקר. כל אחד מקבל באקראי 5 קלפים (ללא החזרה).
- מה ההסתברות שעודד יקבל את כל המלכים וערן את כל המלכות?
  - מה ההסתברות שאחד השחקנים יקבל את הקלף אס-לב?
  - מה ההסתברות שערן יקבל קלפים שחורים בלבד ועודד יקבל שני קלפים שחורים בדיוק?
  - מה ההסתברות שערן יקבל לפחות 3 קלפים שהם מספר (אס אינו מספר)?

- 6 במכללה 4 מסלולי לימוד. בכל מסלול לימוד 5 מזכירות. יש ליצור וועד של 5 מזכירות מתוך כלל המזכירות במכללה. יוצרים וועד באופן אקראי. חשבו את ההסתברויות הבאות:
- א. כל המזכירות בוועד יהיו ממסלול "מדעי ההתנהגות".  
 ב. כל המזכירות בוועד יהיו מאותו המסלול.  
 ג. מכל מסלול תבחר לפחות מזכירה אחת.

$$7 \quad \text{הוכיחו כי: } \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

- 8  $2n$  בנים ו- $2n$  בנות מתחלקים ל-2 קבוצות.
- א. בכמה דרכים שונות ניתן לבצע את החלוקה אם שתי הקבוצות צריכות להיות שוות בגודלן ויש בכל קבוצה מספר שווה של בנים ובנות?  
 ב. בכמה דרכים ניתן לבצע את החלוקה אם יש מספר שווה של בנים ובנות בכל קבוצה אבל הקבוצות לא בהכרח בגודל שווה.

### תשובות סופיות:

- 1 א. 53130      ב. 20475      ג. 3003
- 2 א. 252      ב. 100      ג. 100      ד. 60
- 3 א. 0.1117      ב. 0.1445      ג. 0.9819
- 4 א. 0.02      ב. 0.187      ג. 0.972      ד. 0.00246
- 5 א. 0      ב. 0.1923      ג. 0.009      ד. 0.837
- 6 א.  $6.45 \cdot 10^{-5}$       ב.  $2.58 \cdot 10^{-4}$       ג. 0.3225
- 7 שאלת הוכחה.

$$8 \quad \text{א. } \binom{2n}{n}^2 \quad \text{ב. } \sum_{i=1}^n \binom{2n}{i}^2$$

# סטטיסטיקה והסתברות 01 88165

פרק 22 - קומבינטוריקה - שאלות מסכמות

תוכן העניינים

98 ..... 1. כללי

## קומבינטוריקה – שאלות מסכמות:

### שאלות:

- (1) בכיתה 40 תלמידים. מעוניינים לבחור חמישה מהם לוועד כיתה. בכמה דרכים ניתן להרכיב את הוועד אם:
- בוועד 5 תפקידים שונים ותלמיד יכול למלא יותר מתפקיד אחד.
  - בוועד 5 תפקידים שונים ותלמיד לא יכול למלא יותר מתפקיד אחד.
  - אין תפקידים שונים בוועד.
- (2) במשרד 30 עובדים, יש לבחור ארבעה עובדים למשלחת לחו"ל. בכמה דרכים ניתן להרכיב את המשלחת?
- במשלחת ארבע משימות שונות שיש למלא וכל עובד יכול למלא יותר ממשימה אחת.
  - כמו בסעיף א' רק הפעם עובד לא יכול למלא יותר ממשימה אחת.
  - מעוניינים לבחור ארבעה עובדים שונים למשלחת שבה לכולם אותו התפקיד.
- (3) מעוניינים להרכיב קוד סודי. הקוד מורכב מ-2 ספרות שונות ו-3 אותיות שונות באנגלית (26 אותיות אפשריות).
- כמה קודים שונים ניתן להרכיב?
  - כמה קודים שונים ניתן להרכיב אם הקוד מתחיל בספרה ונגמר בספרה?
  - כמה קודים ניתן להרכיב אם הספרות חייבות להיות צמודות זו לזו?
  - בכמה קודים הספרות לא מופיעות ברצף?
- (4) בארונית 4 מגירות. ילד התבקש על ידי אמו לסדר 6 משחקים בארונית. הילד מכניס את המשחקים באקראי למגירות השונות. כל מגירה יכולה להכיל את כל המשחקים יחד.
- מה ההסתברות שהילד יכניס את כל המשחקים למגירה העליונה?
  - מה ההסתברות שהילד יכניס את כל המשחקים לאותה מגירה?
  - מה ההסתברות ש"דומינו" יוכנס למגירה העליונה ויתר המשחקים לשאר המגירות.
  - מה ההסתברות ש"דומינו" לא יוכנס למגירה העליונה?

- (5) בעיר מסוימת מתמודדות למועצת העיר 4 מפלגות שונות: "הירוקים", "קדימה", "העבודה" ו"הליכוד". 6 אנשים אינם יודעים למי להצביע, ולכן בוחרים באקראי מפלגה כלשהי.
- מה ההסתברות שכל ה-6 יבחרו באותה מפלגה?
  - מה ההסתברות שמפלגת ה"ירוקים" לא תקבל קולות?
  - מה ההסתברות שמפלגת ה"ירוקים" תקבל בדיוק 3 קולות וכל מפלגה אחרת תקבל קול 1 בלבד?
  - מה ההסתברות שמפלגת "הירוקים" תקבל 2 קולות, מפלגת "העבודה" תקבל 2 קולות ומפלגת "הליכוד" תקבל 2 קולות?
- (6) 5 חברים נפגשו ורצו לראות סרט. לרשותם ספריה המונה 8 סרטים שונים. כל אחד התבקש לבחור סרט באקראי.
- מה ההסתברות שכולם יבחרו את אותו הסרט?
  - מה ההסתברות שכולם יבחרו את "הנוסע השמיני"?
  - מה ההסתברות שכל אחד יבחר סרט אחר?
  - מה הסיכוי שלפחות שניים יבחרו את אותו הסרט?
  - מה ההסתברות שיוסי וערן ייבחרו את "הנוסע השמיני" וכל השאר סרטים אחרים?
  - מה ההסתברות שהנוסע השמיני לא ייבחר על ידי אף אחד מהחברים?
  - לקחו את 8 הסרטים ויצרו מהם רשימה. נתון שברשימה 3 סרטי אימה, מה ההסתברות שברשימה שנוצרה יופיעו 3 סרטי האימה ברצף?
- (7) בקבוצה 10 אנשים. יש ליצור שתי וועדות שונות מתוך הקבוצה: אחת בת 4 אנשים והשנייה בת 3 אנשים. כל אדם יכול להיבחר רק לוועדה אחת. חשבו את מס' הדרכים השונות ליצירת הוועדות הללו כאשר:
- אין בוועדות תפקידים.
  - בכל וועדה יש תפקיד אחד של אחראי הוועדה.
  - בכל וועדה כל התפקידים שונים.
- (8) 4 גברים ו-3 נשים מתיישבים על כסאות בשורה של כסאות תיאטרון. בכל שורה 10 כסאות. בכמה דרכים שונות ניתן לבצע את ההושבה:
- ללא הגבלה.
  - כל הגברים ישבו זה ליד זה וגם כל הנשים תשבנה זו ליד זו.
  - שני גברים בקצה אחד ושני הגברים האחרים בקצה שני.
- (9) בהגרלה ישנם 10 מספרים מ-1 עד 10. נבחרו באקראי 5 מספרים. מה ההסתברות שהמספר 7 הוא השני בגודלו מבין המספרים שנבחרו?

- 10** 6 אנשים עלו לאוטובוס שעוצר ב-10 תחנות.  
 כל אדם בוחר באופן עצמאי ואקראי באיזו תחנה לרדת.  
 א. מה ההסתברות שכל אחד יורד בתחנה אחרת?  
 ב. מה ההסתברות שבדיוק 3 ירדו בתחנה החמישית?  
 ג. מה ההסתברות שרונית תרד בתחנה השנייה והשאר לא?  
 ד. מה ההסתברות שכולם ירדו בתחנות 5,6 ולפחות אחד בכל אחת מהתחנות הללו?

- 11** ברכבת 4 מקומות ישיבה עם כיוון הנסיעה 41 מקומות ישיבה נגד כיוון הנסיעה.



- 4 זוגות התיישבו במקומות אלו באקראי.  
 א. בכמה דרכים שונות ניתן להתיישב?  
 ב. מה ההסתברות שהזוג כהן ישבו זה לצד זה עם כיוון הנסיעה?  
 ג. מה ההסתברות שהזוג כהן ישבו זה לצד זה?  
 ד. מה ההסתברות שהזוג כהן ישבו כל אחד ליד החלון? (בכל שורה יש חלון).  
 ה. מה ההסתברות שהזוג כהן יישבו כך שכל אחד בכיוון נסיעה מנוגד?  
 ו. מה ההסתברות שהזוג כהן יישבו אחד מול השני פנים מול פנים.  
 ז. מה ההסתברות שכל הגברים ייסעו עם כיוון הנסיעה וכל הנשים תשבנה נגד כיוון הנסיעה?  
 ח. מה ההסתברות שכל זוג ישב אחד מול השני?

- 12** סיסמא מורכבת מ-5 תווים, תווים אלו יכולים להיות ספרה (0-9) ואותיות ה-ABC (26 אותיות). כל תו יכול לחזור על עצמו יותר מפעם אחת.  
 א. כמה סיסמאות שונות יש?  
 ב. כמה סיסמאות שונות יש שבהן כל התווים שונים?  
 ג. כמה סיסמאות שונות יש שבהן לפחות ספרה אחת ולפחות אות אחת?

- 13** מתוך קבוצה בת  $n$  אנשים רוצים לבחור 3 אנשים לוועדה. בכמה דרכים שונות ניתן לבצע את הבחירה? בטא את תשובתך באמצעות  $n$ .  
 א. בוועדה אין תפקידים ויש לבחור 3 אנשים שונים לוועדה.  
 ב. בוועדה תפקידים שונים. וכל אדם לא יכול למלא יותר מתפקיד אחת.  
 ג. בוועדה תפקידים שונים ואדם יכול למלא יותר מתפקיד אחד.

- 14** שני אנשים מטילים כל אחד מטבע  $n$  פעמים. בטאו באמצעות  $n$  את הסיכוי שלכל אחד מהם אותו מספר פעמים של התוצאה "ראש".

- 15** יוצרים קוד עם  $a$  ספרות (מותר לחזור על אותה ספרה בקוד).  
 חשבו את ההסתברויות הבאות (בטאו את תשובותיכם באמצעות  $a$ ):
- בקוד אין את הספרה 5.
  - בקוד מופיעה הספרה 3.
  - בקוד לא מופיעות ספרות אי זוגיות.
- 16** זוג קוביות הוטלו מספר פעמים. כמה פעמים יש להטיל את זוג הקוביות בכדי שבהסתברות של לפחות 0.5 תתקבל לפחות הטלה אחת (של הזוג) עם סכום תוצאות 12?
- 17** בוחרים באופן מקרי מספר בין 6 ספרות.
- מה הסיכוי שהספרה 5 תופיע בדיוק פעם אחת במספר?
  - מה הסיכוי שהספרה 4 תופיע לפחות פעם אחת וגם הספרה 0 תופיע לפחות פעם אחת במספר?
- 18** במשרד של דנה 5 תיקיות אותן היא מסדרת באקראי בטור. 3 תיקיות הן אדומות ו-2 תיקיות הן כחולות. דנה רשמה שני פתקים ושמה כל פתק במקום אקראי בין התיקיות (לכל פתק יש 4 אפשרויות למיקום).
- מה הסיכוי ששני הפתקים יהיו במקומות שונים?
  - מה הסיכוי שבין שני הפתקים יש שתי תיקיות אדומות ואין תיקיות כחולות?
  - מה הסיכוי שבין שני הפתקים יש בדיוק תיקיה אחת?
  - מה הסיכוי שבין שני הפתקים יש שתי תיקיות ואחת מהן כחולה?
- 19** לירון 6 עטים אותם הוא מכניס באקראי ל-3 קלמרים שונים. לכל עט הוא בוחר באופן מקרי קלמר.
- מה הסיכוי שיש בדיוק 2 קלמרים שבכל קלמר בדיוק 2 עטים?
  - מה הסיכוי שיש בדיוק קלמר אחד שבו בדיוק 2 עטים?
  - מה הסיכוי שיש בדיוק 3 קלמרים שבכל אחד בדיוק 2 עטים?
- 20** מסדרים  $n$  כדורים שונים ב  $n$  תאים שונים (תא יכול להכיל יותר מכדור אחד). מה הסיכוי שבתא  $i$  ( $1 \leq i \leq n$ ) יהיו בדיוק  $k$  כדורים?
- 21** בתחרות ריצה עלו לגמר 6 מתמודדים. רק בשלושת המקומות הראשונים זוכים במדליות. נניח שכל המתמודדים מסיימים את התחרות.
- כמה אפשרויות יש לסיים את התחרות?
  - כמה אפשרויות יש לכך שמתמודד מספר 6 יקבל מדליה?
  - כמה אפשרויות יש לכך שמתמודד מספר 6 יקבל מדליה או שמתמודד מספר 2 יקבל מדליית זהב?

22) מטילים קובייה הוגנת  $k$  פעמים.

- א. מה הסיכוי שהתוצאה הכי גדולה שהתקבלה היא  $j$  ?
- ב. מה הסיכוי שהתוצאה הכי קטנה שהתקבלה היא  $i$  ?
- ג. עבור  $i \leq j$ , מה הסיכוי שהתוצאה הכי גדולה היא  $j$  וגם התוצאה הכי קטנה היא  $i$  ?

## תשובות סופיות:

- (1) א. 102,400,000    ב. 78,960,960    ג. 658008
- (2) א. 810,000    ב. 657,720    ג. 27,405
- (3) א. 14,040,000    ב. 1,404,000    ג. 5,616,000    ד. 8,424,000
- (4) א. 0.00024    ב. 0.00098    ג. 0.05933    ד. 0.75000
- (5) א. 0.00098    ב. 0.17798    ג. 0.02929    ד. 0.02197
- (6) א.  $\frac{1}{4096}$     ב.  $\frac{1}{32,768}$     ג. 0.205    ד. 0.795
- ה. 0.0105    ו. 0.5129    ז. 0.1071
- (7) א. 4,200    ב. 50,400    ג. 604,800
- (8) א. 604,800    ב. 2,880    ג. 2,880
- (9) 0.238
- (10) א. 0.1512    ב. 0.014    ג. 0.059    ד.  $\frac{62}{10^6}$
- (11) א. 40,320    ב. 0.1071    ג. 0.2142    ד. 0.0357
- ה. 0.5714    ו. 0.1429    ז. 0.0143    ח. 0.0095
- (12) א. 60,466,176    ב. 45,239,040    ג. 48,484,800
- (13) א.  $\frac{n!}{3!(n-3)}$     ב.  $n \cdot (n-1)(n-2)$     ג.  $n^3$
- (14)  $\frac{1}{4^n} \cdot \sum_{i=0}^n \binom{n}{i}^2$
- (15) א.  $0.9^a$     ב.  $1-0.9^a$     ג.  $0.5^a$
- (16) לפחות 25 פעמים.
- (17) א. 0.35721    ב. 0.1759
- (18) א. 0.75    ב. 0.075    ג. 0.375    ד. 0.15
- (19) א. 0    ב.  $\frac{450}{729}$     ג.  $\frac{90}{729}$
- (20)  $\frac{\binom{n}{k} (n-1)^{n-k}}{n^n}$
- (21) א. 720    ב. 360    ג. 432

$$(22) \quad \text{א. } \frac{j^k - (j-1)^k}{6^k} \quad \text{ב. } \frac{(7-i)^k - (6-i)^k}{6^k} \\ \text{ג. } \frac{(j-i+1)^k - 2 \cdot (j-i)^k + (j-i-1)^k}{6^k}$$

# סטטיסטיקה והסתברות 01 88165

פרק 23 - הסתברות מותנית-במרחב מדגם אחיד

תוכן העניינים

105 ..... 1. כללי

## הסתברות מותנית – במרחב מדגם אחיד:

**רקע:**

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

ההסתברות של  $A$  בהינתן ש- $B$  כבר קרה:  $P(A|B)$ .

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

דוגמה (פתרון בהקלטה):

נטיל קובייה.

נגדיר:

$A$  - התוצאה זוגית.

$B$  - התוצאה גדולה מ-3.

נרצה לחשב את:  $P(A|B)$ .

## שאלות:

- (1) נבחרה ספרה זוגית באקראי. מה הסיכוי שהספרה גדולה מ-6?
- (2) יוסי הטיל קובייה.  
מה הסיכוי שקיבל את התוצאה 4, אם ידוע שהתוצאה שהתקבלה זוגית?
- (3) הוטלו צמד קוביות. נגדיר:  
 $A$  - סכום התוצאות בשתי ההטלות הינו 7.  
 $B$  - מכפלת התוצאות 12.  
 חשבו את  $P(A|B)$ .
- (4) מטבע הוטל פעמיים.  
ידוע שהתקבל לכל היותר ראש אחד, מה הסיכוי שהתקבלו שני ראשים?
- (5) זוג קוביות הוטלו והתקבל שהתוצאות זהות.  
מה הסיכוי שלפחות אחת התוצאות 5?
- (6) זוג קוביות הוטלו והתקבל לפחות פעם אחת 4.  
מה הסיכוי שאחת התוצאות 5?
- (7) נבחרה משפחה בת שני ילדים, שמהם אחד הוא בן.  
מה ההסתברות שבמשפחה שני בנים בקרב הילדים?
- (8) נבחרה משפחה בת שלושה ילדים, ונתון שהילד האמצעי בן.  
מה הסיכוי שיש בנות בקרב הילדים?
- (9) בכיתה 6 בנים ו-7 בנות. נבחרו 4 ילדים מהכיתה.  
אם ידוע שנבחרו 2 בנים ו-2 בנות, מה הסיכוי שאלעד לא נבחר?
- (10) חמישה חברים יצאו לבית קולנוע והתיישבו זה לצד זה באקראי,  
בכיסאות מספר 5 עד 9. ידוע שערך ודין התיישבו זה ליד זה.  
מה ההסתברות שהם יושבים בכיסאות מספר 6 ו-7?

**תשובות סופיות:**

(1) 0.2

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

(3) 0.5

(4) 0

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

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

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

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

(9)  $\frac{2}{3}$

(10)  $\frac{1}{4}$

# סטטיסטיקה והסתברות 01 88165

פרק 24 - הסתברות מותנית - מרחב לא אחיד

תוכן העניינים

108 ..... 1. כללי

## הסתברות מותנית – מרחב לא אחיד:

רקע:

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

במונה: הסיכוי לחיתוך של שני המאורעות, זה הנשאל וזה הנתון שהתרחש.  
 במכנה: הסיכוי למאורע 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, בקרב 15% מהמשפחות הללו שתי המכוניות הן מתוצרת אירופאית. אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית?

## שאלות:

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

4) נלקחו 200 שכירים ו-100 עצמאים. מתוך השכירים 20 הם אקדמאיים, ומתוך העצמאיים 30 הם אקדמאיים.

- א. בנו טבלת שכיחות משותפת לנתונים.
- ב. נבחר אדם אקראי מה ההסתברות שהוא שכיר?
- ג. מה ההסתברות שהוא שכיר ולא אקדמאי?
- ד. מה ההסתברות שהוא שכיר או אקדמאי?
- ה. אם האדם שנבחר הוא עצמאי מהי ההסתברות שהוא אקדמאי?
- ו. אם האדם שנבחר הוא לא אקדמאי, מה ההסתברות שהוא שכיר?

5) חברה מסוימת פרסמה את הנתונים הבאים לגבי האזרחים מעל גיל 21 :  
 40% מהאנשים מחזיקים כרטיס "ויזה", 52% מחזיקים כרטיס "ישראלכרט",  
 20% מחזיקים כרטיס "אמריקן אקספרס", 15% מחזיקים ויזה וגם ישראלכרט,  
 8% מחזיקים ישראלכרט וגם אמריקן אקספרס ו-7% מחזיקים כרטיס ויזה וגם אמריקן אקספרס. כמו כן, 5% מחזיקים בשלושת הכרטיסים הנ"ל.

- א. אם לאדם יש ויזה, מה הסיכוי שאין לו ישראלכרט?
- ב. אם לאדם שני כרטיסי אשראי, מה הסיכוי שאין לו ישראלכרט?
- ג. אם לאדם לפחות כרטיס אחד, מה הסיכוי שאין לו ישראלכרט?

## תשובות סופיות:

(1) א. 0.833    ב. 0.9375    ג. 0.0625    ד. 0.5    ה. 0.789

(2) א. 5%    ב. 0.0833    ג. 0.786    ד. 0.6875

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

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

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

ה. 0.3    ו. 0.72

(5) א. 0.625    ב. 0.133    ג. 0.402

# סטטיסטיקה והסתברות 01 88165

פרק 25 - דיאגרמת עצים - נוסחת ביס ונוסחת ההסתברות השלמה

תוכן העניינים

1. כללי.....112

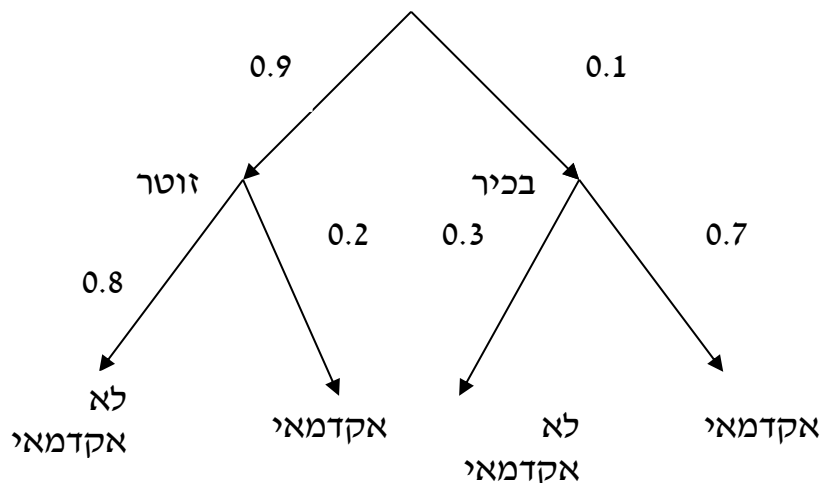
## דיאגרמת עצים – נוסחת בייס ונוסחת ההסתברות השלמה:

### רקע:

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

### דוגמה:

בחברה מסוימת 10% מוגדרים בכירים והיתר מוגדרים זוטרים. מבין הבכירים 70% הם אקדמאים ומבין הזוטרים 20% הם אקדמאים. נשרטט עץ שיתאר את הנתונים, השלב הראשון של העץ אינו מותנה בכלום 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 אקדמאי ?  $0.1 \cdot 0.7 = 0.07$ .

(2) מה הסיכוי שהוא זוטר לא אקדמאי ?  $0.9 \cdot 0.8 = 0.72$ .

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

(3) מה הסיכוי שהוא אקדמאי ?  $0.1 \cdot 0.7 + 0.9 \cdot 0.2 = 0.25$ .

(4) נבחר אקדמאי מה ההסתברות שהוא עובד זוטר? מדובר כאן על שאלה בהסתברות מותנה ולכן נשתמש בעיקרון של הסתברות

$$P(zutar | academay) = \frac{0.9 \cdot 0.2}{0.25} = \frac{0.18}{0.25} = 0.72 \text{ : מותנה}$$

### נוסחת ההסתברות השלמה:

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

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

### נוסחת בייס:

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

## שאלות:

- (1) בשקית סוכריות 4 סוכריות תות ו-3 לימון. מוציאים באקראי סוכריה. אם היא בטעם תות אוכלים אותה ומוציאים סוכריה נוספת, ואם היא בטעם לימון מחזירים אותה לשקית ומוציאים סוכריה נוספת.
- א. מה ההסתברות שהסוכריה הראשונה שהוצאה בטעם תות והשנייה בטעם לימון?
- ב. מה ההסתברות שהסוכריה השנייה בטעם לימון?
- (2) באוכלוסייה מסוימת 30% הם ילדים, 50% בוגרים והיתר קשישים. לפי נתוני משרד הבריאות הסיכוי שילד יחלה בשפעת במשך החורף הוא 80%, הסיכוי שמבוגר יחלה בשפעת במשך 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 ההסתברות שהוא קשיש?
- ד. נבחר ילד, מה ההסתברות שהוא לא יחלה בשפעת במשך החורף?
- (3) בכד א' 5 כדורים כחולים ו-5 כדורים אדומים. בכד ב' 6 כדורים כחולים ו-4 כדורים אדומים. בוחרים באקראי כד, מוציאים ממנו כדור ומבלי להחזירו מוציאים כדור נוסף.
- א. מה ההסתברות ששני הכדורים שיוצאו יהיו בצבעים שונים?
- ב. אם הכדורים שהוצאו הם בצבעים שונים, מה ההסתברות שהכדור השני שהוצא יהיה בצבע אדום?
- (4) חברת סלולר מסווגת את לקוחותיה לפי 3 קבוצות גיל: נוער, בוגרים ופנסיונרים. נתון כי: 10% מהלקוחות בני נוער, 70% מהלקוחות בוגרים והיתר פנסיונרים. מתוך בני הנוער 90% מחזיקים בסמארט-פון, מתוך האוכלוסייה הבוגרת ל-70% יש סמארט-פון ומתוך אוכלוסיית הפנסיונרים 30% מחזיקים בסמארט-פון.
- א. איזה אחוז מלקוחות החברה הם בני נוער עם סמארט-פון?
- ב. נבחר לקוח אקראי ונתון שיש לו סמארט-פון. מה ההסתברות שהוא פנסיונר?
- ג. אם ללקוח אין סמארט-פון, מה ההסתברות שהוא לא בן נוער?

- (5) כדי להתקבל למקום עבודה יש לעבור שלושה מבחנים. המבחנים הם בשלבים, כלומר לאחר כישלון במבחן מסוים אין אפשרות לגשת למבחן הבא אחריו. 70% מהמועמדים עוברים את המבחן הראשון. מתוכם, 50% עוברים את המבחן השני. מבין אלה שעוברים את המבחן השני 40% עוברים את המבחן השלישי.
- א. מה ההסתברות להתקבל לעבודה?  
 ב. מועמד לא התקבל לעבודה. מה ההסתברות שהוא נכשל במבחן הראשון?  
 ג. מועמד לא התקבל לעבודה. מה ההסתברות שהוא עבר את המבחן השני?
- (6) משרד הבריאות פרסם את הנתונים הבאים:
- מתוך אוכלוסיית הילדים והנוער 80% חולים בשפעת בזמן החורף.  
 מתוך אוכלוסיית המבוגרים (עד גיל 65) 60% חולים בשפעת בזמן החורף.  
 30% מהתושבים הם ילדים ונוער. 50% הם מבוגרים. היתר קשישים.  
 כמו כן נתון ש 68% מהאוכלוסייה תחלה בשפעת בחורף.
- א. מה אחוז החולים בשפעת בקרב האוכלוסייה הקשישה?  
 ב. נבחר אדם שלא חלה בשפעת, מה ההסתברות שהוא לא קשיש?
- (7) רדאר שנמצא על החוף צריך לקלוט אנייה הנמצאת ב-1 מ-4 האזורים: A, B, C, D. אם האנייה נמצאת באזור A הרדאר מזהה אותה בסיכוי 0.8, סיכוי זה פוחת ב-0.1 ככל שהאנייה מתקדמת באזור. כמו כן נתון שבהסתברות חצי האנייה נמצאת באזור D, בהסתברות 0.3 באזור C, באזור B היא נמצאת בסיכוי 0.2, אחרת היא נמצאת באזור A.
- א. מה הסיכוי שהאנייה תתגלה ע"י הרדאר?  
 ב. אם האנייה התגלתה ע"י הרדאר, מה ההסתברות שהיא נמצאת באזור C?  
 ג. אם האנייה התגלתה ע"י הרדאר, מה הסיכוי שהיא לא נמצאת באזור B?
- (8) סימפטום X מופיע בהסתברות של 0.4 במחלה A, בהסתברות של 0.6 במחלה B ובהסתברות של 0.5 במחלה C. סימפטום X מופיע אך ורק במחלות הללו, אדם לא יכול לחלות ביותר ממחלה אחת מבין המחלות הללו. לקליניקה מגיעים אנשים כדלקמן: 8% חולים במחלה A, 10% במחלה B, 2% במחלה C והיתר בריאים. כמו כן נתון שבמחלה A, סימפטום X מתגלה בסיכוי של 80%, ובמחלות B, C הסימפטום מתגלה בסיכוי של 90% בכל מחלה.
- א. מה ההסתברות שאדם הגיע לקליניקה וגילו אצלו את סימפטום X?  
 ב. אם התגלה אצל אדם סימפטום X, מה ההסתברות שהוא חולה במחלה A?  
 ג. אם לאדם יש את סימפטום X, מה ההסתברות שהוא חולה במחלה A?  
 ד. אם לא גילו אצל אדם את סימפטום X, מה ההסתברות שהוא בריא?

- (9) סטודנט ניגש למבחן אמריקאי. הסיכוי שהוא יודע תשובה לשאלה מסוימת הוא  $P$ , ואם הוא לא יודע את התשובה הוא מנחש. בכל מקרה הוא עונה על השאלה. נתון שלשאלה יש  $k$  תשובות אפשריות. אם הסטודנט ענה נכון על השאלה, מה הסיכוי שהוא ידע אותה?
- (10) אדם משחק נגד שני מתמודדים, רונית ודולב. האדם צריך לשחק שלושה משחקים ויש לו לבחור איזה סדר משחקים עדיף לו:
- דולב, רונית, דולב.
  - רונית, דולב, רונית.
- בכל משחק מישהו חייב לנצח(אין תיקו). האדם ינצח בטורניר רק אם ינצח בשני משחקים ברציפות. נתון שדולב שחקן טוב יותר מאשר רונית. איזו אפשרות עדיפה יותר על האדם כדי לנצח בטורניר?

### תשובות סופיות:

- (1) א.  $\frac{2}{7}$  ב.  $\frac{23}{49}$
- (2) א. 6% ב. 58% ג. 0.241 ד. 0.2
- (3) א. 0.544 ב. 0.5
- (4) א. 9% ב. 0.09375 ג. 0.9722
- (5) א. 0.14 ב. 0.3488 ג. 0.2442
- (6) א. 70% ב. 0.8125
- (7) א. 0.57 ב. 0.3158 ג. 0.7543
- (8) א. 0.0886 ב. 0.2889 ג. 0.3137 ד. 0.8778
- (9)  $\frac{kp}{1+p(k-1)}$
- (10) א'

# סטטיסטיקה והסתברות 01 88165

פרק 26 - תלות ואי תלות בין מאורעות

תוכן העניינים

1. אי תלות בין מאורעות ( מורחב) ..... 117

## תלות ואי תלות בין מאורעות:

### רקע:

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

הדבר גורר גם ההפך:  $P(A|B) = P(A)$ , כלומר, 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,

חוקר מבצע שני ניסויים בלתי תלויים הסיכוי להצליח בניסוי הראשון הנו 0.7 והסיכוי להצליח בניסוי השני הוא 0.4.

א. מה הסיכוי להצליח בשני הניסויים יחדו?

כיוון שהמאורעות הללו בלתי תלויים:

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

ב. מה הסיכוי להיכשל בשני הניסויים?

באופן דומה:

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

**הרחבה: אי תלות בין  $n$  מאורעות:**

$n$  מאורעות  $A_1, \dots, A_n$  הם בלתי תלויים אם ורק אם:  $P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i)$ .

## שאלות:

- (1) נתון:  $P(A) = 0.2$ ,  $P(B) = 0.5$ ,  $P(A \cup B) = 0.6$ .  
האם המאורעות הללו בלתי תלויים?
- (2) תלמיד ניגש לשני מבחנים שהצלחתם לא תלויה זו בזו. הסיכוי שלו להצליח במבחן הראשון הוא 0.7 והשני 0.4.  
א. מה הסיכוי להצליח בשני המבחנים יחדו?  
ב. מה הסיכוי שניכשל בשני המבחנים?
- (3) במדינה מסוימת יש 8% אבטלה, נבחרו באקראי שני אנשים מהמדינה.  
א. מה ההסתברות ששניהם מובטלים?  
ב. מה ההסתברות שלפחות אחד מהם מובטל?
- (4) מוצר צריך לעבור בהצלחה ארבע בדיקות בלתי תלויות לפני שיווקו, אחרת הוא נפסל ולא יוצא לשוק. הסיכוי לעבור בהצלחה כל אחת מהבדיקות הוא 0.8. בכל מקרה מבוצעות כל 4 הבדיקות.  
א. מה הסיכוי שהמוצר יפסל?  
ב. מה ההסתברות שהמוצר יעבור בהצלחה לפחות בדיקה אחת?
- (5) במדינה מסוימת יש 8% אבטלה, נבחרו באקראי חמישה אנשים מהמדינה.  
א. מה ההסתברות שכולם מובטלים במדגם?  
ב. מה ההסתברות שלפחות אחד מהם מובטל?
- (6) עבור שני מאורעות  $A$  ו- $B$  המוגדרים על אותו מרחב מדגם נתון ש:  
 $P(A|B) = 0.6$ ,  $P(A \cap \bar{B}) = 0.3$ ,  $P(A \cup B) = 0.9$ .  
האם  $A$  ו- $B$  מאורעות בלתי תלויים?
- (7) הוכיח שאם:  $P(A/B) = P(B/A)$ , אז:  $P(A) = P(B)$ .

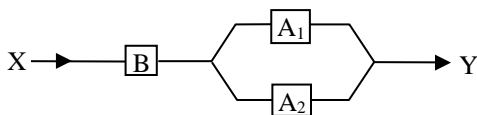
- 8 קבעו אילו מהטענות הבאות נכונות. נמק! א. אם :  $P(A \cup B) = P(A) \cdot P(B)$ , אזי המאורעות בלתי תלויים.  
 ב. מאורע  $A$  כלול במאורע  $B$  :  $0 < P(B) < 1$ ,  $P(A) > 0$ , לכן :  $P(A/B) < P(A)$ .  
 ג.  $A$  ו- $B$  מאורעות זרים שסיכוייהם חיוביים לכן הם מאורעות תלויים.  
 ד.  $A$  ו- $B$  מאורעות תלויים שסיכוייהם חיוביים לכן  $A$  ו- $B$  מאורעות זרים.  
 ה.  $\bar{P}(\bar{A} \cap \bar{B}) = 1 - P(A) - P(B)$  לכן  $A$  ו- $B$  מאורעות זרים.

- 9 זוג מעוניין להביא ילד לעולם, בבדיקה גנטית שהם עשו לאב התגלה שהאב אינו נשא של מחלה  $Q$ . מערכים את הסיכוי של האם להיות נשאית למחלץ  $Q$  להיות 0.2. אם האם נשאית היא תלד בכל פעם ילד חולה בסיכוי 0.5 באופן בלתי תלוי בין הלידות. האם ילדה שני ילדים. האם המאורעות "הילד הראשון בריא" ו-"הילד השני בריא" הם מאורעות בלתי תלויים?

- 10 מטילים פעמיים מטבע עם הסתברות  $p$  לעץ בכל הטלה,  $0 < p < 1$ .  
 $A$  – יצא עץ בהטלה ראשונה.  
 $B$  – יצאו תוצאות שונות.  
 עבור איזה ערכים של  $p$  המאורעות  $A$  ו- $B$  בלתי תלויים?

- 11 הוכח אם  $A$  ו- $B$  בלתי תלויים, אזי  $\bar{A}$ ,  $\bar{B}$  בלתי תלויים.

- 12 נתונה מערכת חשמלית שבשרטוט, כל מתג יכול להיות פתוח או סגור בהסתברויות שונות, אך באופן בלתי תלוי זה בזה.



להלן ההסתברות של כל מתג להיות סגור :  
 $P(A_1) = 0.7$ ,  $P(A_2) = 0.8$ ,  $P(B) = 0.9$

- א. מה ההסתברות שיעבור זרם במערכת החשמלית?  
 ב. אם לא עובר זרם במערכת, מה הסיכוי שמתג  $B$  סגור?

- 13 מטילים שתי קוביות הוגנות. נגדיר שלושה מאורעות :  
 $A$  – תוצאה של קובייה ראשונה זוגית.  
 $B$  – תוצאה של קובייה שניה אי זוגית.  
 $C$  – סכום התוצאות של שתי הקוביות זוגי.  
 האם המאורעות בלתי-תלויים?

14) ענה על הסעיפים הבאים :

א. המאורעות  $A$  ו- $B$  הם מאורעות זרים של ניסוי כלשהו. חוזרים על אותו ניסוי שוב באופן בלתי תלוי זה בזה. הוכיחו שהסיכוי שמאורע  $A$  יתרחש בניסוי לפני שמאורע  $B$  יתרחש

$$\frac{P(A)}{P(A)+P(B)}$$

בניסוי הוא :

ב. מטילים קובייה הוגנת פעם אחרי פעם, מה הסיכוי שתתקבל תוצאה זוגית עוד לפני שתתקבל התוצאה 3?

ג. מטילים קובייה הוגנת פעם אחרי פעם, מה הסיכוי שתתקבל תוצאה זוגית עוד לפני שתתקבל תוצאה גדולה מ-4?

### תשובות סופיות:

- 1) א. כן.
- 2) א. 0.28. ב. 0.18.
- 3) א. 0.0064. ב. 0.1536.
- 4) א. 0.5904. ב. 0.9984.
- 5) א.  $0.08^5$ . ב. 0.3409.
- 6) לא, הם תלויים.
- 7) שאלת הוכחה.
- 8) א. לא נכון. ב. לא נכון. ג. נכון. ד. לא נכון. ה. נכון.
- 9) תלויים.
- 10) 0.5.
- 11) שאלת הוכחה.
- 12) א. 0.846. ב. 0.3506.
- 13) תלויים.
- 14) א. שאלת הוכחה. ב.  $\frac{3}{4}$ . ג.  $\frac{1}{2}$ .

# סטטיסטיקה והסתברות 01 88165

פרק 27 - שאלות מסכמות בהסתברות

תוכן העניינים

1. כללי ..... 121

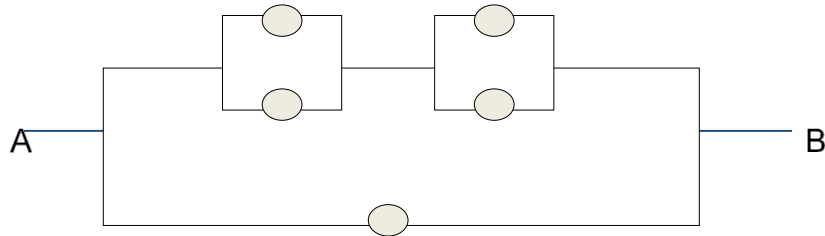
## שאלות מסכמות בהסתברות:

### 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.
- א. מה ההסתברות שמשפחה אקראית בת שתי מכוניות תהיה ללא מכוניות מתוצרת אירופה?
- ב. מה ההסתברות שלפחות מכונית אחת תהיה אירופאית?
- ג. ידוע שלמשפחה יש מכונית אירופאית.
- מה ההסתברות שרק המכונית החדשה שלה היא מתוצרת אירופאית?
- ד. אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית?
- (2) במדינת "שומקום" 50% מהחלב במרכולים מיוצר במחלבה א', 40% במחלבה ב' והיתר במחלבה ג'. 3% מתוצרת מחלבה א' מגיעה חמוצה למרכולים ואילו במחלבה ב' 10%.
- כמו כן ידוע שבמדינת "שומקום" בסך הכול 7.5% מהחלב חמוץ.
- א. איזה אחוז מהחלב שמגיע למרכול ממחלבה ג' חמוץ?
- ב. אם נרכש חלב חמוץ במרכול. מה הסיכוי שהוא יוצר במחלבה ג'?
- ג. ברכישת חלב נמצא שהוא אינו חמוץ. מה הסיכוי שהוא יוצר במחלבה א'?
- ד. האם המאורעות: "חלב חמוץ" ו-"יוצר במחלבה א'" בלתי תלויים?
- (3) רוני ורונה יצאו לבלות במרכז בילויים עם מספר אפשרויות בילוי: בהסתברות של 0.3 הם ייצאו לבאולינג, בהסתברות של 0.5 הם ייצאו לבית קפה ובהסתברות של 0.7 הם ייצאו לפחות לאחד מהם (באולינג/קפה).
- א. מה ההסתברות שהם יצאו רק לבאולינג?
- ב. האם המאורעות "לצאת לבאולינג" ו-"לצאת לבית קפה" זרים?
- ג. האם המאורעות "לצאת לבאולינג" ו-"לצאת לבית קפה" תלויים?
- ד. מה ההסתברות שיום אחד הם יצאו רק לבאולינג וביום למחרת לא יצאו לאף אחד מהמקומות?

- 4) 70% מהנבחנים בסטטיסטיקה עוברים את מועד א'. כל מי שלא עובר את מועד א' ניגש לעשות מועד ב', מתוכם 80% עוברים אותו. מבין אלה שנכשלים בשני המועדים 50% נרשמים לקורס מחדש, והיתר פורשים מהתואר.
- מה הסיכוי שסטודנט אקראי עבר את הקורס?
  - אם סטודנט אקראי עבר הקורס, מה הסיכוי שעבר במועד ב'?
  - מה אחוז הסטודנטים שפורשים מהתואר?
  - נבחרו 2 סטודנטים אקראיים רונית 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 הכול 13% מהאוכלוסייה מובטלת.
- מה אחוז האבטלה בקרב הנשים?
  - נבחר אדם מובטל, מה ההסתברות שזו אישה?
  - נגדיר 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 רגילים (ראש, זנב) ו-3 מטבעות שבשני צדדיהם טבוע ראש. אדם בוחר באקראי מטבע ומטיל אותו פעמיים. נסמן ב-A את ההטלה הראשונה ראש וב-B את ההטלה השנייה ראש.
- חשבו את הסיכויים למאורעות A ו-B.
  - האם המאורע A ו-B בלתי תלויים?
  - ידוע שבהטלה 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 בלפחות אחת מהמדיות.
- מה אחוז האנשים, מאלה שמעוניינים לרכוש רכב משומש, שיראו את 2 המודעות?
  - אם אדם ראה את המודעה באינטרנט, מה ההסתברות שהוא לא ראה את המודעה בעיתון?
  - האם המאורעות: "לראות את המודעה באינטרנט" ו-"לראות את המודעה בעיתון" בלתי תלויים?
  - 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.
    - מה ההסתברות שאדם המעוניין לרכוש רכב משומש יתקשר לערן?
    - אדם המעוניין לרכוש רכב משומש התקשר לערן. מה ההסתברות שהוא ראה את שתי המודעות?

8) נתונה המערכת החשמלית הבאה :



כל יחידה עובדת באופן בלתי תלוי ובהסתברות  $p$ .  
 כדי שהמערכת תפעל צריך לעבור זרם מהנקודה  $A$  לנקודה  $B$ .  
 הוכיחו שהסיכוי שהמערכת תפעל הוא:  $P + (1 - P)(2P - P^2)^2$ .

9) ליאת מעוניינת לתרגל לבחינה בהסתברות. היא מצאה באינטרנט מאגר הכולל 25 שאלות מבחינות. השאלות ממוספרות ו-6 מתוכן עוסקות במשתנה מקרי רציף. ליאת החליטה לבחור באקראי 7 שאלות מהמאגר במטרה לפתור אותן. כל שאלה שלא עוסקת במשתנה הרציף תיפתר על ידי מיכל בסיכוי של 90%, אך אם השאלה עוסקת במשתנה הרציף היא תיפתר בסיכוי של 60%.  
 א. מה הסיכוי שהשאלות שנבחרו הן כולן ממוספרות בסדר עוקב?  
 ב. מה הסיכוי ששאלה 20 היא השאלה עם המספור המקסימלי מבין השאלות שנבחרו?  
 ג. ידוע שליאת בחרה 2 שאלות שעוסקות במשתנה הרציף והיתר לא. מה הסיכוי שתצליח לפתור 6 מתוך השאלות שבחרה?

10) נתונים שלושה מאורעות:  $A$ ,  $B$  ו- $C$ . ידוע ש:  $P(A|B) = 1$ ,  $P(A|C) = 1$ .  
 תנו דוגמא ספציפית למאורעות:  $A$ ,  $B$  ו- $C$  שבה המאורעות  $B$  ו- $C$  תלויים.

11) הוכיחו או הפריכו (על ידי דוגמה נגדית) את הטענה הבאה:  
 אם  $A$  ו- $B$  בלתי תלויים, אז  $A$  ו- $\bar{B}$  בלתי תלויים.

12) משחקים משחק מזל פעמיים, כך שבכל משחק בודד יש אפשרות לנצח או להפסיד. הסיכוי לנצח בכל משחק הוא  $P$  כאשר:  $0 < P < 1$ .  
 נגדיר את המאורעות הבאים:  
 $A$  - תוצאות המשחקים שונות זו מזו.  
 $B$  - המשחק הראשון היה ניצחון.  
 מה ערכו של  $P$ , עבורו  $A$  ו- $B$  יהיו מאורעות בלתי תלויים?

**13** טל מניח בשורה  $N$  קוביות בצבעים שונים. בין שתי קוביות אקראיות כלשהן ערן מניח מכחול. הוכיחו שהסיכוי שהקובייה הכחולה והאדומה יהיו בצדדים

$$\frac{N+1}{3(N-1)} : \text{שונים של המכחול הוא}$$

**14** הוכיחו באמצעות אינדוקציה את אי שוויון בול:  $P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)$ .

## תשובות סופיות:

- (1) א. 0.25    ב. 0.75    ג. 0.6    ד. 0.5
- (2) א. 0.2    ב. 0.267    ג. 0.524    ד. תלויים.
- (3) א. 0.2    ב. אינם זרים.    ג. תלויים.    ד. 0.06
- (4) א. 0.94    ב. 0.255    ג. 0.03    ד. 0.168
- (5) א. 15%    ב. 0.692    ג. לא זרים ותלויים.
- (6) א. 0.65    ב. תלויים.    ג. 0.5384
- (7) א. 8%    ב. 0.733    ג. תלויים.
- ד. i. 0.478    ד. ii. 0.15
- (8) שאלת הוכחה.
- (9) א.  $\frac{19}{480,700}$     ב.  $\frac{27,132}{480,700}$     ג. 0.4015
- (10) ראו סרטון.
- (11) שאלת הוכחה.
- (12)  $\frac{1}{2}$
- (13) שאלת הוכחה.
- (14) שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 28 - המשתנה המקרי הבדיד - פונקציית ההסתברות

תוכן העניינים

1. כללי ..... 126

## המשתנה המקרי הבדיד – פונקציית ההסתברות:

**רקע:**

**משתנה מקרי בדיד:**

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

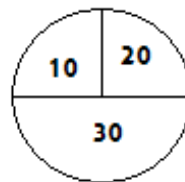
מתארים את המשתנה המקרי על ידי פונקציית ההסתברות.

**פונקציית ההסתברות:**

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

דוגמה (פתרון בהקלטה):

בקזינו יש רולטה כמתואר בשרטוט:



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה בש"ח. בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.

## שאלות:

- (1) ידוע שביישוב מסוים התפלגות מספר המכוניות למשפחה היא :  
 50 משפחות אינן מחזיקות במכונית.  
 70 משפחות עם מכונית אחת.  
 60 משפחות עם 2 מכוניות.  
 20 משפחות עם 3 מכוניות .  
 בוחרים באקראי משפחה מהיישוב, נגדיר את  $X$  להיות מספר המכוניות של המשפחה שנבחרה. בנו את פונקציית ההסתברות של  $X$ .
- (2) מהאותיות :  $A, B, C$  יוצרים קוד דו תווי.  
 א. כמה קודים ניתן ליצור?  
 ב. רשמו את כל הקודים האפשריים.  
 ג. נגדיר את  $X$  להיות מספר הפעמים שהאות  $B$  מופיעה בקוד.  
 בנו את פונקציית ההסתברות של  $X$ .
- (3) תלמיד ניגש בסמסטר לשני מבחנים : מבחן בכלכלה ומבחן בסטטיסטיקה. כמו כן, נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8, הסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9 והסיכוי לעבור את שני המבחנים הנו 0.75. יהי  $X$  מספר המבחנים שהסטודנט עבר. בנו את פונקציית ההסתברות של  $X$ .
- (4) הסיכוי לזכות במשחק מסוים הינו 0.3. אדם משחק את המשחק עד אשר הוא מנצח אך בכל מקרה הוא לא משחק את המשחק יותר מ-4 פעמים. נגדיר את  $X$  להיות מספר הפעמים שהוא שיחק את המשחק. בנו את פונקציית ההסתברות של  $X$ .
- (5) חברה לניהול פרויקטים מנהלת 3 פרויקטים במקביל. הסיכוי שפרויקט א' יצליח הינו 0.7, הסיכוי שפרויקט ב' יצליח הינו 0.8, והסיכוי שפרויקט ג' יצליח הינו 0.9. נתון שהצלחת כל פרויקט בלתי תלויה זו בזו. נגדיר את  $X$  להיות מספר הפרויקטים שיצליחו. בנו את פונקציית ההסתברות של  $X$ .
- (6) להלן פונקציית הסתברות של משתנה מקרי כלשהו :  $k = 1, 2, \dots, 4$  :  $P(X = k) = \frac{k}{A}$ . מצאו את ערכו של  $A$ .

- (7) בגן ילדים 8 ילדים, מתוכם 5 בנים ו-3 בנות. בוחרים באקראי 3 ילדים להשתתף בהצגה. נגדיר את  $X$  כמספר הבנים שנבחרו להצגה. בנו את פונקציית ההסתברות של  $X$ .
- (8) בסקר שנערך בדקו בקרב אנשים האם הם צופים במהדורת החדשות של ערוצים 1,2,10. להלן הנתונים:  
 20% צופים בערוץ 2.  
 8% צופים בערוץ 1.  
 10% צופים בערוץ 10.  
 כמו כן נתון ש 1% צופים בשלושת המהדורות גם יחד.  
 10% צופים בשתי המהדורות מתוך השלושה.  
 נגדיר את  $X$  להיות מספר המהדורות מבין 3 המהדורות המדוברות שאדם אקראי צופה. 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	$P(X)$

(5) להלן טבלה:

3	2	1	0	$X$
0.504	0.398	0.092	0.006	$P(X)$

(6) 0.10

(7) להלן טבלה:

4	3	2	1	$X$
$\frac{10}{56}$	$\frac{30}{56}$	$\frac{15}{56}$	$\frac{1}{56}$	$P(X)$

(8) להלן טבלה:

4	3	2	1	$X$
0.01	0.1	0.15	0.74	$P(X)$

# סטטיסטיקה והסתברות 01 88165

פרק 29 - המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן

תוכן העניינים

1. כללי ..... 130

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

**רקע:**

**תוחלת:**

ממוצע של פונקציית ההסתברות, אם נבצע את התהליך אינסוף פעמים כמה בממוצע נקבל. התוחלת היא צפי של המשתנה המקרי.

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

**שונות:**

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

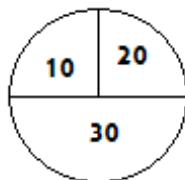
$$\text{מגדירים שונות באופן הבא: } V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = \sigma^2$$

**סטיית תקן:**

שורש של השונות – הפיזור הממוצע הצפוי סביב התוחלת. מסמנים:  $STD = \sigma$ .

**דוגמה:**

בקזינו רולטה כמוראה בשרטוט. אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה ב-ש. הסתברות לקבלת הסכומים השונים:



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

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

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

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

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

## שאלות:

- (1) אדם משחק במשחק מזל. נגדיר את  $X$  להיות סכום הזכייה. להלן פונקציית ההסתברות של  $X$ :

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

מהי התוחלת, השונות וסטית התקן של  $X$ ?

- (2) בישוב מסוים שני סניפי בנק: בנק פועלים ובנק לאומי. מתוך האוכלוסייה הבוגרת בישוב, ל-50% חשבון בנק בסניף הפועלים, ל-40% חשבון בנק בסניף לאומי ול-20% מהתושבים הבוגרים אין חשבון באף אחד מהסניפים. יהי  $X$  מס' סניפי הבנק שלבוגר בישוב יש בהם חשבון. חשבו את:  $E(X)$ .
- (3) ידוע של-20% מהמשפחות יש חיבור לווייני בביתם. בסקר אדם מחפש לראיין משפחה המחוברת ללוויין. הוא מטלפן באקראי למשפחה וממשיך עד אשר הוא מגיע למשפחה המחוברת ללוויין. בכל מקרה הסוקר לא יתקשר ליותר מ-5 משפחות. נגדיר את  $X$  להיות מספר המשפחות שאליהן האדם יתקשר. א. בנו את פונקציית ההסתברות של  $X$ . ב. חשבו את התוחלת וסטית התקן של  $X$ .
- (4) לאדם צרור מפתחות. בצרור 5 מפתחות אשר רק אחד מתאים לדלת של ביתו. האדם מנסה את המפתחות באופן מקרי. לאחר שניסה מפתח מסוים הוא מוציא אותו מהצרור כדי שלא ישתמש בו שוב. נסמן ב- $X$  את מספר הניסיונות עד שהדלת תפתח. א. בנו את פונקציית ההסתברות של  $X$ . ב. חשבו את התוחלת והשונות של  $X$ .

5) נתונה פונקציית ההסתברות של המשתנה המקרי  $X$ :

8	6	4	2	$X$
0.2		0.3		$E(X)$

כמו כן נתון ש:  $E(X) = 4.2$ .

א. מצאו את ההסתברויות החסרות בטבלה.

ב. חשבו את:  $V(X)$ .

6) משתנה מקרי בדיד מקבל את הערכים 5-0 ו-5. נתון שהתוחלת של המשתנה 0 ושהשונות היא 10. מצא את פונקציית ההסתברות.

7) להלן ההתפלגות של משתנה מקרי:

$X$	$P$
1	$\frac{1}{4}$
3	$\frac{1}{2}$
$K$	$\frac{1}{4}$

מהו הערך שייתן ערך מינימלי לשונות של  $X$ ?

### תשובות סופיות:

(1) תוחלת: 2, שונות: 796.

(2) 0.9.

(3) א. ראו סרטון. ב. תוחלת: 3.36, סטיית תקן: 1.603.

(4) א. ראו טבלה: ב. תוחלת: 3, שונות: 2.

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

(5) א. ראו טבלה: ב. 5.16.

8	6	4	2	$X$
0.2	0.1	0.3	0.4	$P(X)$

(6) ראו טבלה:

5	0	-5	$X$
0.2	0.6	0.2	$P(X)$

(7) 2.33.

# סטטיסטיקה והסתברות 01 88165

פרק 30 - המשתנה המקרי הבדיד - תוחלת של פונקציה של משתנה מקרי  
בדיד

תוכן העניינים

1. ראשי.....134

## תוחלת של פונקציה של משתנה מקרי בדיד:

רקע:

יהי  $X$  משתנה מקרי, ותהי  $g(X)$  פונקציה של  $X$ . אז:

$$E(g(X)) = g(x_1)P(X=x_1) + g(x_2)P(X=x_2) + g(x_3)P(X=x_3) + \dots$$

$$= \sum_i g(x_i) \cdot P(x_i)$$

כאשר  $x_1, x_2, x_3, \dots$  הם הערכים שהמשתנה  $X$  מקבל.

דוגמה (פתרון בהקלטה):

נתון:

$X$	0	1	2
$P(X)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

מצאו התפלגות ותוחלת של:  $Y = X^2$ .

## שאלות:

- (1) מסובבים רולטה עליה המספרים 1 עד 4. יהיה  $X$  המספר שהתקבל לאחר סיבוב הרולטה. התפלגות  $X$  היא כדלהלן:

$X$	4	3	2	1
$P(X)$	0.3	0.4	0.2	0.1

א. חשבו את:  $E(X)$ ,  $E\left(\frac{1}{X}\right)$ .

ב. האם:  $E\left(\frac{1}{X}\right) = \frac{1}{E(X)}$ ?

- (2) יהי  $X$  משתנה מקרי בעל פונקציית ההסתברות הבאה:

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

חשבו את התוחלת של:

א.  $X^2$ .

ב.  $2^X$ .

- (3) להלן פונקציית הסתברות של משתנה מקרי כלשהו:  $P(X = k) = \frac{k}{A}$ ,  $k = 1, 2, \dots, 4$ .

א. מצאו את ערכו של  $A$ .

ב. חשבו את:  $E\left([X - E(X)]^2\right)$ .

- (4) בכל יום משחק ערן משחק יחיד בכל אחת מהאפליקציות הבאות: TWODOTS ו-PIANOTILES. בכל אחד מהמשחקים ישנם שלבים שיש לעבור. משחק בודד מסתיים בהצלחה אם ערן עבר את שלב, ובכישלון אם ערן לא עבר את השלב. ההסתברות שבאפליקציית TWODOTS ערן יעבור שלב היא 0.6 בכל יום. ההסתברות שבאפליקציית PIANOTILES ערן יעבור שלב היא 0.35 בכל יום. נניח שמעבר שלב בכל אחד מהמשחקים בלתי תלוי במשחק אחר. נסמן ב- $W$  את מספר המשחקים שערן יעבור שלב בהם מחר.

א. חשבו את  $E(W)$ .

ב. חשבו את  $E(W^3)$ .

(5) יהי  $X$  משתנה מקרי בדיד עם תוחלת ושונויות סופיים:  $Y = aX + b$ , כאשר  $a \neq 0$ .  
 $a, b$  הינם פרמטרים. יש להוכיח ש:  $E(Y) = aE(X) + b$ ,  $V(Y) = a^2 \cdot V(X)$ .

(6) אלעד צופה בסדרה בת 6 פרקים. 3 פרקים מתוך ה-6 הם פרקים שצולמו בישראל ו-3 פרקים אחרים צולמו בבולגריה. פרק אחד מבין הפרקים שצולמו בבולגריה מצולם כולו ביער. אלעד צופה בפרקי הסדרה בסדר אקראי, עד אשר הוא מגיע לפרק שצולם ביער בבולגריה. נגדיר את  $W$  כמספר הפרקים שצולמו בבולגריה שבהם יצפה אלעד.

א. מהי התפלגות  $W$ ?

ב. חשבו:  $E(W^3)$ .

(7) למיקה יש 20 חולצות ו-3 מגירות. כאשר מיקה מסדרת את 20 החולצות במגירות היא בוחרת עבור כל חולצה, באופן מקרי ובלתי תלוי בחולצות האחרות, את המגירה אליה תכניס את החולצה (כל אחת מהמגירות יכולה להכיל את כל החולצות).

נסמן ב- $X$  את מספר המגירות המכילות בדיוק 10 חולצות.

מצאו את התפלגות  $X$  ואת:  $E(\sqrt{X+2})$ .

(8) מטבע מוטל 10 פעמים.  $X =$  מספר הפעמים שהתקבלה התוצאה ראש.

א. בנו את פונקציית ההסתברות של  $X$ .

ב. הרווח במשחק הוא  $4^X$ . מצאו את התוחלת של הרווח במשחק.

$$\text{רמז: היעזרו בבינום של ניוטון: } (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$

### תשובות סופיות:

(1) א.  $E\left(\frac{1}{X}\right) = 0.4083$ ,  $E(X) = 2.9$ . ב. לא.

(2) א. 3.2. ב. 3.45.

(3) א. 10. ב. 1.

(4) א. 0.95. ב. 2.21.

(5) הוכחה.

(6) א.  $X \sim U(1,3)$ . ב. 12.

(7)  $E(\sqrt{X+2}) = 1.4659$ .

(8) א.  $X \sim B(10,0.5)$ . ב.  $2.5^{10}$ .

# סטטיסטיקה והסתברות 01 88165

פרק 31 - המשתנה המקרי הבדיד - טרנספורמציה לינארית

תוכן העניינים

1. כללי ..... 137

## המשתנה המקרי הבדיד – טרנספורמציה לינארית:

### רקע:

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

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

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

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

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

### שלבי העבודה:

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

דוגמה – הרולטה:

בהמשך לנתוני שאלת הרולטה נתון שעלות השתתפות במשחק 15 ₪. מהי התוחלת והשונויות של הרווח במשחק?

פתרון (בהקלטה):

$$\text{חישבנו קודם ש: } E(X) = 22.5 = \mu, \quad V(X) = 68.75 = \sigma^2$$

**שאלות:**

(1) סטודנט ניגש ל-5 קורסים הסמסטר. נניח שכל קורס שסטודנט מסיים מזכה אותו ב-4 נקודות אקדמאיות. חשבו את התוחלת והשונות של סך הנקודות שיצבור הסטודנט כאשר נתון שתוחלת מספר הקורסים שיסיים היא 3.5 עם שונות 2.

(2) תוחלת סכום הזכייה במשחק מזל הינה 10 עם שונות 3. הוחלט להכפיל את סכום הזכייה במשחק. עלות השתתפות במשחק הינה 12. מה התוחלת ומהי השונות של הרווח במשחק?

(3) תוחלת של משתנה מקרי הינה 10 וסטית 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- $V(X) = 3$ .  
 $Y$  הינו משתנה מקרי חדש, עבורו:  $Y = 7 - X$ . חשבו את:  $E(Y)$  ו- $V(Y)$ .

(5) אדם החליט לבטח את רכבו; שווי הרכב 100,000 ₪. להלן התביעות האפשריות והסתברותן: בהסתברות של 0.001 תהיה תביעה טוטאלוסט (כל שווי הרכב). בהסתברות של 0.02 תהיה תביעה בשווי מחצית משווי הרכב. בהסתברות של 5% תהיה תביעה בשווי רבע משווי הרכב. אחרת אין תביעה בכלל. החברה מאפשרת תביעה אחת בשנה. נסמן ב- $X$  את גובה התביעה השנתית, באלפי ₪.  
א. בנו את פונקציית ההסתברות של  $X$ .  
ב. חשבו את התוחלת והשונות של גובה התביעה.  
ג. פרמיית הביטוח היא 4,000 ₪.  
מהי התוחלת ומהי השונות של רווח חברת הביטוח לביטוח הרכב הני"ל?

(6) יהי  $X$  מספר התשובות הנכונות במבחן בו 10 שאלות. פונקציית ההסתברות של  $X$  נתונה בטבלה הבאה:

10	9	8	7	6	5	$X$
		0.3	0.2	0.2	0.1	$P(X)$

כמו כן, נתון שצפי מספר התשובות הנכונות בבחינה הוא 7.35.

- א. השלימו את פונקציית ההסתברות.
  - ב. חשבו את השונות מספר התשובות הנכונות בבחינה.
  - ג. הציון בבחינה מחושב באופן הבא:
- כל שאלה נכונה מזכה ב-10 נקודות. לכל שאלה שגויה, מופחתת נקודה. מהי התוחלת ומה השונות של הציון בבחינה?

(7) להלן פונקציית הסתברות של משתנה מקרי כלשהו:  $P(X = k) = \frac{k}{A}$ ,  $k = 1, 2, \dots, 4$ .

- א. מצא את ערכו של  $A$ .  
 ב. חשב את התוחלת והשונות של המשתנה הנחקר.  
 ג. חשב את:  $E(X^3)$ .  
 ד. חשב את התוחלת והשונות של המשתנה הבא:  $\frac{X}{2} - 4$ .

### 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:  $85,727.5^2$ .  
 (6) ב.  $V(X) = 1.8275$ .  
 (7) א.  $A = 10$ . ב.  $E(X) = 3$ ,  $V(X) = 1$ . ג.  $E(X^3) = 35.4$ ,  $V(X^3) = 616.84$ . ד.  $E(Y) = -2.5$ ,  $V(Y) = 0.25$ .

# סטטיסטיקה והסתברות 01 88165

פרק 32 - תוחלת ושונות של סכום משתנים מקריים

תוכן העניינים

1. כללי ..... 140

## תוחלת ושונות של סכום משתנים מקריים:

---

**רקע:**

אם:  $X_1, X_2, \dots, X_n$  משתנים מקריים אזי:

$$E(T) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

אם:  $X_1, X_2, \dots, X_n$  משתנים מקריים בלתי תלויים בזוגות, אזי:

$$V(T) = V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

**דוגמה:**

אדם משחק בשני משחקי מזל בלתי תלויים. תוחלת סכום הזכייה של המשחק הראשון היא 7 עם סטיית תקן 3. תוחלת סכום הזכייה של המשחק השני היא 2- עם סטיית תקן 4.

מה התוחלת ומהי השונות של סכום הזכייה הכולל של שני המשחקים יחד?

## שאלות:

(1) הרווח ממניה א' הוא עם תוחלת של 5 ושונות 10. הרווח ממניה ב' הוא עם תוחלת של 4 ושונות. ידוע שההשקעות של שתי המניות בלתי תלויות זו בזו. מה התוחלת והשונות של הרווח הכולל מהשקעה בשתי המניות יחד?

(2)  $X$  ו- $Y$  הם משתנים בלתי תלויים, סטיית התקן של  $X$  היא 3. סטיית התקן של  $Y$  היא 4. מהי סטיית התקן של  $X+Y$ ?

(3) אדם משחק בשני משחקי מזל בלתי תלויים זה בזה:  $X$  - סכום הזכייה במשחק הראשון.  $Y$  - סכום הזכייה במשחק השני. נתון:

$$\sigma(X) = 3, \quad E(x) = 10$$

$$\sigma(Y) = 4, \quad E(y) = 12$$

מהי התוחלת ומהי סטיית התקן של סכום הזכייה בשני המשחקים?

(4) ברולטה הסיכוי לזכות ב-30 ש"ח הוא חצי, ב-10 ש"ח רבע וכך גם ב-20 ש"ח. מה היא התוחלת והשונות של סכום הזכייה הכולל לאדם המשחק ברולטה 4 פעמים?

(5) נתון משתנה מקרי בעל פונקציית ההסתברות הבאה:

$$K = 2, 3, 4, 5, \quad P(X = K) = \frac{A}{K-1}$$

$$0 \leq A < 1$$

מצאו את ערכו של  $A$ .

א. חשבו את התוחלת והשונות של  $X$ .

ב. נלקחו  $n$  משתנים מקריים בלתי תלויים מההתפלגות הנ"ל. בטאו באמצעות  $n$  את תוחלת והשונות של סכום המשתנים.

**תשובות סופיות:**

- (1) תוחלת: 9, שונות: 15.
- (2) 5.
- (3) תוחלת: 22, שונות: 5.
- (4) תוחלת: 90, שונות: 275.
- (5) א.  $A = \frac{12}{25} = 0.48$ . ב. תוחלת: 2.92, שונות: 1.1136.  
ג. תוחלת: 2.92, שונות:  $1.1136n$ .

# סטטיסטיקה והסתברות 01 88165

פרק 33 - התפלגויות בדידות מיוחדות - התפלגות בינומית

תוכן העניינים

1. כללי ..... 143

## התפלגויות בדידות מיוחדות – התפלגות בינומית:

### רקע:

נגדיר את המושג ניסוי ברנולי:  
ניסוי ברנולי הנו ניסוי שיש לו שתי תוצאות אפשריות: "הצלחה" ו"כישלון".  
למשל מוצר פגום או תקין, אדם עובד או מובטל, עץ או פלי בהטלת מטבע וכדומה.  
בהתפלגות בינומית חוזרים על אותו ניסוי ברנולי  $n$  פעמים באופן בלתי תלוי זה בזה.  
מגדירים את  $X$  להיות מספר ההצלחות שהתקבלו בסך הכול. נסמן ב- $p$  את הסיכוי  
להצלחה בניסוי בודד, וב- $q$  את הסיכוי לכישלון בניסוי בודד.  
ואז נגיד ש:  $X \sim B(n, p)$ .

### פונקציית ההסתברות של $X$ :

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

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

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

**תוחלת:**  $E(X) = np$

**שונות:**  $V(X) = npq$

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

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

דוגמה (פתרון בהקלטה):

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

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

**שאלות:**

- (1) במדינה 10% מהאוכלוסייה מובטלת. נבחרו 5 אנשים באקראי מאותה אוכלוסייה. נגדיר את  $X$  להיות מספר המובטלים שהתקבלו במדגם.
- מהי ההתפלגות של  $X$  ?
  - מה ההסתברות שיהיה בדיוק מובטל אחד?
  - מה ההסתברות שכולם יעבדו במדגם?
  - מה ההסתברות 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. נגדיר את  $X$  כמספר האנשים שנדגמו 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 התוחלת ומה סטיית התקן של מספר האנשים שנדגמו ולהם סמארט-פון?
- (3) בבית הימורים יש שורה של 6 מכונות מזל מאותו סוג. משחק במכונת מזל כזו עולה 5 ₪. ההסתברות לזכות ב-20 ₪ בכל אחת מהמכונות היא 0.1 וההסתברות להפסיד את ההשקעה היא 0.9 בכל מכונה. מהמר נכנס לבית ההימורים ומכניס 5 ₪ לכל אחת מ-6 המכונות.
- מה ההסתברות שיפסיד בכל המכונות?
  - מה ההסתברות שיזכה בדיוק בשתי מכונות?
  - מה ההסתברות שיזכה ביותר כסף מה-30 ₪ שהשקיע?
  - מהן התוחלת וסטיית התקן של הרווח נטו של המהמר (הזכיות בניכוי ההשקעה)?
- (4) במדינה מסוימת התפלגות ההשכלה בקרב האוכלוסייה מעל גיל 30 היא כזו:

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

- נבחרו 20 אנשים אקראיים מעל גיל 30.
- מה ההסתברות ש-5 מהם אקדמאים?
  - מה התוחלת של מסי' בעלי ההשכלה הנמוכה?

- (5) במכללה מסוימת 20% מהסטודנטים גרים בת"א. מבין הסטודנטים שגרים בת"א 30% מגיעים ברכבם, ומבין הסטודנטים שלא גרים בת"א 50% מגיעים ברכבם למכללה.
- א. השומר בשער המכללה בודק לכל סטודנט את תיקו בהיכנסו למכללה. מה ההסתברות שבקרב 5 סטודנטים שנבדקו ע"י השומר רק 1 מתוכם הגיע למכללה ברכבו?
- ב. בהמשך לסעיף הקודם מה ההסתברות שרוב הסטודנטים בקרב ה-5 הגיעו למכללה ברכבם?
- (6) במבחן אמריקאי 20 שאלות. סטודנט ניגש למבחן והסיכוי שהוא יודע שאלה כלשהי הוא 0.8. אם הוא לא יודע הוא מנחש את התשובה. לכל שאלה 4 תשובות אפשריות שרק אחת מהן נכונה.
- א. מה הסיכוי לענות על שאלה מסוימת נכון?
- ב. מה הסיכוי שיענה נכונה על בדיוק 16 שאלות?
- ג. על כל שאלה שענה נכון התלמיד מקבל 5 נקודות, על כל שאלה ששגה מופחתת נקודה, מה התוחלת ומהי השונות של ציון התלמיד?
- (7) 5% מקו היצור פגום. המוצרים נארזים בתוך קופסת קרטון. בכל קופסא 10 מוצרים שונים. הקופסאות נארזות בתוך מכולה. בכל מכולה 20 קופסאות.
- א. מה ההסתברות שבקופסא אקראית לפחות מוצר פגום אחד?
- ב. מה התוחלת ומהי סטיית התקן של מספר הקופסאות במכולה בהן לפחות מוצר פגום אחד?
- (8) מטבע הוגן מוטל 5 פעמים. נגדיר את  $X$  כמספר הפעמים שהתקבל עץ. חשבו את:  $E(X^2)$ .

### תשובות סופיות:

- (1) א.  $X \sim B(n=5, p=0.1)$     ב. 0.32805    ג. 0.59049  
 ד. 0.0729    ה. 0.40954    ו. תוחלת: 0.5, שונות: 0.45
- (2) א. 0.2335    ב. 0.1493    ג. 0.1493    ד. תוחלת: 7, סטיית תקן: 1.449
- (3) א. 0.5314    ב. 0.0984    ג. 0.1143    ד. תוחלת: -18, סטיית תקן: 14.697
- (4) א. 0.1789    ב. 2
- (5) א. 0.1956    ב. 0.4253
- (6) א. 0.85    ב. 0.182    ג. תוחלת: 82, שונות: 91.8
- (7) א. 0.401    ב. תוחלת: 8.025, סטיית תקן: 2.193
- (8) 7.5

# סטטיסטיקה והסתברות 01 88165

פרק 34 - התפלגויות בדידות מיוחדות - התפלגות גיאומטרית

תוכן העניינים

1. כללי ..... 147

## התפלגויות בדידות מיוחדות – התפלגות גיאומטרית:

### רקע:

חוזרים באופן בלתי תלוי על אותו ניסוי ברנולי.  
 $X$  – מוגדר להיות מספר הניסויים שבוצעו עד ההצלחה הראשונה, כולל.  
 נסמן ב- $p$  את הסיכוי להצלחה בניסויי בודד וב- $q$  את הסיכוי לכישלון בניסוי בודד.  
 $X \sim G(p)$

**פונקציית ההסתברות:**  $P(X = k) = pq^{k-1}$ ,  $k = 1, 2, \dots, \infty$ .

**תוחלת:**  $E(X) = \frac{1}{p}$

**שונות:**  $V(X) = \frac{q}{p^2}$

### תכונות חשובות:

אם  $X$  מתפלג על פי התפלגות גיאומטרית, אזי  $X$  הוא בעל תכונת חוסר זיכרון,  
 דהיינו,  $P(X > k) = q^k \cdot P(X = (n+k) / X > k) = P(X = n)$ .

דוגמה (פתרון בהקלטה):

בכד 10 כדורים ש-3 מהם ירוקים. אדם מוציא באקראי כדור אחר כדור עד שבידו כדור ירוק. ההוצאה היא עם החזרת הכדור לכד בכל פעם מחדש.

- מהי ההתפלגות של מספר הכדורים שהוצאו?
- מה ההסתברות שהוצאו בדיוק 5 כדורים?
- מה ההסתברות שהוצאו יותר מ-5 כדורים?
- אם הוצאו יותר מ-3 כדורים. מה הסיכוי שהוצאו בדיוק 5 כדורים?
- מה התוחלת וסטיית התקן של מספר הכדורים שהוצאו?

## שאלות:

- (1) קו ייצור המוני מייצר מוצרים כך ש-5% מהם פגומים. איש בקרת איכות דוגם באופן מקרי מוצרים מקו הייצור עד אשר בידו מוצר פגום. חשבו את ההסתברויות הבאות:
- שידגום 3 מוצרים.
  - שידגום 4 מוצרים.
  - שידגום 5 מוצרים.
  - שידגום יותר מ-7 מוצרים.
  - שידגום לא פחות מ-8 מוצרים.
- (2) צילום שמבוצע במכון הרנטגן "X-RAY" יתקבל תקין בהסתברות של 0.9. אדם נכנס למכון כדי להצטלם, והוא ייצא מהמכון רק כאשר יש בידו תצלום תקין.
- מה ההסתברות שיצטלם בסך הכול 3 פעמים?
  - מה ההסתברות שהצטלם יותר מ-4 פעמים?
  - מה התוחלת ומה השונות של מספר הצילומים שייבצע?
  - כל צילום עולה למכון 50 ₪. אדם משלם על צילום תקין 100 ₪. מה התוחלת ומה השונות של רווח המכון מאדם שהגיע להצטלם?
- (3) מטילים מטבע עד אשר מתקבלת התוצאה "עץ".
- מה ההסתברות להטיל את המטבע לכל היותר 10 פעמים?
  - מה ההסתברות להטיל את המטבע לכל היותר 5 פעמים, אם ידוע שהמטבע הוטל לפחות 3 פעמים?
  - אם ידוע שבשתי ההטלות הראשונות התקבלה התוצאה "פלי", מה ההסתברות שהאדם הטיל את המטבע 7 פעמים?
  - מה תוחלת מספר הפעמים שהתקבלה התוצאה "פלי"?
- (4) 30% מהמכוניות בארץ הן בצבע לבן. בכל יום נכנסות לחניון כשלהו 10 מכוניות אקראיות.
- מה ההסתברות שביום מסוים בדיוק מחצית מהמכוניות בחניון יהיו לבנות?
  - מה תוחלת מספר הימים שיעברו מהיום עד שלראשונה מחצית מהמכוניות בחניון יהיו לבנות?

- (5) אדם משחק במשחק מזל עד אשר הוא מפסיד. הצפי הוא שישחק את המשחק 10 פעמים. מה הסיכוי להפסיד במשחק בודד?
- א. מה ההסתברות שישחק את המשחק בדיוק 6 פעמים?  
 ב. מה ההסתברות שישחק את המשחק לכל היותר 12 פעמים?  
 ג. ידוע שהאדם שיחק את המשחק יותר מ-6 פעמים.  
 מה ההסתברות שישחק את המשחק בדיוק 10 פעמים?  
 ד. מהי סטיית התקן של מספר הפעמים שישחק את המשחק?
- (6) במאפייה מייצרים עוגות גבינה ועוגות שוקולד שנארזות באריזות אטומות. 40% מהעוגות הן עוגות גבינה והיתר שוקולד. התווית על האריזה מודבקת בשלב מאוחר יותר של הייצור. אדם נכנס למפעל ובוחר באקראי עוגה.
- א. מה ההסתברות שייאלץ לבחור 5 עוגות עד שקיבל עוגות שוקולד?  
 ב. אם הוא דגם פחות מ-7 עוגות עד שיקבל עוגת שוקולד, מה ההסתברות שבפועל הוא דגם יותר מ-4 עוגות?  
 ג. האדם דוגם עוגות עד אשר הוא מוצא עוגת שוקולד. ידוע שעוגת גבינה עולה ליצרן 50 שקלים ועוגת שוקולד 30 שקלים. מהי התוחלת ומהי השונות של עלות הייצור הכוללת של העוגות שדגם?  
 ד. בהמשך לסעיף הקודם, מהי התוחלת ומהי סטיית התקן של מספר עוגות הגבינה שדגם האדם?

### תשובות סופיות:

- (1) א. 0.04512    ב. 0.0428    ג. 0.0407    ד. 0.6983    ה. 0.6983
- (2) א. 0.009    ב. 0.0001    ג. תוחלת: 1.111, שונות: 0.1234  
 ד. תוחלת: 44.4, שונות: 308.5
- (3) א. 0.999    ב. 0.875    ג. 0.03125    ד. 1
- (4) א. 0.1029    ב. 9.72
- (5) א. 0.06    ב. 0.7176    ג. 0.0729    ד. 9.487
- (6) א. 0.015    ב. 0.0215    ג. תוחלת:  $63\frac{1}{3}$ , שונות:  $2777\frac{7}{9}$   
 ד. תוחלת  $\frac{2}{3}$ , שונות 1.054

# סטטיסטיקה והסתברות 01 88165

פרק 35 - התפלגויות בדידות מיוחדות - התפלגות אחידה

תוכן העניינים

1. כללי ..... 150

## התפלגויות בדידות מיוחדות – התפלגות אחידה:

**רקע:**

התפלגות אחידה הינה התפלגות שבה לכל תוצאה יש את אותה הסתברות. הערכים המתקבלים בהתפלגות הם החל מ- $a$  ועד  $b$  בקפיצות של אחד.  $X \sim U(a, b)$ .

**פונקציית ההסתברות:**  $P(X = K) = \frac{1}{b-a+1}$  ,  $K = a, a+1, \dots, b$

**תוחלת:**  $E(X) = \frac{a+b}{2}$

**שונות:**  $V(X) = \frac{(b-a+1)^2 - 1}{12}$

דוגמה (פתרון בהקלטה):

אדם בוחר מספר אקראי בין 1 ל-100 כולל. מהי פונקציית ההסתברות של המספר ומה הצפי שלו?

## שאלות:

- (1) במשחק הלוטו 45 כדורים ממוספרים מ-1 ועד 45. נתבונן במשתנה  $X$  - המספר של הכדור הראשון שנשלף על ידי המכונה.
- חשבו את  $P(X = 2)$ .
  - חשבו את  $P(X \leq 30)$ .
  - חשבו את  $P(X > 4 | X \leq 10)$ .
  - חשבו את  $P(X = k)$ .
- (2) קוסם מבקש לבחור מספר שלם אקראי בין 1 ל-100.
- בהנחה שאין כאן מניפולציות של הקוסם, מהי התוחלת ומהי סטיית התקן של המספר שיבחר?
  - הקוסם ביקש משישה אנשים לבחור מספר:
    - מה ההסתברות ששלושה מהם יבחרו מספר הגדול מ-80?
    - מה התוחלת ומהי סטיית התקן של סכום המספרים שהאנשים בחרו?
- (3) יהי  $X$  התוצאה בהטלת קובייה.
- מהי ההתפלגות של  $X$ ?
  - מה התוחלת של  $X$ ?
  - קובייה הוטלה 4 פעמים. מה התוחלת ומה השונות של סכום התוצאות ב-4 ההטלות?
- (4) בכד 10 כדורים שרק אחד בצבע אדום. כדורים הוצאו ללא החזרה עד שהתקבל הכדור האדום. מה התוחלת ומהי השונות של מספר הכדורים שהוצאו?
- (5) יש לבחור מספר אקראי בין 1 ל-50, כולל.
- מה הסיכוי שהמספר 4 יבחר?
  - מה הסיכוי שהמספר שיבחר גדול מ-20?
  - אם נבחר מספר גדול מ-20, מה ההסתברות שהוא קטן מ-28?
- (6) הוכיחו שאם:  $X \sim U(a, b)$ , אז מתקיים ש:  $E(X) = \frac{a+b}{2}$ .

## תשובות סופיות:

- (1) א.  $\frac{1}{45}$     ב.  $\frac{30}{45}$     ג. 0.6
- (2) א. תוחלת: 50.5, סטיית תקן: 28.87.  
 ב. i. 0.08192    ii. תוחלת: 303, סטיית תקן: 70.71
- (3) א.  $X \sim U(1,6)$     ב. 3.5    ג. תוחלת: 14, שונות: 11.66
- (4) תוחלת: 5.5, שונות: 8.25
- (5) א.  $\frac{1}{50}$     ב.  $\frac{30}{50}$     ג.  $\frac{7}{30}$
- (6) שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 36 - התפלגויות בדידות מיוחדות- התפלגות פואסונית

תוכן העניינים

1. כללי..... 153

## התפלגויות בדידות מיוחדות – התפלגות פואסונית:

### רקע:

התפלגות פואסונית היא התפלגות שמאפיינת את מספר האירועים שמתרחשים ביחידת זמן.

$\lambda$  - פרמטר המאפיין את ההתפלגות הנ"ל. הפרמטר מייצג את קצב האירועים ביחידת זמן. כלומר, כמה אירועים בממוצע קורים ביחידת זמן:  $X \sim pois(\lambda)$ . התפלגות פואסונית חייבת להופיע כנתון בשאלה ולכן לא יהיה צורך לזהותה.

### פונקציית ההסתברות של ההתפלגות הפואסונית נתונה:

$$P(X = K) = \frac{e^{-\lambda} \cdot \lambda^K}{K!}, \quad K = 0, 1, 2, \dots, \infty$$

### התוחלת והשונות של ההתפלגות:

$$E(X) = V(X) = \lambda$$

### תכונות מיוחדות של ההתפלגות:

- בהתפלגות הזו הפרמטר  $\lambda$  פרופורציונלי לאינטרוול הזמן שעליו דנים.
- אינטרוולי זמן לא חופפים בלתי תלויים זה בזה.

דוגמה (פתרון בהקלטה):

במוקד טלפוני מתקבלות פניות בקצב של 5 פניות לדקה. מספר הפניות בדקה מתפלג פואסונית.

- מה ההסתברות שבדקה כלשהי תתקבל פניה 1?
- מה ההסתברות שבשתי דקות יגיעו 12 פניות?
- מה ההסתברות שבדקה אחת תגיע פניה 1 ובשתי דקות שלאחר מכן 12 פניות?
- מה התוחלת וסטיית התקן של מספר הפניות בדקה?

## שאלות:

- (1) במוקד טלפוני מתקבלות פניות בקצב של 5 פניות לדקה. מספר הפניות בדקה מתפלג פואסונית.
- מה ההסתברות שבדקה תתקבל פניה 1?
  - מה ההסתברות שבדקה תתקבל לפחות פניה 1?
  - מה ההסתברות שבדקה יתקבלו לכל היותר 2 פניות?
  - מה שונות מספר הפניות בדקה?
- (2) מספר הטעויות לעמוד בעיתון מתפלג פואסונית עם ממוצע של 4 טעויות לעמוד. בחלק מסוים של עיתון ישנם 5 עמודים.
- מה ההסתברות שבחלק זה ישנן בדיוק 18 טעויות?
  - אם בעמוד הראשון אין טעויות, מה ההסתברות שבסך הכול בכל החלק ישנן 15 טעויות?
  - אם בחלק של העיתון נמצאו בסך הכול 18 טעויות, מה ההסתברות ש-5 מהן בעמוד הראשון?
- (3) מספר תאונות הדרכים הקטלניות במדינת ישראל מתפלג פואסונית עם סטיית תקן של 2 תאונות לשבוע.
- מה תוחלת מספר התאונות בשבוע?
  - מהי ההסתברות שבחודש (הניחו שבחודש יש 4 שבועות) יהיה בדיוק שבוע אחד בו יהיו 3 תאונות דרכים קטלניות?
- (4) לחנות AM:PM השכונתית מספר הלקוחות שנכנסים מתפלג פואסונית עם ממוצע של 2 לקוחות לדקה.
- מה ההסתברות שבדקה כלשהי יהיו בדיוק 3 לקוחות?
  - מה ההסתברות שבדקה כלשהי יגיח לפחות לקוח אחד?
  - מה ההסתברות שבדקה כלשהי יהיו לכל היותר שני לקוחות?
  - מהי התוחלת ומה סטיית התקן של מספר הלקוחות שנכנסים לחנות בדקה?
- (5) מספר הלידות בבית חולים מתפלג פואסונית עם תוחלת של 8 לידות ביום.
- מה ההסתברות שביום א' נולדו 10 תינוקות וביום ב' נולדו 7 תינוקות?
  - מיילדת עובדת במשמרות של 8 שעות. מה ההסתברות שבמשמרת שלה נולדו 3 תינוקות?
  - מהי התוחלת של מספר הימים בשבוע בהם נולדים ביום עשרה תינוקות?

- 6) במערכת אינטרנט לתשלום חשבונות, מספר החשבונות המשולמים בשעה מתפלג פואסונית עם תוחלת של 30.
- א. כמה שעות צפויות לעבור עד אשר תתקבל שעה עם בדיוק 33 חשבונות?
- ב. בין השעה 08:00 ל-08:20 היו 18 חשבונות, מה ההסתברות שבין 08:00 ל-08:10 היו בדיוק 6 חשבונות?

### תשובות סופיות:

- |           |           |           |                              |   |
|-----------|-----------|-----------|------------------------------|---|
| א. 0.0337 | ב. 0.9933 | ג. 0.1246 | ד. 0.5                       | 1 |
| א. 0.084  | ב. 0.099  | ג. 0.151  |                              | 2 |
| א. 4      | ב. 0.407  |           |                              | 3 |
| א. 0.1804 | ב. 0.8647 | ג. 0.6767 | ד. תוחלת: 2, סטיית תקן: 1.41 | 4 |
| א. 0.0139 | ב. 0.2196 | ג. 0.6948 |                              | 5 |
| א. 16.7   | ב. 0.0708 |           |                              | 6 |

# סטטיסטיקה והסתברות 01 88165

פרק 37 - התפלגויות בדידות מיוחדות-התפלגות היפרגאומטרית

תוכן העניינים

1. כללי ..... 156

## התפלגויות בדידות מיוחדות – התפלגות היפרגאומטרית:

### רקע:

נתונה אוכלוסייה המכילה  $N$  פריטים, מתוכם  $D$  פריטים בעלי תכונה מסוימת – פריטים אלה נקראים "מיוחדים". בוחרים מאותה אוכלוסייה  $n$  פריטים ללא החזרה.  $X$  מוגדר להיות מספר הפריטים ה"מיוחדים" שנדגמו. משתנה מקרי היפרגאומטרי עם הפרמטרים  $(N, D, n)$  יסומן על ידי:  $X \sim H(N, D, n)$ .

$$P(X = k) = \frac{\binom{D}{k} \binom{N-D}{n-k}}{\binom{N}{n}} \quad \text{פונקציית ההסתברות של ההתפלגות:}$$

$$E(X) = n \cdot \frac{D}{N} \quad \text{התוחלת של ההתפלגות:}$$

$$V(X) = n \cdot \frac{D}{N} \cdot \left(1 - \frac{D}{N}\right) \cdot \frac{N-n}{N-1} \quad \text{השונות של ההתפלגות:}$$

דוגמה (הפתרון בהקלטה):

בכתה 40 תלמידים, שמתוכם 10 בנות והשאר בנים. בוחרים קבוצה של ארבעה תלמידים שיסעו למשלחת.

- א. כיצד מספר הבנים במשלחת מתפלג?
- ב. מה התוחלת ומהי השונות של מספר הבנים במשלחת?
- ג. מה הסיכוי שבמשלחת יהיו 3 בנים?

## שאלות:

- (1) בכד 5 כדורים אדומים ו-4 כדורים ירוקים. מוציאים באקראי שלושה כדורים מהכד.  
 א. בנו את פונקציית ההסתברות של מספר הכדורים האדומים שהוצא בטבלה.  
 ב. חשבו את התוחלת והשונות של מספר הכדורים האדומים שהוצאו, פעם מתוך פונקציית ההסתברות ופעם מתוך הנוסחאות להתפלגות היפרגאומטרית.  
 ג. מה הייתה התוחלת והשונות של מספר הכדורים האדומים אם ההוצאה הייתה עם החזרה?
- (2) בחידון 10 שאלות משלושה תחומים שונים: 3 בתחום הספורט, 4 בתחום הבידור והיתר בתחום המדעים. משתתף בחידון שולף באקראי 4 שאלות.  
 נגדיר את  $X$  להיות מספר השאלות מתחום הספורט שנשלפו.  
 א. בנו את פונקציית ההסתברות של  $X$  בנוסחה (לא בטבלה).  
 ב. מה התוחלת וסטיית התקן של  $X$ ?  
 ג. חשבו את ההסתברות הבאה:  $P(X = 2 | X > 1)$ .
- (3) נדגמו 6 אנשים מתוך אוכלוסייה שבה 60% בעלי רישיון נהיגה. אנו מתעניינים במספר האנשים שנדגמו עם רישיון נהיגה. זהו בסעיפים הבאים את ההתפלגות, וחשבו לכל התפלגות את התוחלת והשונות:  
 א. האוכלוסייה גדולה מאד.  
 ב. האוכלוסייה בת 10 אנשים.
- (4) בארגון עובדים 7 מהנדסים, 3 טכנאים ו-5 הנדסאים. בוחרים באופן מקרי משלחת של 4 עובדים לכנס במדריד.  
 א. מהי ההסתברות שייבחרו רק מהנדסים?  
 ב. מה תוחלת מספר הטכנאים שייבחרו?

## תשובות סופיות:

(1) א. ב. תוחלת:  $1\frac{2}{3}$ , שונות:  $\frac{5}{9}$ .

3	2	1	0	x
$\frac{10}{84}$	$\frac{40}{84}$	$\frac{30}{84}$	$\frac{4}{84}$	$P(x)$

ג. תוחלת:  $1\frac{2}{3}$ , שונות:  $\frac{20}{27}$ .

(2) א.  $\frac{\binom{3}{k} \cdot \binom{7}{4-k}}{\binom{10}{4}}$ . ב. תוחלת: 1.5, סטיית תקן: 0.748. ג. 0.9.

(3) א. תוחלת: 3.6, שונות: 1.44. ב. תוחלת: 3.6, שונות: 0.64.

(4) א. 0.0256. ב. 0.8.

# סטטיסטיקה והסתברות 01 88165

פרק 38 - התפלגויות בדידות מיוחדות - התפלגות בינומית שלילית

תוכן העניינים

1. כללי ..... 159

## התפלגויות בדידות מיוחדות – התפלגות בינומית שלילית:

**רקע:**

בהתפלגות זו חוזרים על אותו ניסוי ברנולי בזה אחר זה באופן בלתי תלוי עד אשר מצליחים בפעם ה- $r$ .  $X =$  מספר החזרות עד שהתקבלו  $r$  הצלחות:  $X \sim NB(r, p)$ .

**פונקציית ההסתברות:**  $P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$ ,  $k = r, r+1, \dots, \infty$ .

**תוחלת:**  $E(X) = \frac{r}{p}$

**שונות:**  $V(X) = \frac{r(1-p)}{p^2}$

דוגמה (פתרון בהקלטה):

קובייה מוטלת עד שמקבלים 3 פעמים תוצאה שגדולה מ-4.

א. מה הסיכוי להטיל את הקובייה 6 פעמים?

ב. מה תוחלת ושונות מספר הפעמים שנטיל את הקובייה?

## שאלות:

- (1) בכד 4 כדורים שחורים ו-6 כדורים לבנים. כדור מוצא באקראי פעם אחר פעם ומוחזר בין הוצאה להוצאה. נסמן ב- $X$  את מספר הכדורים שהוצאו עד שהתקבלו 2 כדורים לבנים בסך הכול (לא בהכרח ברצף).
- חשבו את  $P(X = 2)$ .
  - חשבו את  $P(X = 3)$ .
  - חשבו את  $P(X = 4)$ .
  - חשבו את  $P(X = k)$ .
- (2) הסיכוי לזכות במשחק מזל הוא 0.4. אדם משחק במשחק ומפסיק ברגע שהוא ניצח פעמיים (לא בהכרח ברצף).
- מה הסיכוי שישחק פעמיים?
  - מה הסיכוי שישחק 3 פעמים?
  - מה הסיכוי שישחק 4 פעמים?
  - מה הסיכוי שישחק 5 פעמים?
  - מה הסיכוי שישחק  $k$  פעמים?
- (3) הראו שההתפלגות הגאומטרית היא מקרה פרטי של ההתפלגות הבינומית השלילית.
- (4) מטבע מוטל שוב ושוב עד שמתקבל שלוש פעמים עץ בסך הכול.
- בנו את פונקציית ההסתברות של מספר ההטלות הכולל.
  - מהי התוחלת ומהי השונות של מספר ההטלות הכולל?
  - חוזרים על התהליך שלעיל 5 פעמים. מה ההסתברות שפעמיים מתוך ה-5 חזרות נאלץ להטיל את המטבע בדיוק 4 פעמים?
- (5) יהיה  $X_i$  מספר החזרות עד ההצלחה הראשונה בניסיונות ברנוליים בלתי תלויים זה בזה, כאשר  $i = 1, 2, \dots, n$ .
- הוכיחו שהתוחלת והשונות של  $\sum_{i=1}^n X_i$  זהות לתוחלת והשונות של ההתפלגות הבינומית השלילית  $NB(n, p)$ .

### תשובות סופיות:

- (1) א. 0.36    ב. 0.288    ג. 0.0576    ד.  $0.6^2 \cdot 0.4^{k-2}$
- (2) א. 0.16    ב. 0.192    ג. 0.1728    ד. 0.13824    ה.  $0.4^2 \cdot 0.6^{k-2}$
- (3) שאלת הוכחה.
- (4) ב. תוחלת: 6, שונות: 6.    ג. 0.1886
- (5) שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 39 - המשתנה המקרי הבדיד - שאלות מסכמות

תוכן העניינים

1. כללי ..... 162

## המשתנה המקרי הבדיד – שאלות מסכמות:

### שאלות:

(1) נתון כי:  $X \sim B\left(4, \frac{1}{2}\right)$ ,  $Y \sim B\left(10, \frac{1}{4}\right)$ .

- א. חשבו את התוחלת וסטיית התקן של  $X$ .
- ב.  $W = 2X - 4$ , חשבו את התוחלת וסטיית התקן של  $W$ .
- ג.  $T = X + Y$ , חשבו את התוחלת של  $T$ .
- האם ניתן לדעת מה סטיית התקן של  $T$ ?
- (2) ערן משחק בקזינו בשתי מכונות הימורים, בכל מכונה משחק אחד (במכונה א' ובמכונה ב'). הסיכוי שלו לנצח במשחק במכונה א' הינו 0.08 והסיכוי שלו לנצח רק במכונה א' הינו 0.05. הסיכוי שלו להפסיד בשני המשחקים ביום מסוים הוא 0.88.

- א. מה הסיכוי שערן ניצח בשני המשחקים?
- ב. מה התוחלת ומהי השונות של מספר הניצחונות של ערן?
- ג. אם ערן נכנס לקזינו 5 פעמים ובכל פעם שיחק את שני המשחקים, מה ההסתברות שערן ינצח בשני המשחקים בדיוק פעם אחת מתוך חמשת הפעמים?
- (3) לאדם צרור מפתחות. בצרור 5 מפתחות אשר רק אחד מתאים לדלת של ביתו. האדם מנסה את המפתחות באופן מקרי. לאחר שניסה מפתח מסוים הוא מוציא אותו מהצרור כדי לא להשתמש בו שוב. נסמן ב- $X$  את מספר הניסיונות עד שהדלת תפתח.
- א. בנו את פונקציית ההסתברות של  $X$ .
- ב. חשבו את התוחלת והשונות של  $X$ .
- ג. כל ניסיון לפתוח הדלת אורך חצי דקה. מה התוחלת ומה השונות של הזמן הכולל לפתיחת הדלת?

- (4) מספר התקלות בשידור "ערוץ 1" מתפלג פואסונית בקצב של 6 תקלות ביום.
- א. מה ההסתברות שביום מסוים הייתה לפחות תקלה אחת?
- ב. מה ההסתברות שבשבוע (7 ימי שידור) יהיו בדיוק 6 ימים בהם לפחות תקלה אחת?
- ג. מה תוחלת מספר הימים שיעברו מהיום ועד היום הראשון בו לפחות תהיה תקלה אחת?

- (5) בעל חנות גדולה בקניון שם לב ש-40% מהמוצרים בחנותו נרכשים עבור ילדים, 35% נרכשים עבור נשים ו-25% נרכשים עבור גברים. 10% מהמוצרים הנרכשים עבור ילדים הם מתוצרת חוץ, וכך גם 60% מהמוצרים הנרכשים עבור נשים ו-50% מאלה הנרכשים עבור גברים.
- מה ההסתברות למכור בחנות זו מוצר מתוצרת חוץ?
  - יהי  $X$  מספר המוצרים שימכרו בחנות זו מפתחתה ביום א' בבוקר, עד שלראשונה יימכר מוצר מתוצרת הארץ (כולל). מהי פונקציית ההסתברות של  $X$ ?
  - מהי תוחלת מס' המוצרים מתוצרת חוץ שימכרו, עד שלראשונה יימכר מוצר מתוצרת הארץ?
  - ביום ב' נמכרו בחנות 7 מוצרים. מה ההסתברות שבדיוק 3 מהם הם מתוצרת חוץ?
- (6) חברת הפקות של סרטים הפיקה 3 סרטים, אשר הופקו לטלוויזיה המקומית. חברת ההפקות מנסה למכור את הסרטים הללו לחו"ל. להלן ההסתברויות למכירת הסרטים לחו"ל:
- הסרט "הצב" יימכר לחו"ל בסיכוי של 0.6.
  - הסרט "לעולם לא" יימכר לחו"ל בסיכוי של 0.7.
  - הסרט "מוות פתאומי" יימכר לחו"ל בסיכוי של 0.2.
- ידוע כי כל סרט עלה להפקה חצי מיליון שקלים. כמו כן, כל סרט הביא להכנסה של 200,000 שקלים מהטלוויזיה המקומית. במידה וסרט יימכר לחו"ל, כל סרט יימכר ב-600,000 שקלים.
- בנו את פונקציית ההסתברות של מספר הסרטים שיימכרו לחו"ל.
  - מהי התוחלת והשונות של מספר הסרטים שיימכרו?
  - מהי התוחלת ומהי סטיית התקן של הרווח (במאות אלפי שקלים) של חברת ההפקה?
- (7) במפעל מייצרים סוכריות כך ש-20% מהסוכריות בטעם תות. הייצור הוא ייצור המוני. שאר הסוכריות בטעמים שונים, השקיות נארזות ובכל שקית בדיוק 5 סוכריות.
- נבחרה שקית ונתון שבשקית פחות מ-3 סוכריות אדומות. מה ההסתברות שבשקית סוכריה אדומה אחת?
  - בוחרים באקראי שקית אחר שקית, במטרה למצוא שקית ללא סוכריות אדומות. מה ההסתברות שייאלצו לדגום יותר מ-6 שקיות?

8) מבחן בנוי משני חלקים: בחלק א' 10 שאלות ובחלק ב' 10 שאלות. תלמיד התכוון רק לחלק א' של המבחן ובחלק זה בכל שאלה יש סיכוי של 0.8 שיענה נכון, בחלק השני לכל שאלה יש 4 תשובות כשרק אחת נכונה. בחלק זה הוא מנחש את התשובות.

- מהי ההסתברות שבחלק הראשון הוא יענה נכון על 7 שאלות בדיוק?
- מהי ההסתברות שבחלק השני הוא יענה נכון על פחות מ-3 שאלות?
- מה התוחלת ומהי השונות של מספר התשובות הנכונות בחלק הראשון?
- מהי התוחלת ומהי השונות של מספר התשובות הנכונות בבחינה כולה?

9) יהי  $X$  משתנה מקרי המקיים:  $E(X) = 2$  וכן:  $V(X) = 1$ .

חשבו:  $E(X - 5)^2$ .

10) הסיכוי לעבור מבחן נהיגה הינו  $P$ . בוחרים באקראי ארבעה נבחנים.

ההסתברות ששניים מהם יעברו את מבחן הנהיגה גבוה פי  $\frac{8}{3}$  מהסיכוי שכל

הארבעה יעברו את המבחן.

א. חשבו את ערכו של  $P$ .

ב. תלמיד ניגש לבחינה עד אשר הוא עובר אותה.

מה ההסתברות שיעבור את מבחן הנהיגה רק במבחן הרביעי?

ג. מה ההסתברות שיאלץ לגשת לפחות לחמישה מבחנים בסך הכול?

ד. מה התוחלת ומהי השונות של מספר המבחנים שבהם יכשל?

ה. ידוע שהתלמיד ניגש לשלושה מבחנים ועדיין לא עבר. מה ההסתברות שבסופו של דבר יעבור במבחן הנהיגה החמישי?

11) רובוט נמצא בנקודה 0 על ציר המספרים. הרובוט מבצע  $n$  צעדים ובכל צעד

הוא נע בסיכוי  $P$ . ימינה ביחידה אחת ובסיכוי  $1 - P$  שמאלה ביחידה אחת.

נסמן ב- $X$  את המספר עליו עומד הרובוט לאחר  $n$  צעדים.

רשמו את פונקציית ההסתברות של  $X$  באמצעות  $P$  ו- $n$ .

12) למטבע יש סיכוי  $P$  לקבל את התוצאה ראש. מטילים את המטבע. אם יוצא

ראש בפעם הראשונה מפסידים שקל ומפסיקים את המשחק. אחרת,

ממשיכים לזרוק וזוכים במספר שקלים לפי מספר הפעמים שהטלנו את

המטבע מההתחלה ועד שהתקבל ראש.

א. בנו את פונקציית ההסתברות של רווח המשחק (באמצעות  $P$ ).

ב. בטאו את תוחלת הרווח באמצעות  $P$ .

ג. לאלו ערכי  $P$  המשחק כדאי?

- 13** מטבע הוגן מוטל עד שמתקבל  $m+1$  פעמים עץ. רשמו את פונקציית ההסתברות של מספר הפעמים שהתקבל פלי.
- 14** נתונות  $N$  מגירות ממוספרות מ-1 ועד  $N$ . מתוך  $n$  חולצות, יש לבחור באופן אקראי לכל חולצה מגירה. כל מגירה יכולה להכיל את כל החולצות. נגדיר את  $X_1$  - כמספר החולצות שהונחו במגירה מספר 1. נגדיר את  $X_N$  - כמספר החולצות שהונחו במגירה מספר  $N$ . חשבו את:  $V(X_1 + X_N)$ .
- 15**  $n$  אנשים יושבים במסעדה. בזמן שמגיע העת לשלם, האנשים פועלים לפי העיקרון הבא: כל אחד מהם מטיל מטבע הוגן עד אשר אחד מהם מקבל תוצאה שונה מכל השאר והוא זה שמשלם. מהי תוחלת מספר הסבבים שיבוצעו עד שימצא משלם?
- 16** הסיכוי לעבור בקורס מסוים את מועד א' הוא 0.7. סטודנט שנכשל במועד א' בהכרח ניגש למועד ב' ואז הסיכוי שלו לעבור אותו הוא 0.8. אם סטודנט נכשל במועד ב' הוא ניגש למועד מיוחד ואחרון. נתון שלמועד א' נגשו כל 20 הסטודנטים הרשומים לקורס. מהי התפלגות מספר הבחינות שיאלץ המרצה לחבר?
- 17** לקניון 3 כניסות שונות. בכל כניסה מספר האנשים שנכנסים לקניון מתפלג פואסונית באופן בלתי תלוי בכניסה האחרת. מספר האנשים שנכנסים בכניסה ה- $i$  מתפלג פואסונית עם קצב של  $i$  אנשים בשנייה. יהי  $Y$  מספר האנשים שנכנסים לקניון בשנייה מכל הכניסות יחדיו. מצאו את:  $E\left[\frac{1}{Y+1}\right]$ .
- 18** לרני 20 טושים אותם הוא מכניס באקראי ל-3 קלמרים. לכל טוש נבחר קלמר באקראי ובאופן בלתי תלוי בטוש אחר. כל קלמר יכול להכיל עד 20 טושים. נסמן ב- $X$  את מספר הקלמרים שיש בהם בדיוק 10 טושים. חשבו את  $E(\sqrt{x+7})$ .

19) בשדרות רוטשילד החליטו לשתול  $n$  ברושים ו-2 אורנים אחד אחרי השני בשורה. סידור העצים בשורה נעשה באקראי. נגדיר את  $X$  להיות מספר הברושים, בין הברוש הגבוה ביותר לברוש הנמוך ביותר שנשתלו.

א. מצאו את ההתפלגות של  $X$ .

ב. הוכיחו שהתוחלת של  $X$  היא  $\frac{n-2}{3}$ .

## תשובות סופיות:

- (1) א. תוחלת: 2, סטיית תקן: 1.      ב. תוחלת: 0, סטיית תקן: 2.
- (2) א. 0.03      ב. תוחלת: 0.15, שונות: 0.1875.
- (3) א. ראו טבלה:      ב. תוחלת: 3, שונות: 2.

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

- ג. תוחלת: 1.5, שונות: 0.5.
- (4) א. 0.9975      ב. 0.0172      ג. 1.0025
- (5) א. 0.375      ב. 0.6
- (6) א. ראו טבלה:      ב. תוחלת: 1.5, שונות: 0.61.

3	2	1	0	$x$
0.084	0.428	0.392	0.092	$P(x)$

- ג. תוחלת: 0, סטיית תקן: 4.68.
- (7) א. 0.4348      ב. 0.0923
- (8) א. 2.013      ב. 0.5256      ג. תוחלת: 8, שונות: 1.6.
- ד. תוחלת: 10.5, שונות: 3.475.
- (9) 10.
- (10) א. 0.6      ב. 0.0384      ג. 0.0256
- ד. תוחלת: 0.67, שונות: 1.11      ה. 0.24

$$P(X = k) = \binom{n}{k+n} \cdot p^{\frac{k+n}{2}} \cdot (1-p)^{\frac{n-k}{2}} \quad (11)$$

$$P(X = k) = \begin{cases} P & k = -1 \\ (1-P)^{k-1} \cdot P & k = 2, 3, \dots, \infty \end{cases} \quad (12)$$

$$0 < p < \sqrt{\frac{1}{2}} \quad \text{ג.}$$

$$. P(X = k) = \binom{m+k}{m} \cdot \left(\frac{1}{2}\right)^{k+m+1}, k = 0, 1, \dots, \infty \quad (13)$$

$$. n \cdot \left(\frac{2}{N}\right) \cdot \left(1 - \frac{2}{N}\right) \quad (14)$$

$$. \frac{2^n}{2n} \quad (15)$$

(16) ראו טבלה :

3	2	1	X
0.7099	0.2893	0.0008	P(X)

$$. \frac{e^{-6}}{6} [e^6 - 1] \quad (17)$$

$$. 2.675 \quad (18)$$

$$. P(X = k) = \frac{n-k-1}{\binom{n}{2}}, k = 0, 1, \dots, n-2. \quad (19)$$

א. ב. הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 40 - המשתנה המקרי הרציף - התפלגויות כלליות ללא אינטגרלים

תוכן העניינים

1. כללי.....169

## המשתנה המקרי הרציף – התפלגויות כלליות ללא אינטגרלים:

### רקע:

בפרק זה נעסוק בהתפלגות של משתנים מקריים רציפים (גובה אדם אקראי, זמן תגובה וכו').

משתנים רציפים הם משתנים שבתחום מסוים מקבלים רצף אינסופי של ערכים אפשריים בניגוד למשתנים בדידים.

נתאר את המשתנה המקרי הרציף על ידי פונקציה הנקראת פונקציית צפיפות. באופן כללי נסמן פונקציית צפיפות של משתנה רציף כלשהו ב- $f(x)$ .

השטח שמתחת לפונקציית הצפיפות נותן את ההסתברות.

פונקציית צפיפות חייבת להיות לא שלילית והשטח הכולל שמתחת לפונקציה יהיה תמיד 1.

בקורס הנוכחי לא נבצע אינטגרציה כדי לחשב את השטחים, אלא נשתמש בצורות הנדסיות מקובלות.

### ריענון מתמטי:

#### נוסחאות לחישוב שטחים:

שטח משולש: גובה ( $h$ ) כפול הבסיס ( $a$ ) חלקי 2:  $S_{triangle} = \frac{h \cdot a}{2}$ .

שטח מלבן: אורך ( $a$ ) כפול רוחב ( $b$ ):  $S_{rectangle} = a \cdot b$ .

#### משוואת קו ישר:

$$y = mx + n$$

$-m$  שיפוע.

$-n$  נקודת החיתוך עם ציר ה- $y$ .

שיפוע של ישר העובר דרך שתי נקודות:  $(X_1, Y_1), (X_2, Y_2)$ :  $m = \frac{Y_2 - Y_1}{X_2 - X_1}$ .

משוואת ישר שעובר דרך נקודה ספציפית  $(X_1, Y_1)$  ושיפועו ידוע  $m$ :

$$y - Y_1 = m(x - X_1)$$

### פונקציית התפלגות מצטברת:

היא פונקציה הנותנת במשתנה רציף את הסיכוי ליפול מתחת לערך מסוים:

$$F(t) = p(X \leq t)$$

כמו כן:

$$p(a < X < b) = F(b) - F(a) \quad p(X > t) = 1 - F(t)$$

### אחוזונים:

האחוזון ה- $P$  הוא ערך (נסמן אותו:  $x_p$ ) שהסיכוי ליפול מתחתיו הוא  $P$ .

$$p(X \leq x_p) = p$$

כלומר:

דוגמה (פתרון בהקלטה):

בשרטוט שלפניכם נתונה פונקציית הצפיפות של המשתנה  $X$ .

$X$  הינו זמן ההמתנה למענה קולי בדקות.

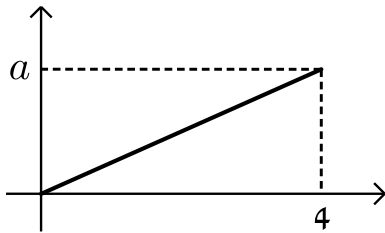
א. מצאו את ערכו של  $a$ .

ב. רשום את נוסחת פונקציית הצפיפות.

ג. חשבו את הסיכוי שזמן ההמתנה נמוך מ-2 דקות.

ד. בנו את פונקציית ההתפלגות המצטברת.

ה. מהו האחוזון ה-80 של ההתפלגות?



תשובות:

$$S_{\Delta} = \frac{a \cdot h}{2} = \frac{4 \cdot a}{2} = 2a$$

א.  $2a = 1 \rightarrow a = \frac{1}{2}$

$$f(x) = \begin{cases} \frac{1}{8} \cdot x & 0 \leq x \leq 4 \\ 0 & \text{זחרת} \end{cases}$$

ב.

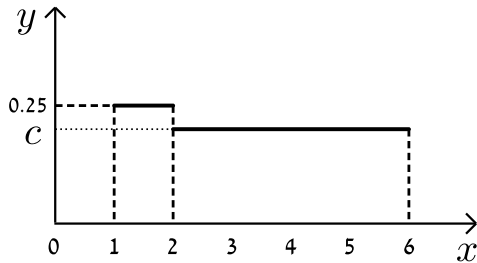
$$p(x < 2) = \frac{2 \cdot \frac{1}{4}}{2} = \frac{1}{4} = p(x \leq 2) \quad \text{ג.}$$

$$p(x \leq t) = f(t) = \begin{cases} 0 & t < 0 \\ \frac{t \cdot \frac{t}{8}}{2} = \frac{t^2}{16} & 0 \leq t \leq 4 \\ 1 & t > 4 \end{cases} \quad \text{ד.}$$

$$p(x < 2) = f(2) = \frac{2^2}{16} = \frac{1}{4} \quad \text{ה.}$$

## שאלות:

1)  $X$  הינו משתנה רציף עם פונקציית צפיפות כמוצג בשרטוט:



א. מצא את ערכו של  $c$ .

ב. בנה את פונקציית ההתפלגות המצטברת.

ג. חשבו את ההסתברויות הבאות:

i.  $P(x < 4)$

ii.  $P(x > 1.5)$

iii.  $P(1.5 < x < 5)$

iv.  $P(5 < x < 10)$

ד. מצא את החציון של המשתנה.

2) נתון משתנה מקרי רציף  $X$  שפונקציית הצפיפות שלו היא:

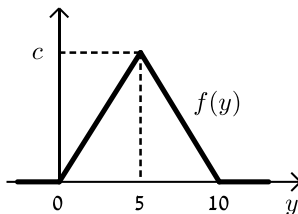
$$f(x) = \begin{cases} cx & 0 \leq x \leq b \\ 0 & \text{אחרת} \end{cases}$$

ידוע ש-  $P(0 < X < 1) = \frac{1}{4}$

א. מצאו במפורש את פונקציית הצפיפות של  $X$ .

ב. מצאו את החציון של  $X$ .

ג. מה הסיכוי ש- $X$  קטן מ-0.5?



3) נתונה פונקציית צפיפות של משתנה מקרי  $Y$ :

א. מצאו את  $c$ .

ב. מצאו את פונקציית ההתפלגות המצטברת של  $Y$ .

ג. חשבו את ההסתברויות:

i.  $P(Y = 7.0)$

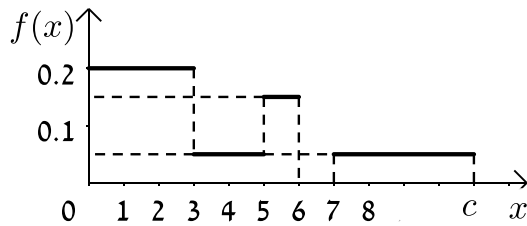
ii.  $P(Y \leq 3.0)$

iii.  $P(7.5 \leq Y \leq 15.5)$

iv.  $P(Y > 4)$

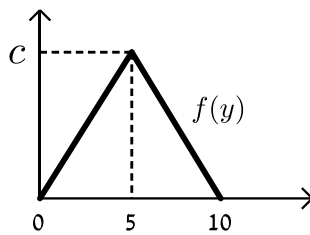
ד. מצאו את העשירון התחתון:  $y_{0.1}$ , הרבעון התחתון:  $y_{0.25}$  והחציון של  $Y$ .

הסיקו מהו העשירון עליון:  $y_{0.9}$ .



(4) נתונה פונקציית צפיפות של משתנה מקרי  $X$ :

- א. מצאו ערך  $c$  שעבורו תתקבל פונקציית צפיפות.  
 ב. מצאו את פונקציית ההתפלגות המצטברת.  
 ג. חשבו את ההסתברויות הבאות:  
 $P(1.0 < X \leq 5.0)$ ,  $P(X \geq -2.0)$ ,  $P(X \geq 4)$



(5) נתונה פונקציית הצפיפות הבאה:

- א. מה ערכו של  $c$ ?  
 ב. מצא אינטרוול (תחום) סימטרי סביב הערך 5 שהסיכוי ליפול בו הינו  $\frac{1}{2}$ .

(6) זמן ההמתנה בדקות של לקוח בתור למכולת השכונתית מתפלג עם פונקציית

$$F(t) = 1 - e^{-0.2t}$$

- א. מה הסיכוי שזמן ההמתנה יהיה לפחות רבע שעה?  
 ב. אם חיכיתי בתור 10 דקות מה ההסתברות שאאלץ לחכות בסך הכל פחות מרבע שעה?  
 ג. מה הזמן ש-90% מהלקוחות מחכים מתחתיו?

## תשובות סופיות:

$$f(t) = \begin{cases} 0 & t > 1 \\ (t-1) \cdot 0.25 & 1 \leq t \leq 2 \\ 0.25 + (t-2) \cdot \frac{3}{16} & 2 < t \leq 6 \\ 1 & t > 6 \end{cases} \quad \text{ב.} \quad \text{א.} \quad \frac{3}{16} \quad (1)$$

$$f(x) = \begin{cases} \frac{1}{2} \cdot x & 0 \leq x \leq 2 \\ 0 & \text{חרת} \end{cases} \quad \text{א.} \quad (2)$$

.1.41 ב.    .3  $\frac{1}{3}$  ג.    .  $\frac{3}{16}$  iv    .  $\frac{11}{16}$  .iii    .  $\frac{7}{8}$  .ii

$$f(t) = p(y \leq t) = \begin{cases} 0 & t < 0 \\ \frac{t \cdot 0.04t}{2} = 0.02t^2 & 0 \leq t \leq 5 \\ 1 - \frac{(10-t)(-0.04(t-10))}{2} = 1 - 0.02(t-10)^2 & 5 < t \leq 10 \\ 1 & t > 10 \end{cases} \quad \text{ב.} \quad \text{א.} \quad 0.2 \quad (3)$$

.0.32 iv    .0.125 .iii    .0.18 .ii    .0 .i .ג  
 .  $Y_{0.1} = 2.24$  ,  $Y_{0.25} = 3.54$  ,  $Y_{0.9} = 7.76$  ,  $Y_{0.5} = 5$  .ד

$$f(t) = \begin{cases} 0 & t < 0 \\ 0.2t & 0 < t \leq 3 \\ 0.6 + (t-3) \cdot 0.05 & 3 < t \leq 5 \\ 0.7 + (t-5) \cdot 0.15 & 5 < t \leq 6 \\ 0.85 & 6 < t \leq 7 \\ 0.85 + (t-7) \cdot 0.05 & 7 < t \leq 10 \\ 1 & t > 10 \end{cases} \quad \text{ב.} \quad \text{א.} \quad 0.10 \quad (4)$$

.  $P(x \geq 4) = 0.35$  ,  $P(x \geq -2) = 1$  ,  $P(1 < x < 5) = 0.5$  .ג

.5  $\pm$  1.46 ב.    .0.2 א. (5)

.115.13 ג.    .0.6321 ב.    .0.0498 א. (6)

# סטטיסטיקה והסתברות 01 88165

פרק 41 - התפלגויות רציפות מיוחדות- התפלגות מעריכית

תוכן העניינים

1. כללי ..... 175

## התפלגויות רציפות מיוחדות – התפלגות מעריכית:

### רקע:

התפלגות זו היא התפלגות רציפה המאפיינת את הזמן עד להתרחשות מאורע מסוים.  $\lambda$  - הוא ממוצע מספר האירועים המתרחשים ביחידת זמן (אותו פרמטר מההתפלגות הפואסונית):  $X \sim \exp(\lambda)$  כאשר  $\lambda > 0$ .

התפלגות זו צריכה להיות נתונה בתרגיל או שיאמר שמספר האירועים ביחידת זמן מתפלג פואסונית ואז הזמן עד להתרחשות המאורע הבא מתפלג מעריכית.

### פונקציית הצפיפות של ההתפלגות:

$$f(x) = \lambda e^{-\lambda x} \quad \text{לכל } x \geq 0.$$

פונקציית ההתפלגות המצטברת:  $F(t) = p(x \leq t) = 1 - e^{-\lambda t}$ .

$$\text{התוחלת: } E(x) = \frac{1}{\lambda}$$

$$\text{השונות: } V(x) = \frac{1}{\lambda^2}$$

להתפלגות זו יש תכונת חוסר הזיכרון:  $P(X > a+b \mid X > a) = P(X > b)$ .

דוגמה (פתרון בהקלטה):

אורך חיי סוללה מתפלג מעריכית עם תוחלת של 8 שעות.

א. מה ההסתברות שסוללה תחזיק מעמד פחות מ-9 שעות?

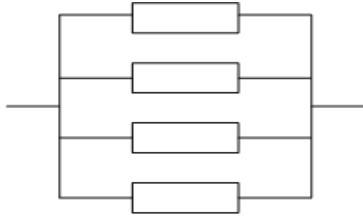
ב. מה סטיית התקן של אורך חיי הסוללה?

ג. אם סוללה כבר חייה מעל שעתיים, מה הסיכוי שהיא תחייה מעל 7 שעות בסך הכול?

## שאלות:

- (1) הזמן שלוקח במערכת עד שתקלה מתרחשת מתפלג מעריכית עם תוחלת של 0.5 שעה.
- מה הסתברות שהתקלה הבאה תתרחש תוך יותר מ-0.5 שעה?
  - מה ההסתברות שהתקלה הבאה תתרחש תוך פחות משעה?
  - מצא את הזמן החציוני להתרחשות תקלה במערכת.
- (2) הזמן שעובר בכביש מסוים עד להתרחשות תאונה מתפלג מעריכית עם תוחלת של 24 שעות.
- מהי סטית התקן של הזמן עד להתרחשות תאונה?
  - מה ההסתברות שהתאונה הבאה תתרחש תוך פחות מיממה?
  - מהי ההסתברות שהתאונה הבאה תתרחש תוך לפחות יומיים?
- (3) משך הזמן X (בדקות) שסטודנטים עובדים רצוף על מחשב מתפלג מעריכית עם תוחלת של 30 דקות.
- מה הסיכוי שעבודת סטודנט על המחשב תארך פחות מרבע שעה?
  - מה הסיכוי שעבודת סטודנט על המחשב תארך בין רבע שעה לחצי שעה?
  - אם סטודנט עובד על המחשב כבר יותר מ-10 דקות, מה ההסתברות שמשך כל עבודתו יעלה על 30 דקות?
  - מהו הזמן שבסיכוי של 90% הסטודנט יעבוד פחות ממנו?
- (4) בממוצע מגיעים לחדר מיון 4 חולים בשעה בזרם פואסוני.
- שולה המזכירה הגיעה לחדר המיון. מה ההסתברות שזמן ההמתנה שלה לחולה הבא יהיה יותר מ-20 דקות?
  - אם שולה המתינה יותר מרבע שעה לחולה הבא. מה ההסתברות שתמתין בסך הכל יותר מחצי שעה?
  - מה ההסתברות שבין החולה הראשון לשני יש להמתין יותר מרבע שעה ובין החולה שני לשלישי יש להמתין פחות מרבע שעה?

5) מערכת חשמלית כוללת 4 רכיבים אלקטרוניים זהים הפועלים במקביל כמתואר בשרטוט:



על מנת שהמערכת תפעל בצורה תקינה נדרש שלפחות אחד מהמרכיבים יהיה תקין. אורך החיים של כל רכיב מתפלג מעריכית עם ממוצע של 100 שעות.

א. מה ההסתברות שהמערכת תפעל בצורה תקינה במשך 100 שעות לפחות?

ב. מעוניינים להוסיף במקביל עוד רכיב למערכת.

עלות הוספת רכיב היא  $K$  ₪.

כמו כן אם המערכת עבדה פחות מ-100 שעות

נגרם הפסד של  $A$  ₪.

מה התנאי שבו יהיה כדאי להוסיף את הרכיב למערכת?

### תשובות סופיות:

1) א. 0.368 ב. 0.865 ג. 0.347

2) א. 24 שעות. ב. 0.632 ג. 0.135

3) א. 0.393 ב. 0.239 ג. 0.513 ד. 69.08

4) א. 0.264 ב. 0.368 ג. 0.233

5) א. 0.8403 ב.  $K < 0.0588A$

# סטטיסטיקה והסתברות 01 88165

פרק 42 - התפלגויות רציפות מיוחדות-התפלגות אחידה

תוכן העניינים

1. כללי ..... 178

## התפלגויות רציפות מיוחדות – התפלגות אחידה:

### רקע:

זו התפלגות שפונקציית הצפיפות שלה קבועה בין  $a$  לבין  $b$ .

$$. X \sim U(a, b)$$

### פונקציית הצפיפות:

$$. f(x) = \frac{1}{b-a}$$

$$a \leq x \leq b$$

### פונקציית ההתפלגות המצטברת:

$$. F(t) = \frac{t-a}{b-a}$$

### התוחלת :

$$. E(X) = \frac{a+b}{2}$$

### השונות:

$$. V(x) = \frac{(b-a)^2}{12}$$

דוגמה (פתרון בהקלטה):

$X$  - משתנה מקרי רציף המתפלג באופן אחיד בין 20 ל-40.

מה הסיכוי ש- $X$  קטן מ-25?

מה התוחלת והשונות של  $X$ ?

$$a = 20, b = 40$$

$$X \sim U(20, 40)$$

$$\text{א. } P(x < 25) = f(25) = \frac{25-20}{40-20} = 0.25$$

$$E(x) = \frac{20+40}{2} = 30$$

$$\text{ב. } V(x) = \frac{(40-20)^2}{12} = 33\frac{1}{3}$$

## שאלות:

- (1) משך (בדקות) הפסקה בשיעור,  $X$ , מתפלג:  $U(13,16)$ .
- מהי התוחלת ומהי סטיית התקן של משך ההפסקה?
  - מהי ההסתברות שהפסקה תמשך יותר מ-15 דקות?
  - מהי ההסתברות שמשך ההפסקה יסטה מהתוחלת בפחות מדקה?
- (2) רכבת מגיעה לתחנה בשעות היום כל עשר דקות. אדם הגיע לתחנה בזמן אקראי.
- הסבר כיצד מתפלג זמן ההמתנה לרכבת?
  - אם זמן ההמתנה לרכבת ארך יותר מ-5 דקות, מהי ההסתברות שבסך הכל האדם ימתין לרכבת פחות מ-8 דקות?
  - מה תוחלת מספר הימים שיעברו עד הפעם הראשונה שהאדם ימתין לרכבת יותר מ-9 דקות?
- (3) מכונה אוטומטית ממלאת גביעי גלידה. משקל הגלידה לגביע מתפלג אחיד בין 100-110 גרם (המשקל הוא של גלידה ללא הגביע).
- מה ההסתברות שמשקל הגלידה בגביע יהיה מעל 108 גרם?
  - נתון שהגלידה בגביע עם משקל נמוך מ-107 גרם. מה ההסתברות שמשקל הגלידה יהיה מעל 105 גרם?
  - מה העשירון העליון של משקל הגלידה בגביע?
  - עלות גביע גלידה היא 0.5 שקל. כל גרם של גלידה עולה 0.22 אגורות. מהי התוחלת ומהי סטיית התקן של עלות הגביע ביחד עם הגלידה?

## תשובות סופיות:

- (1) א. תוחלת: 14.5, שונות: 0.866. ב.  $\frac{1}{3}$ . ג.  $\frac{2}{3}$ .
- (2) א.  $X \sim U(0,10)$ . ב. 0.6. ג. 10.
- (3) א. 0.2. ב.  $\frac{2}{7}$ . ג. 109. ד. תוחלת: 73.1 אגורות, סטיית תקן: 0.635 אגורות.

# סטטיסטיקה והסתברות 01 88165

פרק 43 - התפלגויות רציפות מיוחדות - התפלגות נורמלית

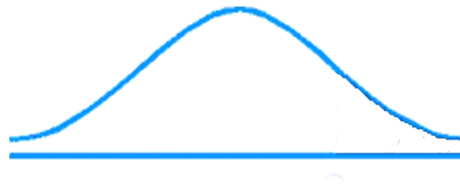
תוכן העניינים

1. כללי ..... 181

## התפלגויות רציפות מיוחדות – התפלגות נורמלית:

### רקע:

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



לעקומה זו קוראים גם עקומת גאוס ועקומה אחת נבדלת מהשנייה באמצעות הממוצע וסטיית התקן שלה.

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

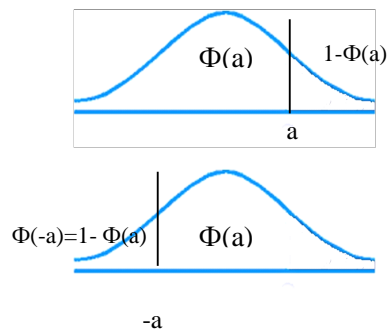
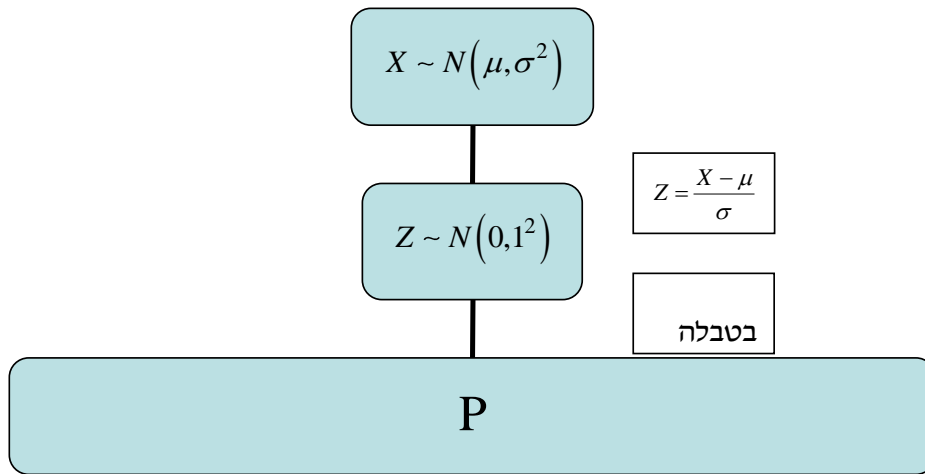
$$\text{נוסחת פונקציית הצפיפות: } f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

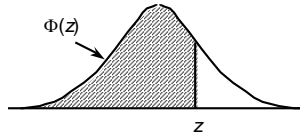
כדי לחשב הסתברויות בהתפלגות נורמלית יש לחשב את השטחים הרלוונטיים שמתחת לעקומה. כדי לחשב שטחים אלה נמיר כל התפלגות נורמלית להתפלגות נורמלית סטנדרטית על ידי תהליך הנקרא תקנון. התפלגות נורמלית סטנדרטית היא התפלגות נורמלית שהממוצע שלה הוא אפס וסטיית התקן היא אחת, והיא תסומן באות  $Z$ :  $Z \sim N(0, 1^2)$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

$$\text{תהליך התקנון מבוצע על ידי הנוסחה הבאה: } Z = \frac{X - \mu}{\sigma}$$

אחרי תקנון מקבלים ערך הנקרא ציון תקן. ציון התקן משמעו בכמה סטיות תקן הערך סוטה מהממוצע.

לאחר חישוב ציון התקן של ערך מסוים נעזרים בטבלה של ההתפלגות הנורמלית הסטנדרטית לחישוב השטח הרצוי, ובאופן כללי נתאר את הסכמה הבאה:



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


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

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

דוגמה (הפתרון בהקלטה):

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

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

## שאלות:

- (1) הגובה של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 170 ס"מ וסטית 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 בדיוק 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 לתרופה מסוימת להשפיע מתפלג נורמלית עם ממוצע של 30 דקות ושונות של 9 דקות רבועות.
- מהי פרופורציית המקרים בהן התרופה תעזור אחרי יותר משעה?
  - מה אחוז מהמקרים שבהן התרופה תעזור בין 35 ל-37 דקות?
  - מה הסיכוי שהתרופה תעזור בדיוק תוך 36 דקות?
  - מה שיעור המקרים שבהן ההשפעה של התרופה תסטה מ-30 דקות בפחות מ-3 דקות?
- (3) המשקל של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 60 ק"ג וסטית תקן של 8 ק"ג.
- מה אחוז האנשים שמשקלם נמוך מ-55 ק"ג?
  - מהי פרופורציית 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 בלא יותר מ-4 ק"ג?
  - מה הסיכוי שאדם אקראי ישקול מתחת ל-140 ק"ג?
- (4) משקל תינוקות ביום היוולדם מתפלג נורמלית עם ממוצע של 3300 גרם וסטית תקן 400 גרם.
- מצאו את העשירון העליון.
  - מצאו את האחוזון ה-95.
  - מצאו את העשירון התחתון.

- (5) ציוני מבחן אינטליגנציה מתפלגים נורמלית עם ממוצע 100 ושונויות 225.
- מה העשירון העליון של הציונים במבחן האינטליגנציה?
  - מה העשירון התחתון של ההתפלגות?
  - מהו הציון ש-20% מהנבחנים מקבלים מעליו?
  - מהו האחוזון ה-20?
  - מהו הציון ש-5% מהנבחנים מקבלים מתחתיו?
- (6) נפח משקה בבקבוק מתפלג נורמלית עם סטיית תקן של 20 מ"ל, ונתון ש-33% מהבקבוקים בעלי נפח שעולה על 508.8 מ"ל.
- מה ממוצע נפח משקה בבקבוק?
  - 5% מהבקבוקים המיוצרים עם הנפח הגבוה ביותר נשלחים לבדיקה, החל מאיזה נפח שולחים בקבוק לבדיקה?
  - 1% מהבקבוקים עם הנפח הקטן ביותר נתרמים לצדקה, מהו הנפח המקסימלי לצדקה?
- (7) אורך חיים של מכשיר מתפלג נורמלית. ידוע שמחצית מהמכשירים חיים פחות מ-500 שעות, כמו כן ידוע ש-67% מהמכשירים חיים פחות מ-544 שעות.
- מהו ממוצע אורך חיי מכשיר?
  - מהי סטית בתקן של אורך חיי מכשיר?
  - מה הסיכוי שמכשיר אקראי יחיה פחות מ-460 שעות?
  - מהו המאיון העליון של אורח חיי מכשיר?
  - 1% מהמכשירים בעלי אורך החיים הקצר ביותר נשלח למעבדה לבדיקה מעמיקה. מהו אורך החיים המקסימלי לשליחת מכשיר למעבדה?
- (8) להלן שלוש התפלגויות נורמליות של שלוש קבוצות שונות ששורטטו באותה מערכת צירים. ההתפלגויות מוספרו כדי להבדיל ביניהן.
- לאיזו התפלגות הממוצע הגבוה ביותר?
  - במה מבין המדדים הבאים התפלגות 1 ו-2 זהות?
    - בעשירון העליון.
    - בממוצע.
    - בשונויות.
  - לאיזו התפלגות סטיית התקן הקטנה ביותר?
    - 1.
    - 2.
    - 3.
    - אין לדעת.



- 9) הזמן שלוקח לאדם להגיע לעבודתו מתפלג נורמלית עם ממוצע של 40 דקות וסטית תקן של 5 דקות.
- א. מה ההסתברות שמשך הנסיעה של האדם לעבודתו יהיה לפחות שלושת רבעי השעה?
- ב. אדם יצא לעבודתו בשעה 08:10 מביתו. הוא צריך להגיע לעבודתו בשעה 09:00. מה הסיכוי שיאחר לעבודתו?
- ג. אם ידוע שזמן נסיעתו לעבודה היה יותר משלושת רבעי השעה. מה ההסתברות שזמן הנסיעה הכולל יהיה פחות מ-50 דקות?
- ד. מה הסיכוי שבשבוע (חמישה ימי עבודה) בדיוק פעם אחת יהיה זמן הנסיעה לפחות שלושת רבעי השעה?
- 10) ההוצאה החודשית לבית אב בעיר "טרירה" מתפלגת נורמלית עם ממוצע של 2000 דולר וסטית תקן של 300 דולר. בחרו באקראי 5 בתי אב. ההסתברות שלפחות אחד מהם מוציא בחודש מעל ל- $T$  דולר היא 0.98976.
- א. מה ערכו של  $T$ ?
- ב. מה הסיכוי שההוצאה החודשית של בית אב בעיר תהיה לפחות סטיית תקן אחת מעל  $T$ ?
- ג. מסתבר שנפלה טעות בנתונים, ויש להוסיף 100 דולר להוצאות החודשית של כל בתי האב בעיר. לאור זאת, מה ההסתברות שההוצאה החודשית של בית אב נמוכה מ-1800 דולר?
- 11) אורך שיר אקראי המשודר ברדיו מתפלג נורמלית עם תוחלת של 3.5 דקות וסטית תקן של שלושים שניות.
- א. מה ההסתברות שאורך של שיר אקראי המנוגן ברדיו יהיה בין 3 ל-2.5 דקות?
- ב. מהו הטווח הבין רבעוני של אורך שיר המשודר ברדיו?
- ג. ביום מסוים מנוגנים 200 שירים ברדיו. כמה שירים מתוכם תצפה שיהיו באורך הנמוך מ-3.5 דקות?
- ד. בשעה מסוימת שודרו 8 שירים. מה ההסתברות שרבע מהם בדיוק היו ארוכים מ-4 דקות והיתר לא?

**תשובות סופיות:**

ה. 50%	ד. 50%	ג. 0	ב. 2.28%	א. 89.25%	<b>(1)</b>
	ד. 68.26%	ג. 0%	ב. 3.76%	א. 0%	<b>(2)</b>
	ד. 0.383	ג. 39.44%	ב. 89.44%	א. 26.43%	<b>(3)</b>
				ה. 100%	
		ג. 2787.2	ב. 3958	א. 3812.8	<b>(4)</b>
	ד. 87.4	ג. 112.6	ב. 80.8	א. 119.2	<b>(5)</b>
		ג. 453.48	ב. 532.9	א. 500	<b>(6)</b>
	ד. 733	ג. 0.3446	ב. 100	א. 500	<b>(7)</b>
				ה. 267	
		ג. 1	ב. בממוצע.	א. 3	<b>(8)</b>
	ד. 0.3975	ג. 0.8563	ב. 0.0228	א. 0.1587	<b>(9)</b>
		ג. 0.1587	ב. 0.2266	א. 1925	<b>(10)</b>
	ד. 0.25	ג. 100	ב. 0.675	א. 0.1359	<b>(11)</b>

# סטטיסטיקה והסתברות 01 88165

פרק 44 - טרנספורמציה על משתנה מקרי רציף

תוכן העניינים

1. כללי ..... 189

## טרנספורמציה על משתנה מקרי רציף:

### רקע:

מצב שבו ידועה לנו התפלגות של משתנה מקרי רציף כלשהו ואז יוצרים משתנה מקרי חדש שהוא פונקציה של המשתנה המקרי הידוע.

דוגמה (פתרון בהקלטה):

נתון משתנה מקרי רציף  $X$  המתפלג אחיד בין 0 ל-1 .  
מצאו את פונקציית ההתפלגות המצטברת של המשתנה  $Y$ ,  
כאשר הקשר בין  $X$  ל- $Y$  נתון על ידי הנוסחה:  $Y = e^x$ .

## שאלות:

- (1) יהי  $W$  משתנה מקרי המתפלג מעריכית עם תוחלת השווה ל-1. הגדירו משתנה חדש:  $Y = e^{-W}$ .  
 א. מצאו את פונקציית ההתפלגות המצטברת של  $Y$ .  
 ב. זהו את  $Y$  כהתפלגות מיוחדת וקבע מהם הפרמטרים.
- (2) נתון:  $X \sim U(0,1)$ , ויוצרים דרך  $X$  משתנה חדש המוגדר להיות:  $R = X^2$ . מצאו את פונקציית הצפיפות של המשתנה החדש  $R$ .
- (3) ידוע ש- $X \sim \exp(\lambda)$ . כמו כן, נתון הקשר הבא:  $Y = \ln(X)$ . הוכיחו שפונקציית הצפיפות של  $Y$  נתונה על ידי הנוסחה הבאה:  $f(Y) = \lambda \cdot e^{-\lambda \cdot e^Y + 1}$ .
- (4) ידוע ש- $X \sim \exp(\lambda=1)$ . כמו כן, נתון הקשר הבא:  $Y = 1 - 2 \cdot e^{-X}$ .  
 א. מצאו את פונקציית ההתפלגות המצטברת של  $Y$ .  
 ב. זהו את ההתפלגות של  $Y$ .
- (5) אורך מקצוע של קובייה מתפלג אחיד בין 1 ל-2. מצאו את פונקציית הצפיפות של נפח הקובייה.
- (6) נתונה פונקציית ההתפלגות המצטברת הבאה:  $F_X(t) = \theta^t - 1$ , עבור התחום:  $0 \leq t \leq 1$ .  
 א. מצאו את ערכו של הפרמטר  $\theta$ .  
 ב. מצאו את פונקציית הצפיפות של המשתנה  $X$ .  
 ג. יהי  $Y = 2^X - 1$ . מצאו את פונקציית הצפיפות של  $Y$  וזהו את התפלגותו.

### תשובות סופיות:

(1) ב.  $Y \sim U(0,1)$ .

(2)  $f(R) = \frac{1}{2\sqrt{R}}$  כאשר  $1 > R > 0$ .

(3) שאלת הוכחה.

(4) ב.  $Y \sim U(-1,1)$ .

(5)  $f(y) = \frac{1}{3}y^{-\frac{2}{3}}$  כאשר  $1 < y < 8$ .

(6) א. 2. ב.  $Y \sim U(0,1)$ .

# סטטיסטיקה והסתברות 01 88165

פרק 45 - פונקציה יוצרת מומנטים

תוכן העניינים

1. כללי ..... 192

## פונקציה יוצרת מומנטים:

### רקע:

פונקציה יוצרת מומנטים של משתנה מקרי כלשהו מוגדרת להיות:  $M_X(t) = E(e^{tx})$ .  
 אם מדובר במשתנה מקרי בדיד, הפונקציה יוצרת המומנטים תהיה:

$$M_X(t) = E(e^{tX}) = \sum_k e^{tk} \cdot P(X = k)$$

אם מדובר במשתנה מקרי רציף, פונקציית יוצרת המומנטים תהיה:

$$M_X(t) = E(e^{tx}) = \int_x e^{tx} \cdot f(x) dx$$

המומנט מסדר  $n$  מוגדר להיות:  $E(X^n)$ .

מומנט מסדר  $n$  של משתנה מקרי  $X$  מתקבל מהנגזרת ה- $n$ ית לפי  $t$  של פונקציית

יוצרת המומנטים  $M_X(t)$  בנקודה שבה  $t = 0$ . כלומר:  $M_X^{(n)}(t)|_{t=0} = E(X^n)$ .

### משפט:

קיימת התאמה חד-חד-ערכית בין משתנה מקרי לבין פונקציית יוצרת המומנטים שלו.

### תזכורת מתמטית לנגזרות:

$$(x^n)' = nx^{n-1}$$

$$(k \cdot f(x))' = k \cdot f'(x)$$

$$(a^x)' = a^x \cdot \ln a$$

$$(e^x)' = e^x$$

$$(\log_a x)' = \frac{1}{x \ln a}$$

$$(\ln x)' = \frac{1}{x}$$

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$$

$$(k)' = 0$$

$$(f(x) \pm g(x))' = f'(x) \pm g'(x)$$

$$(f(x) \cdot g(x))' = f'(x)g(x) + g'(x)f(x)$$

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - g'(x)f(x)}{(g(x))^2}$$

$$[f(g(x))]' = f'(g(x)) \cdot g'(x) \text{ - כלל שרשרת}$$

דוגמה (פתרון בהקלטה):

הראו שהפונקציה יוצרת המומנטים של ההתפלגות המעריכית:  $X \sim \exp(\lambda)$ ,

היא:  $\frac{\lambda}{\lambda - t}$ . מצאו את המומנט הראשון והמומנט השני של ההתפלגות.

## שאלות:

- (1) נתונה פונקציה ההסתברות הבאה למשתנה מקרי בדיד.  
 א. מצאו את פונקציית יוצרת המומנטים.  
 ב. מצאו את התוחלת על סמך סעיף א'.

$X$	1	2	3
$P(X)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

- (2) מצאו את פונקציית יוצרת המומנטים של התפלגות הבינומית:  $X \sim B(n, p)$ , ומצאו את המומנט הראשון והשני של הפונקציה.
- (3) מצאו את פונקציית יוצרת המומנטים של ההתפלגות הגיאומטרית:  $X \sim G(P)$ , וחשבו את תוחלת של ההתפלגות מתוך פונקציית יוצרת המומנטים.
- (4) מצאו את פונקציית יוצרת מומנטים של התפלגות הפואסונית:  $x \sim p(\lambda)$ . מצאו את המומנט הראשון והשני של ההתפלגות.

- (5) יהי  $X$  משתנה מקרי בעל פונקציית הצפיפות הבאה:

$$f(x) = \begin{cases} Ae^{-x} & 0 \leq x \leq 7 \\ 0 & \text{אחר } t \end{cases}$$

- א. מצאו את ערכו של  $A$ .  
 ב. מצאו את הפונקציה יוצרת המומנטים של  $X$ .

- (6) יהי  $X$  משתנה מקרי עם תוחלת 5 ושונות 16, ותהי  $m_x(t)$  פונקציית יוצרת המומנטים של  $X$ .  $Y$  הינו משתנה מקרי עם פונקציית יוצרת מומנטים  $m_y(t)$ , ונתון:  $m_y(t) = t \cdot m_x(t)$ .  
 חשבו את התוחלת והשונות של  $Y$ .

## תשובות סופיות:

- (1) א. פונקציה יוצרת מומנטים:  $\frac{1}{2}e^t + \frac{1}{3}e^{2t} + \frac{1}{6}e^{3t}$  . ב.  $\frac{2}{3}$  .
- (2) פונקציה יוצרת מומנטים:  $(e^t \cdot p + 1 - p)^n$  .
- (3) פונקציה יוצרת מומנטים:  $\frac{e^t p}{1 - e^t \cdot (1 - p)}$  .
- (4) פונקציה יוצרת מומנטים:  $e^{\lambda(e^t - 1)}$  .
- (5) א.  $\frac{1}{1 - e^{-7}}$  . ב. פונקציה יוצרת מומנטים:  $\frac{e^7}{e^7 - 1} \cdot \frac{e^{7(t-1)} - 1}{t - 1}$  .
- (6) תוחלת: 1, שונות: 9.

## נספחים:

פונקציית התפלגות מצטברת $f(X)t$	פונקציית צפיפות $f(X)t$	התפלגות
$f_x(t) = \begin{cases} 0 & t < 0 \\ \frac{t-a}{b-a} & a \leq t \leq b \\ 1 & t > b \end{cases}$	$f_x(t) = \begin{cases} \frac{1}{(b-a)} & a \leq t \leq b \\ 0 & \text{else} \end{cases}$	אחיד $U(a,b)$
$f_x(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & \text{else} \end{cases}$	$f_x(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{else} \end{cases}$	מעריכי $\exp(\lambda)$
$\phi\left(\frac{t-\mu}{\sigma}\right)$	$f_x(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	נורמלית $N(\mu, \sigma^2)$

התפלגות	$E(X)$	$VAR(X)$	$M_X(t)$
אחיד $U(a,b)$	$\frac{b-a}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$
מעריכי $\exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}$
נורמלית $N(\mu, \sigma^2)$	$\mu$	$\sigma^2$	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$

$M_X(t)$	$Var(x)$	$E(x)$	$P_X(x)$	משמעות	משתנה מקרי
$[pe^t + q]^n$	$n \cdot p \cdot q$	$n \cdot p$	$\binom{n}{x} p^x q^{n-x}$ $x = 0, 1, \dots, n$	חוזרים באופן בלתי תלוי על אותו ניסוי ברנולי $n$ פעמים: $P$ ההסתברות להצלחה. $1 - P = q$ ההסתברות לכישלון $X$ - מספר ההצלחות	בינומי $Bin(n, p)$
$\frac{pe^t}{1 - qe^t}$	$\frac{q}{p^2}$	$\frac{1}{p}$	$pq^{x-1}$ $x = 1, 2, \dots, \infty$	חוזרים באופן בלתי תלוי על אותו ניסוי ברנולי עד ההצלחה הראשונה. $X$ - מספר ניסויים עד הצלחה ראשונה.	גיאומטרי $G(P)$
$e^{\lambda(e^t - 1)}$	$\lambda$	$\lambda$	$e^{-\lambda} \frac{\lambda^x}{x!}$	$X$ - מספר ההופעות בלידת זמן. מ"מ המקבל ערכים $.0, 1, \dots, \infty$ .	פואסוני $Pois(\lambda)$

# סטטיסטיקה והסתברות 01 88165

פרק 46 - תכונות של פונקציית יוצרת מומנטים

תוכן העניינים

1. כללי ..... 198

## תכונות של פונקציה יוצרת מומנטים:

### רקע:

להלן מספר תכונות שפונקציית יוצרת מומנטים מקיימת:

- קיימת התאמה חד-חד-ערכית בין משתנה מקרי לבין פונקציית יוצרת המומנטים שלו.
- השפעת טרנספורמציה לינארית על פונקציית יוצרת מומנטים:

$$M_{aX+b}(t) = e^{bt} M_X(at)$$

- אם  $X$  ו- $Y$  משתנים בלתי תלויים מתקיים ש:

$$M_{X+Y}(t) = E(e^{t \cdot X}) \cdot E(e^{t \cdot Y}) = M_X(t) \cdot M_Y(t)$$

## תזכורת:

$F_x(t)$ פונקציית התפלגות מצטברת	$f_x(t)$ פונקציית צפיפות	התפלגות
$f_x(t) = \begin{cases} 0 & t < 0 \\ \frac{t-a}{b-a} & a \leq t \leq b \\ 1 & t < b \end{cases}$	$f_x(t) = \begin{cases} \frac{1}{(b-a)} & a \leq t \leq b \\ 0 & \text{else} \end{cases}$	אחיד $U(a,b)$
$f_x(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & \text{else} \end{cases}$	$f_x(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{else} \end{cases}$	מעריכי $\exp(\lambda)$
$\phi\left(\frac{t-\mu}{\sigma}\right)$	$f_x(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$	נורמלית $N(\mu, \sigma^2)$

התפלגות	$E(X)$	$VAR(X)$	$M_X(t)$
אחיד $U(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$
מעריכי $\exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}$
נורמלית $N(\mu, \sigma^2)$	$\mu$	$\sigma^2$	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$

$M_X(t)$	$Var(x)$	$E(x)$	$P_X(x)$	משמעות	משתנה מקרי
$[pe^t + q]^n$	$n \cdot p \cdot q$	$n \cdot p$	$\binom{n}{x} p^x q^{n-x}$ $x = 0, 1, \dots, n$	חוזרים באופן בלתי תלוי על אותו ניסוי ברנולי $n$ פעמים: $P$ ההסתברות להצלחה $1 - P = q$ ההסתברות לכישלון $x$ : מספר הצלחות	בינומי $Bin(n, p)$
$\frac{pe^t}{1 - qe^t}$	$\frac{q}{p^2}$	$\frac{1}{p}$	$pq^{x-1}$ $x = 1, 2, \dots, \infty$	חוזרים באופן בלתי תלוי על אותו ניסוי ברנולי עד הצלחה הראשונה. $x$ : מספר ניסויים עד הצלחה ראשונה	גיאומטרי $G(p)$
$e^{\lambda(e^t - 1)}$	$\lambda$	$\lambda$	$e^{-\lambda} \frac{\lambda^x}{x!}$ $0, 1, \dots, \infty$	$x$ : מספר ההופעות בלידת זמן. מ"מ המקבל ערכים $0, 1, \dots, \infty$	פואסוני $Pois(\lambda)$

דוגמה (פתרון בהקלטה):

נתון:  $Y \sim P(\lambda = 2)$   $X \sim P(\lambda = 4)$

$X$  ו- $Y$  הינם בלתי תלויים.

א. מהי פונקציית יוצרת המומנטים של:  $5X - 3$ ?

ב. נגדיר את:  $T = X + Y$ . מה ההתפלגות של  $T$ ?

**שאלות:**

(1) נתון ש-  $X_i \sim p(\lambda)$  בלתי תלויים.

א. מצאו את פונקציית יוצרת מומנטים של  $\sum_{i=1}^n X_i$ .

ב. הוכיחו ש-  $\sum_{i=1}^n X_i \sim P(n \cdot \lambda)$ .

(2) נתון:  $Y \sim P(\lambda = 2)$ ,  $X \sim P(\lambda = 10)$ .

$X$  ו- $Y$  הינם בלתי תלויים. נגדיר את:  $T = X + Y$ .

א. מצאו את פונקציית יוצרת המומנטים של  $T$ .

ב. הוכיחו ש-  $T \sim P(\lambda = 12)$ .

ג. הוכיחו ש-  $X/T = 8 \sim B\left(8, \frac{5}{6}\right)$ . כלומר, ההתפלגות של  $X$ ,

בהינתן ש- $T = 8$  היא בינומית עם הפרמטרים:  $n = 8$  ו- $p = \frac{5}{6}$ .

(3) יהי:  $X_i \sim \exp(1)$ ,  $i = 1, 2, \dots, n$  והמשתנים הם בלתי תלויים.

$$T = \sum_{i=1}^n X_i$$

נגדיר את

א. מצאו את פונקציית יוצרת המומנטים של  $T$ .

ב. חשבו את התוחלת והשונות של  $T$ .

ג. יהי:  $Z = \frac{T - E(T)}{\sigma(T)}$  כלומר התקנון של  $T$ .

מצאו את פונקציית יוצרת המומנטים של  $Z$ .

(4) נתון שפונקציית יוצרת מומנטים של ההתפלגות הנורמלית נתונה על ידי

$$M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

הנוסחה הבאה: לכל  $t$ , כאשר:  $X \sim N(\mu, \sigma^2)$ .

א. הוכיחו שאם  $Y = 2X$  אזי  $Y \sim N(2\mu, 4\sigma^2)$ .

ב. הוכיחו שאם  $T = X_1 + X_2$  ו- $X_1$  ו- $X_2$  בלתי תלויים מאותה התפלגות

נורמלית אז מתקיים ש:  $T \sim N(2\mu, 2\sigma^2)$ .

### תשובות סופיות:

(1) א. פונקציה יוצרת מומנטים:  $e^{(n\lambda)(e^t-1)}$ . ב. שאלת הוכחה.

(2) א. פונקציה יוצרת מומנטים:  $e^{12(e^t-1)}$ . ב. שאלת הוכחה.

ג. שאלת הוכחה.

(3) א. פונקציה יוצרת מומנטים:  $\left(\frac{1}{1-t}\right)^n$ . ב. תוחלת:  $n$ , שונות:  $n$ .

ג. פונקציה יוצרת מומנטים:  $e^{-n^2 t} \cdot \left(\frac{1}{1 - \left(\frac{1}{n^2} t\right)}\right)^n$

(4) א. שאלת הוכחה. ב. שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 47 - משתנה דו-מימדי בדיד - פונקציית הסתברות משותפת

תוכן העניינים

1. כללי ..... 203

## משתנה דו מימדי בדיד – פונקציית הסתברות משותפת:

### רקע:

התפלגות דו ממדית הינה התפלגות שדנה בשני משתנים. נרצה כעת לבנות פונקציית הסתברות דו ממדית, בה יש התפלגות של שני משתנים בו זמנית:  $X$  ו- $Y$ .

### דוגמה:

תלמיד ניגש בסמסטר לשני מבחנים: מבחן בכלכלה ומבחן בסטטיסטיקה. כמו כן, נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8, הסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9 והסיכוי לעבור את שני המבחנים הנו 0.75. יהי  $X$  מספר הקורסים שהסטודנט עבר. ויהי  $Y$  משתנה אינדיקטור המקבל את הערך 1 אם הסטודנט עבר את הבחינה בכלכלה, ו-0 אחרת. בנו את פונקציית ההסתברות המשותפת של  $X$  ו- $Y$ .

### נחשב את כל ההסתברויות המשותפות:

$$p(x=0, y=0) = 0.05$$

$$p(x=0, y=1) = 0$$

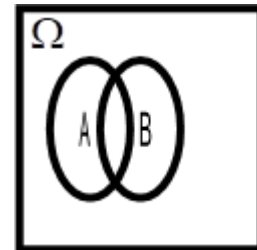
$$p(x=1, y=0) = 0.15$$

$$p(x=1, y=1) = 0.05$$

$$p(x=2, y=0) = 0$$

$$p(x=2, y=1) = 0.75$$

$y/x$	0	1	2
0	0.05	0.15	0
1	0	0.05	0.75



שימו לב שסכום כל ההסתברויות בפונקציית ההסתברות המשותפת הוא 1.

כעת נסכם את השורות ואת העמודות ונקבל את פונקציות הסתברות שוליות:

$Y/X$	0	1	2	$P_Y$
0	0.05	0.15	0	0.2
1	0	0.05	0.75	0.8
$P_X$	0.05	0.2	0.75	1

**משתנים בלתי תלויים:**

$X$  ו- $Y$  יהיו משתנים בלתי תלויים, אם עבור כל  $X$  ו- $Y$  אפשריים התקיים הדבר

$$\text{הבא: } p(x=k, y=l) = p(x=k) \cdot p(y=l).$$

מספיק פעם אחת שהמשתנים אינם מקיימים תנאי זה אזי הם תלויים.

**דוגמה:**

$$p(x=2, y=1) = 0.75 \neq p(x=2) \cdot p(y=1) = 0.75 \cdot 0.8 = 0.6$$

ככלל, אם יש אפס בתוך פונקציית ההסתברות המשותפת ניתן להבין באופן מידי שהמשתנים תלויים, שאז הרי התנאי לא מתקיים. אך אם אין אפס בטבלה, אין זה אומר שהמשתנים בלתי תלויים ויש לבדוק זאת.

**שאלות:**

- (1) אדם נכנס לקזינו עם 75 דולר. הוא ישחק במכונת מזל בה יש סיכוי של 0.3 לנצח. במקרה של ניצחון במשחק הוא יקבל מהקזינו 25 דולר ובמקרה של הפסד הוא ישלם 25 דולר. אותו אדם החליט שיפסיק לשחק ברגע שיהיה לו 100 דולר, אך בכל מקרה לא ישחק יותר מ-3 משחקים. נגדיר את  $X$  להיות הכסף שברשות האדם בצאתו מהקזינו ואת  $Y$  כמספר המשחקים שהאדם שיחק.
- א. בנו את פונקציית ההסתברות המשותפת והשוליות.  
 ב. מה תוחלת מספר המשחקים שישחק האדם?  
 ג. אם האדם יצא מהקזינו שברשותו 100 דולר, מה התוחלת ומהי השונות של מספר המשחקים ששיחק?

- (2) להלן פונקציית ההסתברות המשותפת והשוליות של שני משתנים מקריים בדידים:

$Y / X$	0	1	2	$P(Y)$
2		0.08	0.12	0.4
3	0.1	0.05		
4				0.45
$P(X)$		0.4	0.2	

- א. השלימו את ההסתברויות החסרות בטבלה.  
 ב. האם  $X$  ו- $Y$  תלויים?  
 ג. מצאו את הסתברות ש- $Y = 3$ , אם ידוע ש- $X = 1$ .
- (3) מפעל משווק מוצר הנארז בחבילות בגדלים שונים. ישנו מספר שווה של חבילות בנות שני מוצרים ושלושה מוצרים. ההסתברות שמוצר מסוים יהיה פגום היא  $\frac{1}{10}$ . מהנדס הייצור בוחר באקראי חבילת מוצרים לשם בקרת איכות. יהי  $X$  מספר המוצרים בחבילה, ו- $Y$  מספר המוצרים הפגומים בחבילה.
- א. מה ההתפלגות של המשתנה  $Y$  בהינתן  $x = 3$ .  
 ב. מה ההתפלגות של המשתנה  $Y$  בהינתן  $X$  הינו  $K$  כלשהו.  
 ג. מהי תוחלת מספר המוצרים הפגומים בחבילות בנות 3 מוצרים? נמקו.  
 ד. בנו את פונקציית ההסתברות המשותפת.

- (4) מתוך כד עם 3 כדורים ממוספרים במספרים 2, 4, 8 שולפים באקראי שניים ללא החזרה. יהי  $X$  המספר הקטן מבין השניים ו- $Y$  הגדול מביניהם.
- א. חשבו את ההתפלגות של  $(X, Y)$ .
- ב. אם המספר המינימאלי שנבחר הוא 2, מה הסיכוי שהמקסימאלי הוא 8?
- ג. חשבו את ההתפלגות המותנית של  $X$  בהינתן  $Y = 4$ . מצאו:  $E(X / Y = 4)$ .
- (5) ביישוב שני סניפי בנק. סניף פועלים וסניף לאומי. להלן הנתונים לגבי האוכלוסייה הבוגרת המתגוררת ביישוב: ל-60% יש חשבון בסניף פועלים של היישוב, ל-50% יש חשבון בסניף לאומי של היישוב ול-95% יש חשבון בלפחות אחד מהסניפים. יהי  $X$  מספר הסניפים בישוב אשר לתושב בוגר יש בהם חשבון, ויהי  $Y$  משתנה אינדיקטור:
- 1 – אם יש לתושב חשבון בסניף פועלים.  
 0 – אחרת.
- א. בנו את פונקציית ההסתברות המשותפת של  $X$  ו- $Y$ .
- ב. הוסיפו את פונקציית ההסתברות השולית.
- ג. ידוע שלתושב בוגר חשבון בבנק פועלים, מה ההסתברות שיש לו חשבון בנק בסניף אחד בלבד?

**תשובות סופיות:**

1) א. להלן טבלה: ב. 2.4 ג. תוחלת: 1.348, שונות: 0.575.

$x \setminus y$	0	50	100	$P(y)$
1	0	0	0.3	0.3
3	0.343	0.294	0.063	0.7
$P(x)$	0.343	0.294	0.363	1

2) א. להלן טבלה: ב. תלויים. ג. 0.125.

$x \setminus y$	0	1	2	$P(y)$
2	0.2	0.08	0.12	0.4
3	0.1	0.05	0	0.15
4	0.1	0.27	0.08	0.45
$P(x)$	0.4	0.4	0.2	1

3) א.  $y/x=3 \sim B\left(n=3, p=\frac{1}{10}\right)$  ב.  $y/x=k \sim B\left(n=k, p=\frac{1}{10}\right)$

ג. 0.3 ד. להלן טבלה:

$x \setminus y$	2	3	$P(y)$
0	0.405	0.3645	
1	0.09	0.1215	
2	0.005	0.0135	
3	0	0.0005	
$P(x)$	0.5	0.5	1

4) א. להלן טבלה: ב. 0.5 ג. תוחלת: 2.

$x \setminus y$	2	4	$P(y)$
4	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$
8	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
$P(x)$	$\frac{2}{3}$	$\frac{1}{3}$	1

5) א+ב. להלן טבלה: ג. 0.75.

$x \setminus y$	0	1	2	$P(y)$
0	0.05	0.35	0	0.4
1	0	0.45	0.15	0.6
$P(x)$	0.05	0.8	0.15	1

# סטטיסטיקה והסתברות 01 88165

פרק 48 - משתנה דו מימדי בדיד - מתאם בין משתנים

תוכן העניינים

1. כללי..... 209

## משתנה דו מימדי בדיד – מתאם בין משתנים:

### רקע:

נרצה לבדוק את מידת ההתאמה הלינארית בין שני המשתנים על ידי מקדם המתאם הלינארי שמסומן ב- $\rho$ . מקדם מתאם זה מקבל ערכים בין -1 ל-1.

$$\frac{\quad}{-1 \qquad \qquad \qquad 0 \qquad \qquad \qquad 1}$$

מקדם מתאם -1 או 1 אומר שקיים קשר לינארי מוחלט ומלא בין המשתנים, שניתן לבטאו על ידי הנוסחה:  $y = ax + b$ .

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

מתאם חיובי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחה לינארית שמקשרת את  $X$  ל- $Y$  באופן מוחלט ואילו מתאם שלילי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לרדת אבל לא קיימת נוסחה לינארית שמקשרת את  $X$  ל- $Y$  באופן מוחלט.

### חישוב מקדם המתאם:

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} \quad \text{הנוסחה של מקדם המתאם היא:}$$

$$\text{cov}(x, y) = E[(x - \mu_x)(y - \mu_y)] = E(x \cdot y) - E(x) \cdot E(y) \quad \text{השונויות המשותפת:}$$

תכונות של השונויות המשותפת:

$$\text{cov}(X, Y) = \text{cov}(Y, X) \quad (1)$$

$$\text{cov}(X, X) = \text{Var}(X) \quad (2)$$

### משתנים בלתי מתואמים:

משתנים בלתי מתואמים הם משתנים שמקדם המתאם שלהם אפס, וכדי שדבר כזה יקרה השונויות המשותפת צריכה להתאפס. משתנים בלתי מתואמים הם משתנים שכלל אין ביניהם התאמה לינארית. משתנים בלתי תלויים הם משתנים שאין ביניהם קשר ולכן גם הם בלתי מתואמים, אך משתנים בלתי מתואמים אינם בהכרח בלתי תלויים.

### השפעת טרנספורמציה לינארית על מקדם המתאם:

$$\rho[(aX+b), (cY+d)] = \begin{cases} \rho(X,Y) & \text{if } a \cdot c > 0 \\ -\rho(X,Y) & \text{if } a \cdot c < 0 \end{cases}$$

כלומר, טרנספורמציה לינארית על שני משתנים לא משנה את עוצמת הקשר ביניהם היא עלולה לשנות רק את כיוון הקשר.

### דוגמה (פתרון בהקלטה):

נחזור לדוגמה שהוצגה בפרק הקודם:

תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ומבחן בסטטיסטיקה. כמו כן, נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8, הסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9 והסיכוי לעבור את שני המבחנים הנו 0.75.

יהי  $X$  מספר הקורסים שהסטודנט עבר, ויהי  $Y$  משתנה אינדיקטור המקבל את הערך 1, אם הסטודנט עבר את הבחינה בכלכלה, ו-0 אחרת.

נחשב את מקדם המתאם:

$X/Y$	0	1	2	$P_Y$
0	0.05	0.15	0	0.2
1	0	0.05	0.75	0.8
$P_X$	0.05	0.2	0.75	1

$X$	0	1	2
$P_X$	0.05	0.2	0.75

$$E(X) = \sum_i x_i P(x_i) = \mu = 0 \cdot 0.05 + 1 \cdot 0.2 + 2 \cdot 0.75 = 1.7$$

$$V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = 0^2 \cdot 0.05 + 1^2 \cdot 0.2 + 2^2 \cdot 0.75 - 1.7^2 = 0.31 = \sigma^2$$

$y$	$P_Y$
0	0.2
1	0.8

$$\sigma_x = \sqrt{V(X)} = \sqrt{0.31} = 0.557$$

$$E(y) = \sum_i y_i P(y_i) = 0 + 0.8 = 0.8$$

$$V(y) = \sum_i (y_i - \mu_y)^2 P(y_i) = \sum_i y_i^2 P(y_i) - \mu_y^2 = 0 + 0.8 - 0.8^2 = 0.16 = \sigma_y^2$$

$$\sigma_y = \sqrt{0.16} = 0.4$$

$$E(xy) = 0 \cdot 0 \cdot 0.05 + 0 \cdot 1 \cdot 0 + 1 \cdot 0 \cdot 0.15 + 1 \cdot 1 \cdot 0.05 + 2 \cdot 0 \cdot 0 + 2 \cdot 1 \cdot 0.75 = 1.55$$

$$\text{cov}(x, y) = E(x \cdot y) - E(x) \cdot E(y) = 1.55 - 1.7 \cdot 0.8 = 0.19$$

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} = \frac{0.19}{0.557 \cdot 0.4} = 0.853$$

כל קורס שהסטודנט מסיים מזכה אותו ב-3 נקודות אקדמאיות.  
מה יהיה מקדם המתאם בין נקודות הזכות שיצבור למשתנה  $Y$ ?

### שאלות:

- (1) הסיכוי שסטודנט יעבור את המבחן במועד א' בסטטיסטיקה הוא 0.8. אם הוא נכשל במועד א' הוא ניגש למועד ב' שם הסיכוי לעבור את המבחן מוערך ב-0.9 (סטודנט שעובר את א' לא ניגש לב'). במידה והסטודנט נכשל במועד ב' הוא מגיש בקשה למועד ג' אותה מאשרים בסיכוי של 0.2, והסיכוי שלו לעבור את מועד ג' הוא 0.7.
- נגדיר את  $X$  להיות מספר המבחנים אליהם ניגש הסטודנט, ונגדיר את  $Y$  להיות מספר המבחנים שנכשל בהם.
- בנו את פונקציית ההסתברות המשותפת ואת פוני ההסתברות השולית.
  - האם המשתנים הינם בלתי תלויים?
  - ידוע שהסטודנט ניגש ליותר ממבחן אחד, מה ההסתברות שהוא נכשל בפחות משלושה מבחנים?
  - האם המתאם בין  $X$  ל- $Y$  מלא או חלקי? חיובי או שלילי? הסבירו ללא חישוב.
  - חשבו את מקדם המתאם בין  $X$  לבין  $Y$ .
  - האם המשתנים הם בלתי מתאומים?
- (2) נטיל מטבע שלוש פעמים. נגדיר את  $X$  להיות מספר העצים המתקבלים בשתי ההטלות הראשונות, ואת  $Y$  להיות מספר העצים המתקבלים בשתי ההטלות האחרונות.
- בנו את פונקציית ההסתברות המשותפת של  $X$  ו- $Y$  ואת פונקציות ההסתברות השוליות.
  - האם  $X$  ו- $Y$  הם משתנים בלתי תלויים?
  - מהו מקדם המתאם בין  $X$  ל- $Y$ . האם המשתנים מתאומים?
  - אם בשתי ההטלות הראשונות יצא בדיוק עץ אחד, מה ההסתברות שבשתי ההטלות האחרונות יצאו שני עצים?
  - אם בשתי ההטלות האחרונות יצא לפחות פעם אחת עץ, מה ההסתברות שבשתי ההטלות הראשונות יצא עץ אחד?
- (3) נפזר שלושה כדורים שונים בשלושה תאים. נגדיר את המשתנים הבאים:
- $X$  - מספר הכדורים בתא הראשון.
  - $Y$  - מספר הכדורים בתא השני.
- בנו את פונקציית ההסתברות המשותפת.
  - האם המשתנים בלתי מתאומים?

- 4) קובייה הוגנת הוטלה פעמיים. יהי  $X$  ההטלה הגדולה מבין שתי התוצאות, ויהי  $Y$  מס' ההטלות בהן יצאה תוצאה זוגית.
- מצאו את פונקציית ההסתברות המשותפת של  $X$  ו- $Y$ .
  - חשבו את מקדם המתאם של  $X$  ו- $Y$ .
  - מצאו את ההתפלגות של  $Y$  בהינתן ש- $X = 2$ .
- 5) בבניין שלנו 5 דירות. דירות מספר אחת ושלוש הן דירות משופצות והשאר אינן. הוחלט לבחור שתי דירות באקראי מבין הדירות בבניין. נגדיר את המשתנים הבאים:
- $X$  - מספר הדירות המשופצות שנבחרו.
  - $Y$  - מספר הדירות האי זוגיות שנדגמו.
- בנו את פונקציית ההסתברות המשותפת ואת פונקציית ההסתברות השולית.
  - האם המשתנים מתואמים?
  - מה מקדם המתאם בין  $X$  לבין  $Y$ ?
  - מה יהיה מקדם המתאם:
- בין מספר הדירות המשופצות למספר הדירות הזוגיות שנדגמו.
  - בין מספר הדירות הזוגיות לדירות האי זוגיות שנדגמו.
- ה. כל דירה משופצת עולה 2 מיליון ₪ וכל דירה לא משופצת עולה 1.5 מיליון ₪. מה המתאם בין עלות הדירות שנדגמו למספר הדירות הזוגיות?

**תשובות סופיות:**

(1) א. להלן טבלה: ב. תלויים. ג. 0.994 ד. חלקי חיובי.

$x \setminus y$	1	2	3	$P(y)$
0	0.8	0	0	0.8
1	0	0.18	0	0.18
2	0	0.016	0.0028	0.0188
3	0	0	0.0012	0.0012
$P(x)$	0.8	0.196	0.004	1

ה. 0.963 ו. מתואמים.

(2) א. להלן טבלה: ב. תלויים. ג. מקדם המתאם: 0.5, מתואמים.

$x \setminus y$	0	1	2	$P(y)$
0	$\frac{1}{8}$	$\frac{1}{8}$	0	$\frac{2}{8}$
1	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{4}{8}$
2	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$
$P(x)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	1

ד. 0.25 ה. 0.5

(3) א. להלן טבלה: ב. מתואמים.

$x \setminus y$	0	1	2	3
0	$\frac{1}{27}$	$\frac{3}{27}$	$\frac{3}{27}$	$\frac{1}{27}$
1	$\frac{3}{27}$	$\frac{6}{27}$	$\frac{3}{27}$	0
2	$\frac{3}{27}$	$\frac{3}{27}$	0	0
3	$\frac{1}{27}$	$\frac{6}{27}$	0	0

4) א. להלן טבלה: ב. 0.252.

$x \setminus y$	1	2	3	4	5	6
0	$\frac{1}{36}$	0	$\frac{3}{36}$	0	$\frac{5}{36}$	0
1	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{4}{36}$	$\frac{6}{36}$
2	0	$\frac{1}{36}$	0	$\frac{3}{36}$	0	$\frac{5}{36}$

ג.  $\frac{2}{3}$ .

5) א. להלן טבלה: ב.  $X$  ו- $Y$  מתואמים.

$x \setminus y$	0	1	2	$P(y)$
0	0.1	0	0	0.1
1	0.2	0.4	0	0.6
2	0	0.2	0.1	0.3
$P(x)$	0.3	0.6	0.1	1

ה.  $-\frac{2}{3}$ .

ii. -1.

ד. i.  $-\frac{2}{3}$ .

# סטטיסטיקה והסתברות 01 88165

פרק 49 - המשתנה המקרי הדו ממדי - קומבינציות ליניאריות

תוכן העניינים

1. כללי ..... 216

## המשתנה המקרי הדו ממדי – קומבינציות לינאריות:

**רקע:**

יהיו שני משתנים מקריים  $X$  ו- $Y$ .  
התוחלת והשונות של סכומם היא:

$$E(X+Y) = E(X) + E(Y)$$

$$V(X+Y) = V(X) + V(Y) + 2 \cdot \text{cov}(X, Y)$$

התוחלת והשונות של הפרשם היא:

$$E(X-Y) = E(X) - E(Y)$$

$$V(X-Y) = V(X) + V(Y) - 2 \cdot \text{cov}(X, Y)$$

**קומבינציה ליניארית:**

יוצרים משתנה חדש שהוא קומבינציה ליניארית של שני משתנים אחרים:  
 $W = (aX + b) + (cY + d)$ . אזי:

$$\text{cov}[(aX + b), (cY + d)] = a \cdot c \cdot \text{cov}(X, Y)$$

$$E(W) = E((aX + b) + (cY + d)) = aE(X) + b + cE(Y) + d$$

$$V(W) = V((aX + b) + (cY + d)) = a^2V(X) + c^2V(Y) + 2 \cdot a \cdot c \cdot \text{cov}(X, Y)$$

**דוגמה (פתרון בהקלטה):**

- נתונים שני משתנים מקריים  $X$  ו- $Y$  המקיימים:  
 $\mu_X = 80$ ,  $\sigma_X = 15$ ,  $\mu_Y = 70$ ,  $\sigma_Y = 20$ ,  $\text{cov}(X, Y) = 200$
- מצאו את התוחלת והשונות של סכום המשתנים.
  - מצאו את התוחלת והשונות של  $X$  ו- $Y$ .
  - מצאו את השונות ומה התוחלת של המשתנה  $W = 2X + 3Y$ .

**שאלות:**

(1) נתונה פונקציית ההסתברות המשותפת הבאה:

$Y / X$	1	2	3	$P(X)$
2		0.1	0.3	0.6
3	0.2		0.1	
$P(X)$				

- השלימו את ההסתברויות החסרות.
- האם המשתנים תלויים?
- האם המשתנים בלתי מתואמים?
- חשבו את השונות המשותפת.
- חשבו את התוחלת והשונות של סכום המשתנים.
- חשבו את התוחלת והשונות של הפרש המשתנים.

(2) מבחן בנוי מחלק כמותי וחלק מילולי. תוחלת הציון בחלק הכמותי היא 100, עם סטיית תקן 20. תוחלת הציונים בחלק המילולי היא 90 עם סטיית תקן 15. מקדם המתאם בין הציון הכמותי לציון המילולי הוא 0.8.

- חשבו את השונות המשותפת בין הציון הכמותי לציון המילולי.
- חשבו את התוחלת והשונות של סכום הציונים בחלק הכמותי ובחלק המילולי.
- חשבו את התוחלת והשונות של הפרש הציונים בין החלק הכמותי לחלק המילולי.
- עלות הבחינה 2000 שקלים. הוחלט לזכות שקל עבור כל נקודה שנצברה בחלק המילולי ושני שקלים עבור כל נקודה שנצברה בחלק הכמותי. מהי התוחלת ומהי השונות של עלות הבחינה נטו (העלות לאחר הזיכוי)?

(3) נתון:  $\text{var}(X + 2Y) = 3$ ,  $\text{var}(X - 2Y) = 2$ .  
חשבו:  $\text{cov}(X, Y)$ .

(4) מטילים קובייה  $n$  פעמים. נגדיר את המשתנים הבאים:  
 $X$  = מספר הפעמים שהתקבלה התוצאה 6.  
 $Y$  = מספר הפעמים שהתקבלה התוצאה 5.  
 בטאו את השונות המשותפת באמצעות  $n$ .

## תשובות סופיות:

(1) א. להלן טבלה: ב. תלויים. ג. מתואמים. ד.  $-0.1$ .

$x \setminus y$	1	2	3	$P(y)$
2	0.2	0.1	0.3	0.6
3	0.2	0.1	0.1	0.4
$P(x)$	0.4	0.2	0.4	1

- ה. תוחלת: 4.4, שונות: 0.84. ו. תוחלת:  $-0.4$ , שונות: 1.24.
- (2) א. 240. ב. תוחלת: 190, שונות: 1105.
- ג. תוחלת: 10, שונות: 145. ד. תוחלת: 1710, שונות: 2785.
- (3)  $-0.125$ .
- (4)  $-\frac{n}{36}$ .

# סטטיסטיקה והסתברות 01 88165

פרק 50 - המשתנה המקרי הדו ממדי הבדיד - שאלות מסכמות

תוכן העניינים

1. שאלות מסכמות.....219

## המשתנה המקרי הדו ממדי הבדיד – שאלות מסכמות:

**רקע:**

**משתנים בלתי תלויים:**

יהיו משתנים  $X$  ו- $Y$ . הם יהיו משתנים בלתי תלויים אם עבור כל  $X$  ו- $Y$  אפשריים מתקיים:  $p(x=k, y=l) = p(x=k) \cdot p(y=l)$ .

**מקדם המתאם:**

מגדירים את מקדם המתאם:  $\rho = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y}$ .

**שוונות משותפת:**

מגדירים את השונות המשותפת:

$$\text{cov}(x, y) = E[(x - \mu_x)(y - \mu_y)] = E(x \cdot y) - E(x) \cdot E(y)$$

תכונות של השונות המשותפת:

$$1. \text{cov}(X, Y) = \text{cov}(Y, X)$$

$$2. \text{cov}(X, X) = \text{var}(X)$$

$$3. \text{cov}[(aX + b), (cY + d)] = a \cdot c \cdot \text{cov}(X, Y)$$

**משתנים בלתי מתואמים:**

משתנים בלתי מתואמים הם משתנים שמקדם המתאם שלהם אפס וכדי שדבר כזה יקרה השונות המשותפת צריכה להתאפס.

השפעת טרנספורמציה לינארית על מקדם המתאם:

$$\rho[(aX+b), (cY+d)] = \begin{cases} \rho(X,Y) & \text{if } a \cdot c > 0 \\ -\rho(X,Y) & \text{if } a \cdot c < 0 \end{cases}$$

תוחלת ושונות של סכום משתנים:

$$E(X+Y) = E(X) + E(Y) \quad V(X+Y) = V(X) + V(Y) + 2 \cdot \text{cov}(X, Y)$$

קומבינציות לינאריות:

נגדיר קומבינציה ליניארית כללית באופן הבא:  $W = (aX+b) + (cY+d)$   
אזי מתקיים:

$$E(W) = E((aX+b) + (cY+d)) = aE(X) + b + cE(Y) + d$$

$$V(W) = V((aX+b) + (cY+d)) = a^2V(X) + c^2V(Y) + 2 \cdot a \cdot c \cdot \text{cov}(X, Y)$$

## שאלות:

- (1) יש ליצור סיסמה בת 3 תווים. כל תו יכול להיבחר רק מתוך כלל התווים הבאים:  $1, 2, C, B, A$ . יהי  $X$  מספר הפעמים שהספרה 1 מופיעה בסיסמה, ויהי  $Y$  מספר הפעמים שהספרה 1 מופיעה בקצה הסיסמה (שני הקצוות).
- זהו את ההתפלגויות השוליות של  $X$  ו- $Y$  כהתפלגויות מיוחדות.
  - מצאו את ההתפלגות המשותפת של  $X$  ושל  $Y$ .
  - מצאו את מקדם המתאם בין  $X$  ל- $Y$ .
  - מהו המתאם בין  $2X$  ל- $3Y+5$ ?
- (2) במסיבת סוף שנה ישנו ארגז קרח ובתוכו 7 בקבוקי בירה: 4 "מכבי", 2 "גולדסטאר" ו-1 "טבורג". קרן לקחה 3 בקבוקי בירה באקראי מתוך ארגז הקרח. נסמן ב- $X$  את מספר בקבוקי "מכבי" שנלקחו על ידי קרן, ונסמן ב- $Y$  את מספר בקבוקי "טבורג" שנלקחו על ידי קרן.
- בנו את פונקציית ההסתברות המשותפת של  $X$  ושל  $Y$ .
  - חשבו את התוחלת והשונות של  $X$  ושל  $Y$ .
  - מצאו את השונות המשותפת של  $X$  ושל  $Y$ .
  - נגדיר את  $W$  כמספר בקבוקי ה"גולדסטאר" שנלקחו על ידי קרן. בטאו את  $W$  באמצעות  $X$  ו- $Y$ , וחשבו את התוחלת והשונות של  $W$  על סמך התוצאות שהתקבלו בשני הסעיפים הקודמים בלבד.
  - מהו מקדם המתאם בין מספר בקבוקי "מכבי" שנלקחו על ידי קרן, למספר בקבוקים שאינם "מכבי" שנלקחו על ידי קרן?
- (3) במגירה 6 זוגות נעליים. יהודה הוציא מהמגירה 4 נעליים (לא בהכרח זוגות) באקראי. נסמן ב- $W$  את מספר זוגות הנעליים שהוציא יהודה, ונסמן ב- $R$  את מספר הנעליים השמאליות שהוציא יהודה.
- מצא את ההתפלגות המשותפת של המשתנים שהוצגו.
  - האם המשתנים שהוצגו בלתי תלויים?
  - מצא את התפלגות מספר הנעליים השמאליות שהוצאו אם בסך הכול הוצא זוג נעלים יחיד על ידי יהודה.
  - אם ידוע שהוצאו לפחות 3 נעליים שמאליות מה הסיכוי שהוצא לכל היותר זוג אחד?

- (4) בכד 5 כדורים כחולים, 4 כדורים לבנים ו-3 כדורים ירוקים. בוחרים באקראי וללא החזרה 3 כדורים. נגדיר את המשתנים הבאים:
- $X$  - מקבל את הערך 1 אם נבחר לפחות כדור אחד כחול, ו-0 אחרת.  
 $Y$  - מספר הכדורים הלבנים שנבחרו.
- א. חשבו את  $P(X=1)$ .
- ב. בנו את פונקציית ההסתברות המשותפת של  $X$  ו- $Y$ .
- ג. מה התוחלת של  $Y$ , אם ידוע שלא הוצאו כדורים כחולים?
- ד. מה השונות של  $X$ , אם ידוע שהוצא לכל היותר כדור לבן אחד?
- (5) ביום ההולדת הרביעי של טל הוא מחלק שלושה פרסים שונים באקראי ל-5 ילדים. בכל פעם שטל מחלק פרס הוא בוחר באקראי ילד מתוך ה-5 באופן אקראי ובלתי תלוי בבחירות הקודמות. נגדיר את המשתנים הבאים:
- $X$  - מספר הפרסים שקיבלה יוליה.  
 $Y$  - מספר הילדים שלא קיבלו פרס.
- א. בנו את פונקציית ההסתברות המשותפת והשוליות של  $X$  ו- $Y$ .
- ב. האם  $X$  ו- $Y$  הם משתנים בלתי מתואמים?
- ג. מצאו את התוחלת של  $X \cdot Y^2$ .
- ד. מה מקדם המתאם בין מספר הפרסים שקיבלה יוליה, למספר הילדים שקיבלו פרס?
- (6) קבעו אילו מהטענות הבאות נכונות. נמקו.
- א. אם שני משתנים הם מתואמים, אזי הם תלויים.  
 ב. אם שני משתנים הם תלויים, אזי הם מתואמים.  
 ג. אם שני משתנים הם בלתי תלויים, אזי הם בלתי מתואמים.  
 ד. אם שני משתנים הם בלתי מתואמים, אזי הם בלתי תלויים.
- (7) במקום עבודה 50 עובדים מתוכם 25 גברים ו-25 נשים. כל עובד נתבקש לבחור מתנה לחג. לכל עובד מוצגות 5 אופציות, מתוכן הוא צריך לבחור אחת. העובדים בוחרים מתנה באקראי ובאופן בלתי תלוי זה בזה.
- נסמן  $X_i$  - מספר הגברים שבחרו במתנה  $i$ .  
 נסמן  $Y_i$  - מספר הנשים שבחרו במתנה  $i$ .
- א. האם  $X_1$  ו- $Y_1$  הם משתנים בלתי תלויים? אין צורך לחשב רק להסביר.  
 ב. האם  $X_1$  ו- $X_2$  הם משתנים בלתי תלויים? אין צורך לחשב רק להסביר.  
 ג. מהי ההתפלגות של  $X_1 + X_2$ ?  
 ד. האם המתאם בין  $X_1$  ו- $X_2$  מלא או חלקי? חיובי או שלילי?  
 אין צורך לחשב רק להסביר.

(8) הוכיחו את הזהות הבאה עבור שלושת המשתנים  $X, Y$  ו- $Z$  :  
 $\text{cov}(X+Y, Z) = \text{cov}(X, Z) + \text{cov}(Y, Z)$ .

(9) מספר העלים שנושרים בסתיו מהעץ בגינה מתפלג פואסונית עם תוחלת של 50 עלים בדקה. נסמן ב- $Y$  את מספר העלים שנושרים מהעץ בין 12:00 ל-12:10, ונסמן ב- $Q$  את מספר העלים שנושרים בין 12:05 ל-12:30.  
 א. חשבו את:  $\text{cov}(4Y, Q+6)$ .  
 ב. מה המתאם בין  $Y$  ל- $Q$ ?

(10) בסל יש 20 כדורים אדומים, 20 ירוקים ו-20 כחולים. מוציאים באקראי מהסל 20 כדורים. מצאו את מקדם המתאם בין מספר הכדורים האדומים שהוצאו למספר הכדורים הירוקים שהוצאו.

(11) נתון ש:  $Y \sim B(1, p)$  כאשר  $0 < p < 1$ .  
 הוכיחו שאם מתקיים:  $P(X = x | Y = 0) = P(X = x | Y = 1)$  לכל  $X$ , אז  $X$  ו- $Y$  הם משתנים בלתי תלויים.

(12) נתון ש- $X \sim B(n, p)$  וכן:  $Y \sim B(m, p)$ , שאינם תלויים זה בזה.  
 הוכיחו שמתקיים:  $X | X+Y = k \sim HG(n+m, n, k)$ .

## תשובות סופיות:

$$(1) \quad X \sim B\left(3, \frac{1}{5}\right), Y \sim B\left(2, \frac{1}{5}\right) \quad \text{א.}$$

ב. להלן טבלה: ג. 0.816 ג. ד. 0.816.

$X/Y$	0	1	2	3	$P_Y$
0	$\frac{64}{125}$	$\frac{16}{125}$	0	0	$\frac{80}{125}$
1	0	$\frac{32}{125}$	$\frac{8}{125}$	0	$\frac{40}{125}$
2	0	0	$\frac{4}{125}$	$\frac{1}{125}$	$\frac{5}{125}$
$P_X$	$\frac{64}{125}$	$\frac{48}{125}$	$\frac{12}{125}$	$\frac{1}{125}$	1

$$(2) \quad \text{א. להלן טבלה: ב. } E(X) = \frac{12}{7}, V(X) = \frac{24}{49}, E(Y) = \frac{3}{7}, V(Y) = \frac{12}{49}.$$

$X/Y$	0	1	2	3	$P_Y$
0	0	$\frac{3}{35}$	$\frac{12}{35}$	$\frac{4}{35}$	$\frac{20}{35}$
1	$\frac{1}{35}$	$\frac{8}{35}$	$\frac{6}{35}$	0	$\frac{15}{35}$
$P_X$	$\frac{1}{35}$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{4}{35}$	1

$$\text{ג. } -\frac{8}{49} \quad \text{ד. } E(W) = \frac{6}{7}, V(W) = \frac{20}{49} \quad \text{ה. } -1.$$

(3) א. להלן טבלה: ב. המשתנים תלויים.

$R/W$	0	1	2	$P_R$
0	$\frac{15}{495}$	0	0	$\frac{15}{495}$
1	$\frac{60}{495}$	$\frac{60}{495}$	0	$\frac{120}{495}$
2	$\frac{90}{495}$	$\frac{120}{495}$	$\frac{15}{495}$	$\frac{225}{495}$
3	$\frac{60}{495}$	$\frac{60}{495}$	0	$\frac{120}{495}$
4	$\frac{15}{495}$	0	0	$\frac{15}{495}$
$P_W$	$\frac{240}{495}$	$\frac{240}{495}$	$\frac{15}{495}$	1

ג. להלן טבלה: ד. 1.

$R/w = 1$	1	2	3
$P(R/w = 1)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

א.  $\frac{185}{220}$  ב. להלן טבלה: ג. 1.714 ד. 0.071

$X/Y$	0	1	$P_Y$
0	$\frac{1}{220}$	$\frac{55}{220}$	$\frac{56}{220}$
1	$\frac{12}{220}$	$\frac{100}{220}$	$\frac{112}{220}$
2	$\frac{18}{220}$	$\frac{30}{220}$	$\frac{48}{220}$
3	$\frac{4}{220}$	0	$\frac{4}{220}$
$P_X$	$\frac{35}{220}$	$\frac{185}{220}$	1

א. להלן טבלה: ב.  $X$  ו- $Y$  בלתי מתואמים. (5)

$X/Y$	0	1	2	3	$P_Y$
2	$\frac{24}{125}$	$\frac{36}{125}$	0	0	$\frac{60}{125}$
3	$\frac{36}{125}$	$\frac{12}{125}$	$\frac{12}{125}$	0	$\frac{60}{125}$
4	$\frac{4}{125}$	0	0	$\frac{1}{125}$	$\frac{5}{125}$
$P_X$	$\frac{64}{125}$	$\frac{48}{125}$	$\frac{12}{125}$	$\frac{1}{125}$	1

ג. 4.128 ד. 0.

א. נכון. ב. לא נכון. ג. נכון. ד. לא נכון. (6)

- 7) א. בלתי תלויים.      ב. תלויים.      ג.  $x_1 + x_2 \sim B\left(n = 25, p = \frac{2}{5}\right)$ .
- ד. חלקי ושלילי.
- 8) שאלת הוכחה.
- 9) א. 1000.      ב. 0.316.
- 10) -0.5.
- 11) שאלת הוכחה.
- 12) שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 51 - קומבינציות לינאריות על התפלגות נורמלית

תוכן העניינים

1. כללי ..... 227

## קומבינציות לינאריות על התפלגות נורמלית:

**רקע:**

כל קומבינציה לינארית של משתנים המתפלגים נורמאלית – מתפלגת נורמאלית בעצמה.

**דוגמה (פתרון בהקלטה):**

הגובה של גברים במדינת ישראל מתפלג נורמלית עם תוחלת של 175 ס"מ וסטיית תקן של 10 ס"מ, וגובהן של הנשים במדינה מתפלג נורמלית עם תוחלת של 165 ס"מ וסטיית תקן של 8 ס"מ.  
מה הסיכוי שגבר אקראי במדינה יהיה גבוה מאישה אקראית?

## שאלות:

- (1) המשקל של גברים במדינת ישראל מתפלג נורמלית עם תוחלת של 75 ק"ג וסטיית תקן של 10 ק"ג, והמשקל של נשים במדינה מתפלג נורמלית עם תוחלת של 65 ק"ג וסטיית תקן של 8 ק"ג. מה הסיכוי שאישה אקראית תהיה בעלת משקל גבוה יותר מגבר אקראי?
- (2) ההוצאה השנתית על ביגוד לאדם מתפלגת נורמלית עם תוחלת של 3000 ₪ וסטיית תקן של 1000 ₪. ההוצאה השנתית על בילויים מתפלגת נורמלית עם תוחלת של 4000 ₪ וסטיית תקן של 1500 ₪. מקדם המתאם בין ההוצאה השנתית על ביגוד וההוצאה השנתית על בילויים הינו 0.6.  
 א. מה התוחלת ומהי סטיית התקן של התפלגות ההוצאה השנתית הכוללת על ביגוד ובילוי?  
 ב. מה הסיכוי שההוצאה השנתית הכוללת על ביגוד ובילוי תעלה על 8000 ₪?  
 ג. מהו העשירון העליון של ההוצאה השנתית הכוללת על ביגוד ובילוי?
- (3) צריכת הירקות היומית במסעדה מתפלג נורמלית עם תוחלת של 50 ק"ג וסטיית תקן של 4 ק"ג. נתון שמחיר ק"ג ירק הוא 6 ₪ לקילו.  
 א. מה התוחלת ומהי השונות של העלות היומית של ירקות למסעדה?  
 ב. מה ההסתברות שהעלות היומית על ירקות תהיה נמוכה מ-290 ₪?  
 ג. מהו האחוזון ה-40 של התפלגות העלות היומית של המסעדה על ירקות?
- (4) נפח יין בבקבוק מתפלג נורמלית עם תוחלת של 750 מ"ל וסטיית תקן של 20 מ"ל. אדם קנה מארז של 4 בקבוקי יין.  
 א. מהי התוחלת ומהי סטיית התקן של נפח היין במארז.  
 ב. את היין שבמארז האדם מזג לכלי שקיבולתו 3.1 ליטר. מה ההסתברות שהיין יגלוש מהכלי?
- (5) לדוד משה הייתה חווה. בחווה פרה ועיזה. תנובת החלב של הפרה מתפלג נורמלית עם ממוצע של 20 ליטר ביום וסטיית תקן של 5 ליטר ותנובת החלב של העזה מתפלג גם כן נורמלית עם ממוצע של 10 ליטר וסטיית תקן של 2 ליטר. כל ליטר חלב פרה נמכר ב-2 ₪ וליטר חלב עזה נמכר ב-3 ₪.  
 א. מה הסיכוי שהפדיון היומי של דוד משה מחלב יהיה לפחות 62 ₪?  
 ב. מה הסיכוי שמתוך 5 ימים יהיו לפחות 4 ימים בהם תנובת החלב מהפרה והעזה ביחד תהיה מתחת ל-30 ליטר?  
 מה הסיכוי שביום מסוים תנובת הפרה תהיה נמוכה מתנובת העזה?

### תשובות סופיות:

- (1) 0.2177
- (2) א. תוחלת: 7000, סטיית תקן: 2247. ב. 0.3264. ג. 0.9881
- (3) א. תוחלת: 300, שונות: 576. ב. 0.3372. ג. 0.294
- (4) א. תוחלת: 3000 מ"ל, סטיית תקן: 40 מ"ל. ב. 0.0062
- (5) א. 0.7549. ב. 0.1875. ג. 0.0314

# סטטיסטיקה והסתברות 01 88165

פרק 52 - המשתנה המקרי הדו ממדי הרציף

תוכן העניינים

1. משתנה דו ממדי רציף.....230

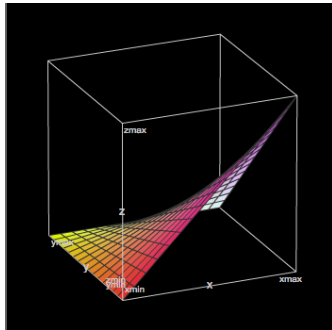
## משתנה מקרי דו ממדי רציף:

### רקע:

יהיו  $X$  ו- $Y$  משתנים מקריים רציפים המוגדרים בתחום  $R$  מסוים.  
 פונקציית הצפיפות המשותפת שלהם תסומן על ידי  $f(x, y)$ .  
 פונקציית צפיפות משותפת צריכה לקיים את שני התנאים הבאים:

$$(1) \quad f(x, y) \geq 0 \quad \text{לכל } (x, y) \in R$$

$$(2) \quad \iint_R f(x, y) dx dy = 1$$



### דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} : \text{נתונה הפונקציה}$$

הראו שפונקציה זו יכולה להיות פונקציית צפיפות משותפת.

### פונקציית צפיפות שולית:

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy : \text{פונקציית הצפיפות השולית של } X \text{ תתקבל באופן הבא}$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx : \text{פונקציית הצפיפות השולית של } Y \text{ תתקבל באופן הבא}$$

### דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} : \text{מצאו לפונקציית הצפיפות}$$

את פונקציית הצפיפות השולית של  $X$ , וחשבו את  $E(X)$  דרכה.

### אי-תלות בין משתנים רציפים:

$X$  ו- $Y$  יהיו משתנים מקרים בלתי תלויים, אם עבור כל  $X$  ו- $Y$  בתחום ההגדרה  $R$   
 מתקיים ש:  $f(x, y) = f(x) \cdot f(y)$

**דוגמה (פתרון בהקלטה):**

האם  $X$  ו- $Y$ , המתפלגים לפי פונקציית הצפיפות המשותפת:

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases}$$

הם משתנים בלתי תלויים?

**חישוב הסתברויות עבור משתנה מקרי רציף דו ממדי:**

הנפח הכלוא מתחת למשטח  $f(x, y)$  בתחום מסוים ייתן את ההסתברות ש- $X$

$$\text{ו-} Y \text{ יהיו בתחום הזה: } P[(x, y) \in A] = \iint_A f(x, y) dx dy$$

**דוגמה (פתרון בהקלטה):**

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \text{ : משתנים מתפלגים לפי פונקציית הצפיפות:}$$

חשבו את הסיכוי:  $P(X < 0.5 \cap Y < 0.5)$ .

**פונקציית התפלגות מצטברת משותפת:**

פונקציית התפלגות מצטברת משותפת הינה פונקציה של שני משתנים רציפים המחזירה את הסיכוי שהמשתנים יהיו קטנים מערכים מסוימים:

$$F(s, t) = P(X \leq s \cap Y \leq t) = \int_{-\infty}^t \int_{-\infty}^s f(x, y) dx dy$$

**דוגמה (פתרון בהקלטה):**

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \text{ : משתנים מתפלגים לפי פונקציית הצפיפות:}$$

מצאו את פונקציית ההתפלגות המצטברת המשותפת

ועל פיה חשבו את הסיכוי:  $P(X < 0.5 \cap Y < 0.5)$ .

**פונקציית צפיפות מותנית:**

אם ל- $X$  ול- $Y$  ישנה פונקציית צפיפות משותפת  $f(x, y)$ , אז מגדירים את פונקציית הצפיפות המותנית של  $X$ , בהינתן ש- $Y = y$  לכל ערכי  $y$

$$. f(x|y) = \frac{f(x, y)}{f(y)} \quad : f(y) > 0 \text{ על ידי}$$

באופן דומה, פונקציית הצפיפות המותנית של  $Y$  בהינתן ש- $X = x$  לכל ערכי  $x$

$$. f(y|x) = \frac{f(x, y)}{f(x)} \quad : f(x) > 0 \text{ על ידי}$$

**דוגמה (פתרון בהקלטה):**

$$. f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \quad : \text{משתנים מתפלגים לפי פונקציית הצפיפות}$$

. מצאו את:  $f(x|y)$

**תוחלת מותנית:**

ל- $X$  ול- $Y$  ישנה פונקציית צפיפות משותפת  $f(x, y)$ .

$$. E(X|Y=y) = \int_{-\infty}^{\infty} x \cdot f(x|y) dx \quad : \text{תהיה } Y = y \text{ בהינתן ש-} Y = y \text{ תהיה}$$

ובאופן דומה, התוחלת של  $Y$  בהינתן ש- $X = x$  תהיה:

$$. E(Y|X=x) = \int_{-\infty}^{\infty} y \cdot f(y|x) dy$$

**דוגמה (פתרון בהקלטה):**

$$. f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \quad : \text{משתנים מתפלגים לפי פונקציית הצפיפות}$$

. מצאו את:  $E(X|Y)$

## שאלות:

- (1) נתונה פונקציית הצפיפות הבאה:  $f(x, y) = x + y$ , המוגדרת בתחום שבו:  $0 \leq x \leq 1$  וגם:  $0 \leq y \leq 1$ . הוכיחו שמדובר בפונקציית צפיפות.
- (2) נתונה פונקציית הצפיפות הבאה:  $f(x, y) = Ax(x - y)$ , המוגדרת בתחום שבו:  $0 \leq x \leq 2$  וגם:  $-x \leq y \leq x$ . מצאו את ערכו של הפרמטר  $A$ .
- (3) נתונה פונקציית הצפיפות הבאה:  $f(x, y) = \frac{(x \cdot y)^3 + x \cdot y}{C}$ , המוגדרת בתחום שבו:  $0 \leq x \leq 1$  וגם:  $0 \leq y \leq 1$ .  
 א. מצאו את ערכו של  $C$ .  
 ב. מצאו את  $f(y)$ .  
 ג. האם  $X$  ו- $Y$  הינם משתנים בלתי תלויים?
- (4) משתנה מקרי דו ממדי מתפלג לפי פונקציית הצפיפות הבאה:  $f(x, y) = \frac{1}{800}$ , המוגדרת בתחום שבו:  $60 \leq x \leq y$  וגם:  $60 \leq y \leq 100$ .  
 א. הראו שפונקציה זו מקיימת את התנאים של פונקציית צפיפות.  
 ב. מצאו את פונקציית הצפיפות השולית של  $Y$ .  
 ג. חשבו את  $E(X)$ ,  $V(X)$ .  
 ד. האם  $X$  ו- $Y$  הם משתנים בלתי תלויים?  
 ה. חשבו את מקדם המתאם בין  $X$  ל- $Y$ .  
 ו. חשבו את הסיכוי:  $P(Y > X + 10)$ .
- (5) משתנה מקרי דו ממדי מתפלג לפי פונקציית הצפיפות הבאה:  
 $f(x, y) = \lambda \mu \cdot e^{-(\lambda x + \mu y)}$ , המוגדרת בתחום שבו:  $x, y > 0$ .  
 א. מצאו את פונקציית הצפיפות של  $X$  ואת פונקציית הצפיפות של  $Y$ .  
 ב. האם  $X$  ו- $Y$  הם משתנים תלויים?  
 ג. מהו מקדם המתאם בין  $X$  ל- $Y$ ?  
 ד. חשבו את הסיכוי:  $P(Y > X)$ .

(6)  $Y$  הינו משתנה מקרי אחיד רציף המתפלג בקטע:  $[2, 4]$ .

בנוסף, נתון ש- $X$  הינו משתנה מקרי רציף המקיים:  $0 \leq x \leq y$ ,  $f(x|y) = \frac{2x}{y^2}$ .  
מצאו את השונות המשותפת של  $X$  ו- $Y$ .

(7) נתונים שני משתנים מקרים רציפים  $X$  ו- $Y$ . פונקציות הצפיפות

$$f(x, y) = \begin{cases} x & 0 < y < 1 \\ 0 & \text{else} \end{cases} \quad 1 - y \leq x \leq 1 + y$$

המשותפות שלהם היא:

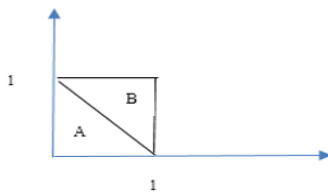
- מצאו את  $f(x)$ .
- מצאו את  $f(y|x)$ .
- מצאו את  $E(Y|X)$ .

(8) יהיו  $X$  ו- $Y$  משתנים רציפים המתפלגים אחיד בתוך משולש שקדקודיו:  $(-1, 2)$ ,  $(-1, 0)$ ,  $(0, 1)$ .

- רשמו את פונקציית הצפיפות המשותפת.
- מצאו את פונקציית הצפיפות השולית של  $X$  ו- $Y$ .
- חשבו את התוחלת של  $X$  ו- $Y$ .
- האם  $X$  ו- $Y$  משתנים בלתי מתואמים?
- האם  $X$  ו- $Y$  משתנים בלתי תלויים?

(9) פונקציית צפיפות משותפת מקבלת ערך אחיד באופן הבא:

הצפיפות על פני משולש  $A$  הינה 1.5 והצפיפות על פני משולש  $B$  היא 0.5.  
האם פונקציית הצפיפות המשותפת היא לגיטימית?



- מצאו את  $f(x)$ .
- מצאו את  $f(x|y)$ .

(10) נתונה פונקציית הצפיפות המשותפת:  $f(x, y) = cx$ . פונקציה זו מוגדרת

בתחום שבו:  $0 \leq x \leq 1$  וכן:  $0 \leq y \leq x^2$ .

- מצאו את הקבוע  $C$ .
- חשבו את ההסתברות ש- $6Y < 1 - X$ .

**(11)** נתונים  $X$  ו- $Y$  שני משתנים מקריים רציפים כך ש:  $Y \sim U(0,1)$

ו- $X|Y=y \sim U(0, \sqrt{y})$ . חשבו את:  $E(Y|X=0.5)$ .

**(12)** נתונה פונקציית הצפיפות:  $f(x, y) = 2e^{-x} \cdot e^{-2y}$  בתחום שבו:  $x, y \geq 0$ .

חשבו את הסיכוי:  $P(X < Y)$ .

**(13)** נתונה פונקציית הצפיפות המשותפת:  $f(x, y) = \frac{e^{-y} \cdot e^{-\frac{x}{y}}}{y}$ , המוגדרת לרביע

הראשון. חשבו את:  $P(X > 1|Y = 2)$ .

**(14)** יוסי וערן עובדים באותו המשרד. הם מגיעים לעבודה בכל יום בין 8:00 ל-9:00.

נניח שזמן ההגעה של כל אחד מתפלג אחיד ובאופן בלתי תלוי זה בזה.

מה הסיכוי שיוסי יצטרך לחכות לערן יותר מ-10 דקות?

**(15)** נתונים שני משתנים מקרים רציפים:  $X \sim N(Y, 1)$  ו- $Y \sim U(0, 2)$ .

א. מצאו את פונקציית הצפיפות המשותפת של  $X$  ו- $Y$ .

ב. מצאו את  $E(X^2|Y)$ .

ג. מצאו את  $E(X)$ .

**(16)** פונקציית הצפיפות המשותפת של  $X$  ו- $Y$  היא:  $f(x, y) = 1$ .

פונקציה זו מוגדרת בתחומי:  $0 \leq x, y \leq 1$ .

הוכיחו ש:  $E(|X - Y|^n) = \frac{2}{(n+1)(n+2)}$ .

**(17)**  $X \sim \exp(1)$  וכן:  $Y \sim \exp(1)$ , הינם משתנים מקרים בלתי תלויים.

נגדיר את:  $Z = \frac{X}{X+Y}$ .

הוכיחו:  $Z \sim U(0,1)$ .

## תשובות סופיות:

(1) שאלת הוכחה.

(2)  $A = \frac{1}{8}$ .

(3) א.  $\frac{5}{16}$  ב.  $f(y) = 0.8y^3 + 1.6y$  ג. תלויים.

(4) א. שאלת הוכחה. ב.  $f(y) = \frac{y-60}{800}$  ג.  $E(X) = 73\frac{1}{3}$ ,  $V(X) = 88\frac{8}{9}$

ד. לא. ה. 0.5 ו. 0.5625

(5) א.  $f(y) = \mu e^{-\mu y}$ ,  $f(x) = \lambda e^{-\lambda x}$  ב. לא.

ג. 0. ד.  $\frac{\lambda}{\lambda + \mu}$

(6)  $\frac{2}{9}$

(7) א.  $f(x) = \begin{cases} x^2 & 0 \leq x < 1 \\ 2x - x^2 & 1 \leq x \leq 2 \\ 0 & \text{else} \end{cases}$

ג.  $E(y|x) = \begin{cases} 1 - \frac{x}{2} & 0 \leq x < 1 \\ \frac{x}{2} & 1 \leq x \leq 2 \\ 0 & \text{else} \end{cases}$

ב.  $f(y|x) = \begin{cases} \frac{1}{x} & 0 \leq x < 1 \quad 1 - x < y < 1 \\ \frac{1}{2-x} & 1 \leq x \leq 2 \quad x-1 < y < 1 \\ 0 & \text{else} \end{cases}$

(8) א.  $f(x, y) = \begin{cases} 1 & 1+x < y < 1-x-1 < x < 0 \\ 0 & \text{else} \end{cases}$

ב.  $f(y) = \begin{cases} y & 0 \leq y < 1 \\ 2-y & 1 \leq y \leq 2 \\ 0 & \text{else} \end{cases}$ ,  $f(x) = \begin{cases} -2x & -1 < x < 0 \\ 0 & \text{else} \end{cases}$

ג.  $E(X) = -\frac{2}{3}$ ,  $E(Y) = 1$  ד. כן.

ה. לא.

$$f(x) = \begin{cases} 1.5-x & 1 < x < 0 \\ 0 & \text{else} \end{cases} \quad \text{ב.} \quad \text{9 א. כן.}$$

$$f(x|y) = \begin{cases} \frac{1.5}{1.5-y} & 0 \leq x < 1-y \\ \frac{0.5}{1.5-y} & 1-y \leq x \leq 1 \\ 0 & \text{else} \end{cases} \quad \text{ג.}$$

$$\text{10 א. 4. ב. } 0.0947$$

$$\frac{7}{12} \quad \text{11}$$

$$\frac{1}{3} \quad \text{12}$$

$$e^{-\frac{1}{2}} \quad \text{13}$$

$$\frac{25}{72} \quad \text{14}$$

$$f(x, y) = \begin{cases} \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2}} & 0 < y < 2 \\ 0 & \text{else} \end{cases} \quad \text{א. 15}$$

$$\text{ב. } y^2 + 1$$

$$\text{ג. 1.}$$

16 שאלת הוכחה.

17 שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 53 - קונבולוציה - התפלגות סכום משתנים בלתי תלויים

תוכן העניינים

1. קונבולוציה - התפלגות סכום משתנים בלתי תלויים ..... 238

## קונבולוציה – התפלגות סכום משתנים בלתי תלויים:

**רקע:**

יהיו  $X$  ו- $Y$  שני משתנים מקריים בלתי תלויים ונתעניין בהתפלגות סכומם:  
 $T = X + Y$  - שגם הוא משתנה מקרי.  
 אם מדובר במשתנים מקריים רציפים עם פונקציות צפיפות  $f_X$  ו- $f_Y$ , פונקציית הצפיפות של  $T = X + Y$ , תינתן על ידי נוסחת הקונבולוציה הבאה:

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(t-y) \cdot f_Y(y) dy$$

**דוגמה (פתרון בהקלטה):**

נתון:  $X \sim \exp(1)$  וכן:  $Y \sim \exp(2)$ . מצאו את פונקציית הצפיפות של:  $T = X + Y$ .

## שאלות:

- (1) נתון ש- $Y, X \sim \exp(\lambda)$ . כמו כן ידוע ש- $X$  ו- $Y$  בלתי תלויים. מצאו את פונקציית הצפיפות של  $X + Y$ .
- (2) נתון ש- $X$  ו- $Y$  משתנים בלתי תלויים המתפלגים נורמלית סטנדרטית. הוכיחו ש- $T = X + Y$  מתפלג נורמלית עם תוחלת 0 ושונות 2.
- (3) סוללה מסוג A בעלת אורך חיים המתפלג אחיד בין 1 ל-3 שעות. כמו כן נתונה סוללה מסוג B בעלת אורח חיים המתפלג מעריכית עם תוחלת חיים של שעה. מכשיר מופעל על ידי סוללה A וברגע שהסוללה מתרוקנת אוטומטית מופעלת סוללה B. נסמן ב- $Z$  את הזמן הכולל של פעילות המכשיר. א. מצאו את פונקציית הצפיפות של  $Z$ . ב. מה הסיכוי שהמכשיר יפעל פחות מ-4 שעות?
- (4)  $X$  ו- $Y$  משתנים מקריים רציפים ובלתי תלויים בעלי פונקציות הצפיפות הבאות:  $f_x(x) = \frac{1}{4}$   $-2 \leq x \leq 2$ ,  $f_y(y) = \begin{cases} y+1 & -1 \leq y \leq 0 \\ 1-y & 0 \leq y \leq 1 \\ 0 & \text{else} \end{cases}$ . מצאו את פונקציית הצפיפות של  $X + Y$ .
- (5) יהיו  $X$  ו- $Y$  משתנים מקריים רציפים ובלתי תלויים בעלי התפלגות אחידה:  $X \sim U(2,3)$   $Y \sim U(1,5)$ . א. מהי ההתפלגות של סכום המשתנים הללו? ב. מה הרבעון העליון של סכום המשתנים?
- (6) יהיו  $X, Y$  ו- $Z$  מתפלגים אחיד רציף באופן בלתי תלוי בין 0 ל-1. מצאו את פונקציית הצפיפות של:  $X + Y + Z$ .
- (7) הוכיחו את נוסחת הקונבולוציה עבור המקרה הרציף. (רמז: היעזרו בפונקציית הצפיפות המשותפת ובהגדרה של משתנים מקריים רציפים בלתי תלויים).

## תשובות סופיות:

$$f_T(t) = \lambda^2 \cdot e^{-\lambda t} \cdot t \quad t \geq 0 \quad (1)$$

(2) שאלת הוכחה.

$$f_z(z) = \begin{cases} \frac{1}{2}(1 - e^{1-z}) & 1 \leq z \leq 3 \\ \frac{1}{2}(e^{3-z} - e^{1-z}) & z > 3 \\ 0 & \text{else} \end{cases} \quad \text{א. ב. 0.841} \quad (3)$$

$$f_T(t) = \begin{cases} \frac{1}{8}(t+3)^2 & -3 \leq t \leq -2 \\ \frac{1}{8}(2 - (t+1)^2) & -2 < t < -1 \\ \frac{1}{4} & -1 \leq t \leq 1 \\ \frac{1}{8}(2 - (t-1)^2) & 1 < t < 2 \\ \frac{1}{8}(t-3)^2 & 2 \leq t \leq 3 \\ 0 & \text{else} \end{cases} \quad (4)$$

$$f_T(t) = \begin{cases} \frac{t-3}{4} & 3 \leq t \leq 4 \\ \frac{1}{4} & 4 < t < 7 \\ \frac{8-t}{4} & 7 \leq t \leq 8 \end{cases} \quad \text{א. ב. 4.5} \quad (5)$$

$$f_w(w) = \begin{cases} \frac{w^2}{2} & 0 \leq w \leq 1 \\ -w^2 + 3w - 1.5 & 1 < w < 2 \\ \frac{(3-w)^2}{2} & 2 \leq w \leq 3 \end{cases} \quad (6)$$

(7) שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 54 - חישוב תוחלת ושונות על ידי פירוק לאינדקטורים

תוכן העניינים

1. כללי ..... 241

## חישוב תוחלת ושונות על ידי פירוק לאינדיקטורים:

### רקע:

נלמד שיטה לחישוב תוחלת ושונות של משתנה מקרי, על ידי פירוק לסכום של משתני אינדיקטור. אינדיקטור הינו משתנה שפונקציית ההסתברות של נראית כך:

$X$	1	0
$P(X)$	$P$	$1-P$

נגיד ש- $X_i$  הינו משתנה אינדיקטור כאשר:  $i=1,2,\dots,n$  ו- $X = \sum_{i=1}^n X_i$ .

ניעזר בנוסחאות תוחלת ושונות סכום משתנים מקרים כדי לחשב את התוחלת והשונות של  $X$ .

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$$

$$V(X) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \cdot \sum_{i < j} COV(X_i, X_j)$$

כאשר עבור משתנים אינדיקטורים מתקיים ש:

$$E(X_i) = P(X_i = 1)$$

$$V(X_i) = P(X_i = 1) \cdot P(X_i = 0)$$

$$COV(X_i, X_j) = P(X_i = 1, X_j = 1) - P(X_i = 1)P(X_j = 1)$$

### דוגמא (פתרון בהקלטה):

יוסי החליט להזמין 8 חברים למסיבת יום הולדתו. הוא הכין 8 הזמנות שעליהן רשם את השם של כל אחד מהחברים. ההזמנות הוכנסו למעטפות וחולקו באקראי ל-8 החברים. נסמן ב- $X$  את מספר ההזמנות שהגיעו לחבר הנכון. חשבו את  $E(X)$  ואת  $V(X)$ .

## שאלות:

- (1) יהיו  $X$  ו- $Y$  משתני אינדיקטורים. הוכיחו ש:
- $E(X) = P(X=1)$ .
  - $V(X) = P(X=1) \cdot [1 - P(X=1)]$ .
  - $COV(X, Y) = P(X=1, Y=1) - P(X=1)P(Y=1)$ .
- (2) 400 אנשים נבחרו מכלל האוכלוסייה.
- חשבו את הסיכוי שביום מסוים בשנה יהיה בדיוק אדם אחד מתוך ה-400 שיש לו יום הולדת.
  - נגדיר את  $X_i$  משתנה אינדיקטור המקבל את הערך 1 אם ביום  $i$  בדיוק אדם אחד מתוך ה-400 עם יום הולדת באותו היום. חשבו את התוחלת והשונות של  $X_i$ .
  - חשבו את התוחלת והשונות של מספר הימים בשנה שבהם יש יום הולדת בדיוק לאחד מ-400 האנשים הללו.
- (3) 3 משחקים הוכנסו באקראי ל-5 מגרות. מגירה יכולה להכיל יותר ממשחק אחד. נסמן ב- $W$  את מספר המגרות בהן בדיוק משחק אחד. חשבו את התוחלת והשונות של  $W$  על ידי פירוק לאינדיקטורים.
- (4)  $A, B$  ו- $C$  הם שלושה מאורעות כך ש:  $P(A) = 0.3, P(B) = 0.2, P(C) = 0.1$ .
- נגדיר את  $Y$  להיות מספר המאורעות מתוך השלושה שמתקיימים. חשבו את התוחלת והשונות של  $Y$  כאשר:
- המאורעות בלתי תלויים זה בזה.
  - $C \subset B \subset A$ .
  - $A, B$  ו- $C$  זרים זה לזה.
- (5) נטיל קובייה 10 פעמים. נסמן ב- $W$  את מספר התוצאות השונות שהתקבלו.
- מצאו את  $E(W)$ .
  - מצאו את  $V(W)$ .

- (6) נסדר בשורה 6 כוסות קולה ו-4 כוסות מים. רצף של שתי כוסות נקרא "גינק", אם שתי הכוסות הן ברצף של קולה. נסמן ב- $X$  את מספר הרצפים מסוג "גינק" שיש לשתי כוסות. למשל, הסידור הבא:  
 קולה, מים, קולה, מים, קולה, מים, קולה, מים, קולה, קולה, קולה,  $X = 2$ .  
 חשבו את התוחלת והשונות של  $X$ .
- (7) נסדר בשורה  $n$  זוגות גרביים באקראי (בסך הכול  $2n$  גרביים).  
 חשבו את התוחלת והשונות של מספר הזוגות מתוך  $n$  הזוגות שבהם זוג הגרביים אינם עומדים זה לצד זה.
- (8) בקייטנה 100 ילדים. מחלקים לכל ילד 2 ארטיקים מתוך 200 הארטיקים שנרכשו לקייטנה. מתוך 200 הארטיקים שנרכשו 100 בטעם תות ו-100 הם בטעם לימון. נסמן ב- $X$  את מספר הילדים שקיבלו 2 ארטיקים בטעמים שונים. נסמן ב- $Y$  את מספר הילדים שקיבלו שני ארטיקים בטעם לימון.  
 א. חשבו את התוחלת והשונות של  $X$ .  
 ב. בטאו את  $Y$  כפונקציה של  $X$  וחשבו את התוחלת והשונות של  $Y$ .  
 ג. מהי השונות המשותפת של  $X$  ו- $Y$ ?

### תשובות סופיות:

- (1) שאלת הוכחה.
- (2) א. 0.3667. ב. תוחלת: 0.3667, שונות: 0.2322.  
 ג. תוחלת: 133.85, שונות: 88.89.
- (3) תוחלת: 1.92, שונות: 1.1136.
- (4) א. תוחלת: 0.6, שונות: 0.46. ב. תוחלת: 0.6, שונות: 1.04.  
 ג. תוחלת: 0.6, שונות: 0.24.
- (5) א. 5.03. ב. 0.568.
- (6) תוחלת: 3, שונות:  $\frac{2}{3}$ .
- (7) תוחלת:  $n-1$ , שונות:  $1 - \frac{1}{n} + \frac{n-1}{2(2n-1)}$ .
- (8) א. תוחלת: 50.251, שונות: 25.126.  
 ב.  $Y = -0.5X + 50$ . ג. -12.563.

# סטטיסטיקה והסתברות 01 88165

פרק 55 - נוסחת התוחלת השלמה

תוכן העניינים

1. כללי ..... 244

## נוסחת התוחלת השלמה:

רקע:

כאשר התפלגות של משתנה  $X$  תלויה במשתנה אחר  $Y$ , מתקיים:  $E(X) = E[E(X/Y)]$ .

עבור משתנה  $Y$  בדיד כלשהו:  $E(X) = \sum_y E(X | Y = y) \cdot P(Y = y)$ .

עבור משתנה  $Y$  רציף כלשהו:  $E(X) = \int_{-\infty}^{\infty} E(X | Y = y) \cdot f(y) dy$ .

**דוגמה (הפתרון בהקלטה):**

מרצה מלמד שתי כיתות. הוא מעוניין לבדוק 5 מחברות בחינה.

בכיתה מספר 1: 10 בנים ו-20 בנות.

בכיתה מספר 2: 15 בנים ו-15 בנות.

המרצה בוחר כיתה באופן מקרי וממנה בוחר 5 מחברות בחינה באקראי.

יהי  $X$  מספר מחברות הבחינה של בנים שנבחרו.

יש לחשב את  $E(X)$ .

## שאלות:

(1) בכד 2 כדורים ירוקים, 4 כדורים כחולים ו-4 אדומים. שולפים כדור באקראי. אם הכדור ירוק מטיילים קובייה עד אשר מתקבלת התוצאה 1. אם הכדור כחול מטיילים קובייה עד אשר מתקבלת התוצאה 5. אם הכדור אדום מטיילים קובייה עד אשר מתקבלת תוצאה זוגית. חשבו את תוחלת מספר הפעמים שהקובייה הוטלה.

(2) מטיילים  $n$  מטבעות ומוציאים מהמשחק את כל המטבעות שהראו ראש. כעת מטיילים את כל המטבעות שנותרו. בטאו באמצעות  $n$  את תוחלת מספר הראשים שהתקבלו בסבב השני של ההטלות.

(3) בהגרלה מבצעים את התהליך הבא: בסיכוי 0.25 מגרילים מספר ממכונה A בסיכוי 0.75 מגרילים מספר ממכונה B. במכונה A המספר המוגרל מתפלג פואסונית עם תוחלת 6. במכונה B המספר המוגרל מתפלג פואסונית עם תוחלת 2. אם הוגרל המספר 0 זוכים ב-15₪. אחרת זוכים ב-50₪. חשבו את תוחלת סכום הזכיה.

(4) נתון ש- $Y/X \sim U(0, X)$ , כאשר פונקציית הצפיפות של  $X$

$$f(x) = \begin{cases} 0.1 & 2 \leq x \leq 6 \\ 0.2 & 6 < x \leq 9 \\ 0 & \text{אחרת} \end{cases} \quad \text{הינה:}$$

חשבו את  $E(Y)$ .

(5) בתכנית הריאליטי "המרוץ למיליון" מגיעים לנקודה שבה שלוש אפשרויות בפני המתמודדים: נקודה A - שבה חוזרים אחרי 1 שעות לנקודת המוצא. נקודה B - שבה חוזרים אחרי 2 שעות לנקודת המוצא ונקודה C - המובילה תוך 2 שעות לנקודת הסיום. המתמודדים בוחרים בכל פעם את הנקודה באופן מקרי. נסמן ב- $X$  את זמן ההגעה לנקודת הסיום. חשבו את  $E(Y)$ .

**תשובות סופיות:**

(1) 4.4

(2)  $\frac{n}{4}$ 

(3) 46.4

(4) 3.05

(5) .5

# סטטיסטיקה והסתברות 01 88165

פרק 56 - נוסחת השונות השלמה ( שונות של משתנה מותנה )

תוכן העניינים

1. נוסחת השונות השלמה ( שונות של משתנה מותנה) ..... 247

## נוסחת השונות השלמה (המותנית):

**רקע:**

כאשר התפלגות של משתנה  $X$  תלויה במשתנה אחר  $Y$ , מתקיים:  $E(X) = E[E(X|Y)]$ .

כמו כן, מתקיים לגבי השונות:  $\text{Var}(X) = \text{Var}(E[X|Y]) + E[\text{Var}(X|Y)]$ .

**דוגמה (הפתרון בהקלטה):**

מרצה מלמד שתי כיתות. הוא מעוניין לבדוק 5 מחברות בחינה.

בכיתה מספר 1: 10 בנים ו-20 בנות.

בכיתה מספר 2: 15 בנים ו-15 בנות.

המרצה בוחר כיתה באופן מקרי וממנה בוחר 5 מחברות בחינה באקראי.

יהי  $X$  מספר מחברות הבחינה של בנים שנבחרו.

יש לחשב את  $V(X)$ .

## שאלות:

- (1) בכד 2 כדורים ירוקים, 4 כדורים כחולים ו-4 אדומים. שולפים כדור באקראי. אם הכדור ירוק מטיילים קובייה עד אשר מתקבלת התוצאה 1. אם הכדור כחול מטיילים קובייה עד אשר מתקבלת התוצאה 5. אם הכדור אדום מטיילים קובייה עד אשר מתקבלת תוצאה זוגית. חשבו את התוחלת והשונות של מספר הפעמים שהקובייה הוטלה.

- (2) נטיל קובייה:  $Y+4$  פעמים. נתון ש- $Y \sim P(4)$ .  
 נגדיר את  $X$  כמספר הפעמים שהתקבלה התוצאה 2.  
 א. מצאו את התוחלת של  $X$ .  
 ב. מצאו את השונות של  $X$ .

- (3) נתון ש- $Y|X \sim U(0, X)$ , כאשר פונקציית הצפיפות של  $X$

$$f(x) = \begin{cases} 0.1 & 2 \leq x \leq 6 \\ 0.2 & 6 < x \leq 9 \\ 0 & \text{אחרת} \end{cases} \quad \text{הינה:}$$

חשבו את  $V(Y)$ .

## תשובות סופיות:

- (1) תוחלת: 4.4, שונות: 22.64.  
 (2) תוחלת:  $\frac{4}{3}$ , שונות:  $\frac{11}{9}$ .  
 (3) 4.4.

# סטטיסטיקה והסתברות 01 88165

פרק 57 - הסקה סטטיסטית - הקדמה

תוכן העניינים

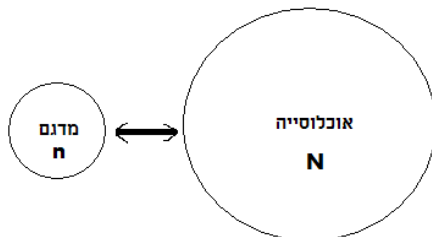
1. כללי ..... 249

## הסקה סטטיסטית – הקדמה:

### רקע:

#### אוכלוסייה:

קבוצה שאליה מפנים שאלה מחקרית. למשל, חברת תרופות שמעוניינת לפתח תרופה למחלת הסוכרת מתעניינת באוכלוסיית חולי הסוכרת בעולם.



#### מדגם:

חלק מתוך האוכלוסייה. למשל, אם נדגום באקראי 10 אנשים מתוך חולי הסוכרת אז זהו מדגם מתוך אוכלוסיית חולי הסוכרת.

במקרים רבים אין אפשרות לחקור את כל האוכלוסייה כיוון שאין גישה לכולה, היא גדולה מידי, או מוגבלים בזמן ובאמצעים טכניים ולכן מבצעים מדגם במטרה לבצע הסקה סטטיסטית מהמדגם לאוכלוסייה. הדגימה בקורס תהיה דגימה מקרית - הכוונה לדגימה שבה לכל תצפית באוכלוסייה יש את אותו סיכוי להיכלל במדגם.

#### סטטיסטי:

גודל המחושב על המדגם.

#### פרמטר:

גודל המתאר את האוכלוסייה.

**הסימונים לפרמטר וסטטיסטי הם שונים:**

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

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

**דוגמה (פתרון בהקלטה):**

25% מאזרחי המדינה תומכים בהצעת החוק של חבר כנסת 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 שמתכננים להוציא מהמדגם?

ו. האם הפרמטר או הסטטיסטי הוא משתנה מקרי?

## שאלות:

- (1) מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מי האוכלוסייה?
  - מה המשתנה?
  - מהם הפרמטרים?
  - מהו גודל המדגם?
- (2) להלן התפלגות מספר מקלטי הטלוויזיה למשפחה בישוב "העוגן". נגדיר את  $X$  להיות מספר המקלטים של משפחה אקראית. מתכננים לדגום מאוכלוסייה זו 4 משפחות ולהתבונן בממוצע מספר מקלטי הטלוויזיה במדגם.
- מיהי האוכלוסייה ומהו המשתנה הנחקר?
  - מהו הסטטיסטי שיילקח מהמדגם ומה סימונו?

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

- (3) נתון כי 20% מהשכירים במדינה הם אקדמאיים. נבחרו באקראי 10 שכירים באותה אוכלוסייה ומתכננים לפרסם את מספר האקדמאיים שנדגמו.
- מיהי האוכלוסייה?
  - מה המשתנה באוכלוסייה?
  - מהם הפרמטרים?
  - מהו הסטטיסטי?

## תשובות סופיות:

- (1) א. כלל הסטודנטים במכללה שסיימו סטטיסטיקה א. ב. ציון. ג. ממוצע: 78, סטיית תקן: 15. ד. 2.
- (2) א. האוכלוסייה: 1000 משפחות בישוב העוגן, המשתנה הנחקר: מס' מקלטים. ב.  $\bar{X}$  = ממוצע מדגם.
- (3) א. השכירים במדינה. ב. השכלה: אקדמאי, לא אקדמאי. ג. שיעור ההצלחות באוכלוסייה: 0.2. ג. מס' האקדמאים במדגם.

# סטטיסטיקה והסתברות 01 88165

פרק 58 - התפלגות הדגימה ומשפט הגבול המרכזי

תוכן העניינים

1. התפלגות ממוצע המדגם ומשפט הגבול המרכזי. 252
2. התפלגות סכום תצפיות בלתי תלויות ומשפט הגבול המרכזי. 260
3. התפלגות מספר ההצלחות במדגם - קירוב נורמלי להתפלגות הבינומית. 263
4. התפלגות פרופורציית ההצלחות במדגם. 268
5. חוק המספרים הגדולים. 272

## התפלגות ממוצע המדגם ומשפט הגבול המרכזי:

### רקע:

בפרק זה נדון בהתפלגות של ממוצע המדגם:  $\bar{x} = \frac{\sum x_i}{n}$

מכיוון שממדגם למדגם אנו יכולים לקבל ממוצע מדגם שונה, אזי ממוצע המדגם הוא משתנה מקרי ויש לו התפלגות.

גדלים המתארים התפלגות כלשהי או אוכלוסייה כלשהי נקראים פרמטרים.

להלן רשימה 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 התפלגות:

ממוצע כל ממוצעי המדגם האפשריים שווה לממוצע האוכלוסייה:  $E(\bar{x}) = \mu_x = \mu$ .

שונות כל ממוצעי המדגם האפשריים שווה לשונות האוכלוסייה מחולק ב- $n$ .

תכונה זו נכונה רק במדגם מקרי:  $V(\bar{x}) = \sigma_x^2 = \frac{\sigma^2}{n}$ .

יש יחס הפוך בין גודל המדגם לבין שונות ממוצעי המדגם.

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

טעות תקן:  $\sigma(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$ .

### דוגמה (פתרון בהקלטה):

השכר הממוצע במשק הינו 9000 ₪ עם סטיית תקן של 4000. דגמו באקראי 25 עובדים.

א. מיהי אוכלוסיית המחקר? מהו המשתנה הנחקר?

ב. מהם הפרמטרים של האוכלוסייה?

ג. מה התוחלת ומהי סטיית התקן של ממוצע המדגם?

**דגימה מהתפלגות נורמאלית:**

אם נדגום מתוך אוכלוסייה שהמשתנה בה מתפלג נורמאלית עם ממוצע  $\mu$  ושוונות  $\sigma^2$ .

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad Z_{\bar{x}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

ממוצע המדגם גם יתפלג נורמאלית:

**דוגמה (פתרון בהקלטה):**

משקל תינוק ביום היוולדו מתפלג נורמאלית 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  $\mu$  ושוונות  $\sigma^2$  אזי עבור מדגם מספיק

$$\bar{x} \rightsquigarrow N\left(\mu, \frac{\sigma^2}{n}\right) \quad (n \geq 30)$$

ממוצע המדגם מתפלג בקירוב נורמאלי:

**דוגמה (פתרון בהקלטה):**

משקל חפיסת שוקולד בקו ייצור מתפלג עם ממוצע 100 גרם וסטיית תקן של 4 גרם. דגמו מקו הייצור 36 חפיסות שוקולד אקראיות. מה ההסתברות שהמשקל הממוצע של חפיסות השוקולד שנדגמו יהיה מתחת ל-102 גרם?

## שאלות:

- (1) מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מיהי האוכלוסייה?
  - מה המשתנה?
  - מהם הפרמטרים?
  - מהו גודל המדגם?
  - מהו תוחלת ממוצע המדגם?
  - מהי טעות התקן?

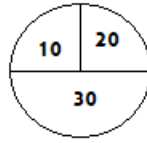
- (2) להלן התפלגות מספר מקלטי הטלויזיה למשפחה בישוב מסוים:

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

- בנו את פונקציית ההסתברות של  $X$ .
  - חשבו את התוחלת, השונות וסטיית התקן של  $X$ .
  - אם נדגום 4 משפחות מהישוב עם החזרה מה תהיה התוחלת, מהי השונות ומהי סטיית התקן של ממוצע המדגם?
- (3) אם נטיל קובייה פעמיים ונתבונן בממוצע התוצאות שיתקבלו, מה תהיה התוחלת ומה תהיה סטיית התקן של ממוצע זה?
- (4) משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.
- מה ההסתברות שתינוק אקראי בעת הלידה ישקול פחות מ-3800 גרם? נתון כי ביום מסוים נולדו 4 תינוקות.
  - מה ההסתברות שהמשקל הממוצע שלהם יעלה על 4 ק"ג?
  - מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-2.5 ק"ג?
  - מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת בלא יותר מ-50 גרם?
  - הסבירו ללא חישוב כיצד התשובה לסעיף הקודם הייתה משתנה אם היה מדובר על יותר מ-4 תינוקות?

- (5) הגובה של המתגייסים לצה"ל מתפלג נורמאלית עם תוחלת של 175 ס"מ וסטיית תקן של 10 ס"מ. ביום מסוים התגייסו 16 חיילים.
- מה ההסתברות שהגובה הממוצע שלהם יהיה לפחות 190 ס"מ?
  - מה ההסתברות שהגובה הממוצע שלהם יהיה בדיוק 180 ס"מ?
  - מה ההסתברות שהגובה הממוצע שלהם יסטה מתוחלת הגבהים בפחות מ-5 ס"מ?
  - מהו הגובה שבהסתברות של 90% הגובה הממוצע של המדגם 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 הפתרון הניחו שזמן הנסיעה לעבודה מתפלג נורמאלית.
- מה ההסתברות שבמשך שבוע משך הנסיעה הממוצע יהיה מעל 33 דקות?
  - מהו הזמן שבהסתברות של 90% ממוצע משך הנסיעה השבועי יהיה גבוה ממנו?
  - מה ההסתברות שממוצע משך הנסיעה השבועי יהיה מרוחק מ-30 דקות בלפחות 2 דקות?
  - כיצד התשובה לסעיף הקודם הייתה משתנה אם האדם היה נוסע לעבודה 6 פעמים בשבוע?
- (7) נפח היין בבקבוק מתפלג נורמאלית עם תוחלת של 750 סמ"ק וסטיית תקן של 10 סמ"ק.
- בארגו 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגו יהיה בדיוק 755 סמ"ק?
  - בארגו 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגו יהיה יותר מ-755 סמ"ק?
  - בארגו 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגו יהיה לפחות 755 סמ"ק?
  - בקבוקי היין שבארגו נמזגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?
- (8) משתנה מתפלג נורמאלית עם תוחלת של 80 וסטיית תקן 4.
- מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?
  - מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה שגודל המדגם הוא 16?
  - הסבר את ההבדל בתשובות של שני הסעיפים.

9) בקזינו ישנה רולטה. על הרולטה רשומים המס' הבאים כמוראה בשרטוט:



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.

- בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.
- מה התוחלת ומה השונות של סכום הזכייה?
- אם האדם ישחק את המשחק 5 פעמים מה התוחלת ומה השונות של ממוצע סכום הזכייה בחמשת המשחקים?
- אם האדם משחק את המשחק 50 פעם מה ההסתברות שבסה"כ יזכה ב-1050 ₪ ומעלה?

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

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

- 12) אורך צינור שמפעל מייצר הינו עם ממוצע של 70 ס"מ וסטיית תקן של 10 ס"מ.
- נלקחו באקראי 100 מוטות, מה ההסתברות שממוצע אורך המוטות יהיה בין 68 ל 78 ס"מ?
  - יש לחבר 2 בניינים באמצעות מוטות. המרחק בין שני הבניינים הינו 7200 ס"מ. מה ההסתברות ש-100 המוטות יספיקו למלאכה?
  - מה צריך להיות גודל המדגם המינימאלי, כדי שבהסתברות של 5% ממוצע המדגם יהיה קטן מ-69 ס"מ. היעזרו במשפט הגבול המרכזי.

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 שממוצע המדגם יהיה קטן מ-5?

14 נתון ש- $X \sim N(\mu, \sigma^2)$ . דגמו 5 תצפיות מאותה התפלגות והתבוננו במוצע

המדגם  $\bar{X}$ . לכן:  $P(\bar{X} > \mu)$  יהיה (בחרו בתשובה הנכונה):

- 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 נדגום  $n$  תצפיות מתוך ההתפלגות ונגדיר:  $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ ,

אזי (בחרו בתשובה הנכונה):

- א.  $\mu$  ו- $\bar{X}$  יהיו משתנים מקריים.
- 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.

17 משקל חפיסת שוקולד בקו ייצור מתפלג עם ממוצע 100 גרם. החפיסות נארזות בקרטון המכיל 36 חפיסות שוקולד אקראיות. ההסתברות שהמשקל הממוצע של חפיסות השוקולד בקרטון יהיה מעל 99 גרם הוא 0.9932.

- א. מהי סטיית התקן של משקל חפיסת שוקולד בודדת?
- ב. מה הסיכוי שמתוך 4 קרטונים בדיוק קרטון אחד יהיה עם משקל ממוצע לחפיסה הנמוך מ-100 גרם?

**18** משתנה מקרי כלשהו מתפלג עם סטיית תקן של 20. מה הסיכוי שאם נדגום 100 תצפיות בלתי תלויות מאותה התפלגות אזי ממוצע המדגם יסטה מתוחלתו בפחות מ-2?

**19** מספר המכוניות הנכנסות לחניון "בציר" במשך היום מתפלג פואסונית עם קצב של מכונית אחת לדקה. שומר מסר נתונים על מספר המכוניות שנכנסות בכל שעה לגבי 40 שעות שאסף נתונים. מה ההסתברות שממוצע מספר המכוניות שנכנסו לחניון לשעה בשעות אלה יהיה לפחות 63?

**20** הוכיחו שאם משתנה מתפלג כלשהו עם תוחלת  $\mu$  ושונות  $\sigma^2$ , ומבצעים מדגם בגודל  $n$  של תצפיות בלתי תלויות מהמשתנה, אזי מתקיימות התכונות הבאות

$$E(\bar{x}) = \mu \text{ ו- } V(\bar{x}) = \frac{\sigma^2}{n}.$$

## תשובות סופיות:

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

(2) א. להלן טבלה:

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

- ב.  $\sigma = 0.973$ ,  $\sigma^2 = 0.9475$ ,  $\mu = 2.05$   
 ג.  $\sigma(\bar{X}) = 0.486$ ,  $\sigma_{\bar{x}}^2 = 0.2369$ ,  $\mu_{\bar{x}} = 2.05$

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

- (4) א. 0.8413      ב. 0.0013  
 (5) א. 0      ב. 0  
 (6) א. 0.0465      ב. 27.71  
 (7) א. 0      ב. 0.1587  
 (8) א. 0.5468      ב. 0.6826  
 (9) א. להלן טבלה:

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

- ב. התוחלת: 22.5, השונות: 68.75.  
 ג. התוחלת: 22.5, השונות: 13.75.  
 ד. 0.8997  
 (10) 0.0475  
 (11) 0.1814  
 (12) א. 0.9772      ב. 0.0228      ג. 271  
 (13) 0.5  
 (14) ב'  
 (15) ד'  
 (16) ג'  
 (17) א. 2.429      ב. 0.25  
 (18) 0.6826  
 (19) 0.0071  
 (20) שאלת הוכחה.

## התפלגות סכום תצפיות בלתי תלויות ומשפט הגבול המרכזי:

**רקע:**

כעת נדון בסטטיסטי המבטא את סכום התצפיות במדגם:  $T = \sum_{i=1}^n X_i$ .  
 כאשר כל התצפיות נדגמו באקראי מאותה אוכלוסייה, כלומר, היו:  $X_1, \dots, X_n$  -  
 משתנים מקריים בלתי תלויים בעלי התפלגות זהה שתוחלתה  $\mu$  ושונותה  $\sigma^2$  אזי:  
 התוחלת והשונות של סכום התצפיות:  $E(T) = n\mu$ ,  $V(T) = n\sigma^2$ .

**דגימה מתוך התפלגות נורמלית:**

אם:  $X \sim N(\mu, \sigma^2)$ , אזי:  $Z = \frac{T - n\mu}{\sqrt{n\sigma^2}}$ ,  $T \sim N(n\mu, n\sigma^2)$

**משפט הגבול המרכזי:**

אם  $X$  מתפלג כלשהו וידוע כי:  $E(X) = \mu$ ,  $V(X) = \sigma^2$ , אזי עבור מדגם מספיק גדול (לפחות 30):  $T \rightsquigarrow N(n\mu, n\sigma^2)$

**דוגמה (פתרון בהקלטה):**

- בעיר מסוימת המשכורת הממוצעת של עובד הינה 8000 ₪. עם סטיית תקן של 2000 ₪. נדגמו 100 עובדים מהעיר שמפקידים את משכורתיהם לסניף בנק.
- מה התוחלת וסטיית התקן של סך המשכורות שיופקדו לסניף הבנק על ידי העובדים הללו?
  - מה ההסתברות שלסניף יופקד פחות מ-780 אלף ₪ ע"י אותם עובדים?

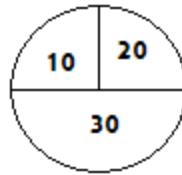
**שאלות:**

- (1) המשקל באוכלוסייה מסוימת מתפלג נורמאלית עם תוחלת של 60 ק"ג וסטיית תקן של 10 ק"ג.
- א. מה הסיכוי שאדם אקראי מהאוכלוסייה ישקול מתחת ל-65 ק"ג?  
 ב. מה הסיכוי שהמשקל הממוצע של 4 אנשים אקראיים יהיה מתחת ל-65 ק"ג?  
 ג. מה הסיכוי שהמשקל הכולל של 4 אנשים אקראיים יהיה מתחת ל-240 ק"ג?
- (2) נפח יין בבקבוק מתפלג נורמאלית עם תוחלת של 750 מ"ל וסטיית תקן של 20 מ"ל. אדם קנה מארז של 4 בקבוקי יין.
- א. מהי התוחלת ומהי סטיית התקן של נפח היין במארז?  
 ב. את היין שבמארז האדם מזג לכלי שקיבולתו 3.1 ליטר.  
 מה ההסתברות שהיין יגלוש מהכלי?  
 ג. אם לא היה נתון שנפח היין מתפלג נורמאלית. האם התשובה לסעיף א' הייתה משתנה? האם התשובה לסעיף ב' הייתה משתנה?
- (3) בספר כלשהו 500 עמודים. קצב הקריאה הממוצע הוא עמוד אחד ב-4 דקות עם סטיית תקן של 1 דקות.
- א. מה ההסתברות לסיים את הפרק הראשון (40 עמודים) תוך שעתיים וחצי?  
 ב. מהו האחוזון ה-95 לזמן סיום קריאת הספר?
- (4) במגדל נבנו 40 יחידות דיור. כמו כן נבנו 135 מקומות חנייה לבניין. להלן פונקציית ההסתברות של מספר המכוניות ליחידת דיור:

$x$	1	2	3	4	5
$P(X = x)$	0.1	0.2	0.3	0.25	0.15

- נניח שמספר המכוניות ליחידת דיור בלתי תליות זו בזו ועם אותה פונקציית ההסתברות לכל יחידת דיור (אין צורך בתיקון רציפות).
- א. מהי ההסתברות שיהיה מקום בחניון המגדל לכל מכוניות הבניין?  
 ב. בהינתן ויש מקום במגדל לכל המכוניות, מה הסיכוי שבפועל מספר המכוניות נמוך מ-130?

5) בקזינו ישנה רולטה עליה מסומנים המספרים הבאים:



אדם מסובב את הרולטה וזוכה בסכום הרשום.

א. אם האדם משחק 50 פעמים, מה ההסתברות שבסך הכול יזכה בסכום של 1050 ₪ ומעלה?

ב. האדם מגיע בכל יום לקזינו ומשחק 50 פעם, עד אשר מגיע היום בו הוא יזכה ב-1050 ₪ ומעלה. מה התוחלת ומהי השונות של מספר הימים שיבלה בקזינו?

6) נתון ש- $X_i \sim \exp(\lambda=1)$ , כאשר:  $i = 1, 2, \dots, 100$ .

$$\text{חשבו את הסיכוי: } P\left(\sum_i X_i \geq 115\right)$$

7) אורך חיי סוללה (בשעות) הוא בעל פונקציית הצפיפות הבאה:  $0 < x < 1$ ,  $f(x) = 2x$ . ברגע שסוללה מתרוקנת מחליפים אותה במיידית בסוללה אחרת. כמה סוללות יש להחזיק במלאי אם רוצים שבסיכוי של 90% לפחות המלאי יספיק עבור 35 שעות לפחות?

### תשובות סופיות:

1) א. 0.6915      ב. 0.8413      ג. 0.5

2) א. תוחלת: 3000 מ"ל, סטיית תקן: 40 מ"ל.      ב. 0.0062.

ג. סעיף א' - לא משתנה, סעיף ב' - לא פתיר, התבסס על התפלגות נורמלית.

3) א. 0.0571      ב. 2036.8

4) א. 0.883      ב. 0.7949

5) א. 0.8997      ב. תוחלת 1.111, שונות 0.1239.

6) 0.0668

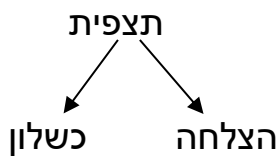
7) 56

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

רקע:

תזכורת על התפלגות בינומית:

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



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

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

תוחלת:  $E(y) = np$ .

שונות:  $V(y) = npq$ .

קירוב נורמלי עבור התפלגות בינומית:

אם לפנינו התפלגות בינומית:  $Y \sim B(n, p)$ , ומתקיים ש:

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

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

$$y \rightsquigarrow N(np, npq)$$

$$\cdot Z_y = \frac{y - np}{\sqrt{npq}} \quad \text{אז}$$

### תיקון רציפות:

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

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

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

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

### הערות:

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

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

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

- ישנם מרצים שנותנים את התנאי המחמיר הבא:

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

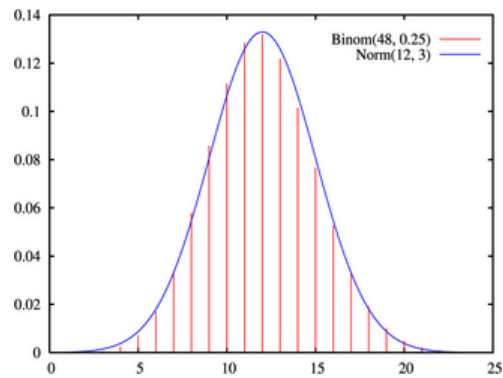
$$2. \quad n \cdot (1 - p) \geq 10$$

- וישנם מרצים שהתנאי שהם נותנים הוא:  $(n \geq 30)$ .
- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור מהתפלגות בינומית לנורמלית.
- הערה נוספת היא לגבי תיקון רציפות. ישנם מרצים שלא מחייבים לבצע תיקון רציפות שהמדגמים גדולים (בדרך כלל מעל 100 תצפיות) בפתרונות שאציג תמיד אבצע תיקון רציפות במעבר מבינומי לנורמלי כיוון שכך הפתרון יהיה יותר מדויק (בכל מקרה שהמדגמים גדולים העניין זניח).

**דוגמה (הפתרון בהקלטה):**

נתון שבקרב אוכלוסיית הנוער 25% זקוקים למשקפיים. נדגמו באקראי 48 בני נוער.

1. מה הסיכוי שבדיוק 14 מתוכם יהיו זקוקים למשקפיים?
2. מה הסיכוי שלכל היותר 13 מתוכם זקוקים למשקפיים?



## שאלות:

- (1) נתון ש-20% מאוכלוסייה מסוימת אקדמאית. נבחרו באקראי 10 אנשים באותה אוכלוסייה.
- א. מה ההסתברות ששלושה מהם אקדמאים?  
 ב. מה ההסתברות שלכל היותר אחד מהם אקדמאי?  
 ג. מה התוחלת ומהי סטיית התקן של מספר האקדמאים במדגם?
- (2) במפעל 10% מהמוצרים פגומים. נלקחו 100 מוצרים באקראי מקו הייצור.
- א. מה ההסתברות שנדגמו לפחות 6 מוצרים פגומים?  
 ב. מה ההסתברות שמספר המוצרים הפגומים יהיה לכל היותר 11 במדגם?
- (3) ציוני פסיכומטרי בקרב הנרשמים למוסד מסוים מתפלגים נורמאלית עם ממוצע 500 וסטיית תקן 100. למוסד מסוים הוחלט לקבל אך ורק סטודנטים שקיבלו מעל 600 בפסיכומטרי. 100 סטודנטים אקראיים נרשמו למוסד. מה ההסתברות שלפחות 20 יתקבלו?
- (4) מטילים מטבע 50 פעמים.
- א. מה ההסתברות לקבל לכל היותר 30 עצים?  
 ב. מה ההסתברות לקבל 28 עצים לפי התפלגות הבינומית ולפי הקירוב הנורמאלי?
- (5) במטוס מקום ל-400 נוסעים. נרשמו לטיסה 430 אנשים (overbooking). מנתונים סטטיסטיים ידוע שהסיכוי שאדם שנרשם לטיסה אךן יגיע הוא 0.9.
- א. מה ההסתברות שלא יהיו מקומות ישיבה לכל האנשים שהגיעו לטיסה?  
 ב. מה צריך להיות גודל המטוס כדי שבסיכוי שלפחות 95% המטוס יספיק לכמות הנרשמים?
- (6) מפעל לייצור ארטיקים טוען שהסיכוי שארטיק שהוא מייצר יהיה פגום הוא 0.01. מוכר הזמין 1000 ארטיקים מהמפעל. מה ההסתברות שהמוכר יקבל 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 הסיכוי שהרווח של המהמר יהיה לכל היותר 10?

### תשובות סופיות:

- |   |            |            |     |
|---|------------|------------|-----|
| ג. התוחלת: 2, סטיית התקן: 1.2649.             | ב. 0.3758. | א. 0.201.  | (1) |
|   | ב. 0.6915. | א. 0.9332. | (2) |
|   |            | 0.1611.    | (3) |
| ב. בינומית - 0.0788, קירוב לנורמלית - 0.0778. | ב. 0.9406. | א. 0.015.  | (4) |
|   | ב. 0.398.  | א. 0.015.  | (5) |
|   |            | 0.9996.    | (6) |
|   |            | 0.8643.    | (7) |

## התפלגות פרופורציית ההצלחות במדגם:

### רקע:

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

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

למשל, שיעור המובטלים במדגם -  $n = 200$  :

מספר המובטלים :  $Y = 20$  .

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

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

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

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

משפט הגבול המרכזי עבור הפרופורציה המדגמית:

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

### הערות:

- התנאים לקרוב הנורמאלי הם נזילים, כלומר משתנים ממרצה אחד לשני.  
 התנאי שהצגתי כאן הוא הפופולרי ביותר:
  1.  $n \cdot p \geq 5$
  2.  $n \cdot (1 - p) \geq 5$
- ישנם מרצים שנותנים את התנאי המחמיר הבא:
  1.  $n \cdot p \geq 10$
  2.  $n \cdot (1 - p) \geq 10$
- וישנם מרצים המשתמשים בתנאי:  $(n \geq 30)$  .
- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור לנורמלית.

- כיוון שפרופורציה אינה חייבת להיות מספר שלם בהכרח לא נהוג לבצע כאן תיקון רציפות.

**דוגמה (פתרון בהקלטה):**

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

- א. מה ההסתברות שהשכיחות היחסית ( $\hat{p}$ ) של הזכאים לבגרות במדגם תעלה על 60%?
- ב. מה ההסתברות שפרופורציית הזכאים לבגרות במדגם תעלה על 70%?

## שאלות:

- (1) במדינה מסוימת 10% מכלל האוכלוסייה הינם מובטלים. נדגמו באקראי 140 אנשים מהמדינה.
- מה התוחלת ומהי השונות של פרופורציות המובטלים שנדגמו?
  - מה ההסתברות שבמדגם לפחות 10% יהיו מובטלים?
  - מה ההסתברות שלכל היותר 9% מהמדגם יהיו מובטלים?
- (2) נניח כי 30% מהאוכלוסייה תומכים בהצעת חוק מסוימת. אם נדגום מהאוכלוסייה 200 איש. חשבו את ההסתברויות הבאות:
- לפחות 35% יתמכו בהצעת החוק במדגם.
  - לכל היות 25% יתמכו בהצעת החוק במדגם.
  - יותר מ-27% יתמכו בהצעת החוק במדגם.
- (3) לפי נתוני משרד התקשורת 40% מהאוכלוסייה מחזיקים בטלפון נייד מסוג "סמארטפון". נדגמו 400 אנשים מהאוכלוסייה.
- מה ההסתברות שבמדגם לכל היותר ל-40% יש סמארטפון?
  - מה ההסתברות שבמדגם לרוב יש סמארטפון?
  - מה ההסתברות שפרופורציית בעלי הסמארטפון במדגם תסטה מהפרופורציה באוכלוסייה בלא יותר מ-4%?
  - כיצד התשובה לסעיף הקודם הייתה משתנה אם הינו מגדילים את גודל המדגם?
- (4) נתון כי 80% מבתי האב מחוברים לאינטרנט. נדגמו 400 בתי אב אקראיים.
- מה ההסתברות שלפחות 340 מהם מחוברים לאינטרנט?
  - מה ההסתברות שפרופורציית המחוברים לאינטרנט במדגם תסטה מהפרופורציה האמיתית ביותר מ-4%?
  - כמה בתי אב יש לדגום כדי שהסטייה בין הפרופורציה המדגמית לפרופורציה האמיתית לא תעלה על 3% בהסתברות של 90%?
  - מהו העשירון התחתון של התפלגות פרופורציית המדגם?
- (5) נתון שציוני פסיכומטרי מתפלגים נורמלית עם תוחלת 500 וסטיית תקן 100. ב"מועדון ה-700" נכללים נבחנים שמקבלים ציון מעל 700 בפסיכומטרי. מה הסיכוי שבמועד בו נבחנו 2000 נבחנים אקראיים יהיו לפחות 3% המשתייכים למועדון?

6 נתון ש- $X \sim B(n, p)$ , ונגדיר את המשתנה הבא:  $\hat{P} = \frac{X}{n}$ .

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

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

### תשובות סופיות:

- 1 א. התוחלת: 0.1, השונות: 0.00064. ב. 0.5. ג. 0.3446.
- 2 א. 0.0618. ב. 0.0618. ג. 0.8238.
- 3 א. 0.5. ב. 0. ג. 0.8968. ד. גדלה.
- 4 א. 0.0062. ב. 0.0456. ג. 0.481. ד. 0.77436.
- 5 0.0154.
- 6 א. שאלת הוכחה. ב. 0.5.

## חוק המספרים הגדולים:

### רקע:

חוק המספרים הגדולים מתייחס להשפעת הגדלת גודל המדגם על הסיכוי של פרופורציית המדגם (או ממוצע המדגם) להיות קרובה מהפרופורציה האמיתית (או מהממוצע האימתי).

### החוק לגבי פרופורציה:

נניח שמבצעים מדגם מקרי מתוך אוכלוסייה אינסופית בה פרופורציית ההצלחות היא  $p$ , ככל שהמדגם גדול יותר: כך הסיכוי שפרופורציית המדגם ( $\hat{p}$ ) תהיה בקרבת הפרופורציה באוכלוסייה ( $P$ ) גבוה יותר.

ולכן הסיכוי לקבל ערך חריג הרחוק מהפרופורציה של האוכלוסייה קטן יותר. בכתיבה מתמטית רושמים את חוק המספרים הגדולים לגבי הפרופורציה באופן

$$\text{הבא: } \lim_{n \rightarrow \infty} \hat{P}_n = P$$

בספרות המקצועית קוראים לחוק הזה החוק החזק של המספרים הגדולים.

את החוק החלש של המספרים הגדולים רושמים באופן הבא:  $P(|\hat{P} - P| < \varepsilon) \xrightarrow{n \rightarrow \infty} 1$ .

**הערה:** ככל שהמדגם גדל הסיכוי שפרופורציית המדגם תהיה בדיוק הפרופורציה האמיתית הולך וקטן.

**החוק לגבי ממוצע:** נניח שמבצעים מדגם מקרי מתוך התפלגות שבה התוחלת  $\mu$  והשונות סופית. ככל שהמדגם גדול יותר, כך הסיכוי שממוצע המדגם ( $\bar{X}$ ) יהיה בקרבת הממוצע באוכלוסייה ( $\mu$ ) גבוה יותר. לכן, הסיכוי לקבל ערך חריג הרחוק מהממוצע של האוכלוסייה קטן יותר. בכתיבה מתמטית רושמים את חוק המספרים הגדולים לגבי הממוצע באופן הבא:

$$\lim_{n \rightarrow \infty} \bar{X}_n = \mu$$

בספרות המקצועית קוראים לחוק הזה החוק החזק של המספרים הגדולים.

את החוק החלש של המספרים הגדולים רושמים באופן הבא:  $P(|\bar{X} - \mu| < \varepsilon) \xrightarrow{n \rightarrow \infty} 1$ .

### דוגמה (פתרון בהקלטה):

באוכלוסייה מסוימת 20% מהאוכלוסייה מובטלת. איזה סיכוי יותר גבוה? במדגם בגודל 100 פרופורציית המובטלים תסטה מהפרופורציה האמיתית בלא יותר מ-4%. במדגם בגודל 200 פרופורציית המובטלים תסטה מהפרופורציה האמיתית בלא יותר מ-4%. הסבירו.

## שאלות:

- (1) באוכלוסייה מסוימת 20% מהאוכלוסייה מובטלת. איזה סיכוי יותר גבוה?  
 א. אחד מתוך מדגם של חמישה יהיה מובטל.  
 ב. שניים מתוך עשרה יהיו מובטלים. הסבירו וחשבו.
- (2) באוכלוסייה מסוימת 20% מהאוכלוסייה מובטלת. איזה סיכוי יותר גבוה?  
 א. לפחות 3 מתוך 10 יהיו מובטלים  
 ב. לפחות 30 מתוך 100 יהיו מובטלים. הסבירו.
- (3) גובה של אוכלוסייה מסוימת מתפלג נורמלית עם ממוצע 170 ס"מ וסטיית תקן של 10 ס"מ. דוגמים 4 אנשים באקראי.  
 א. מה ההסתברות שהגובה הממוצע שלהם יהיה מעל 176 ס"מ.  
 ב. כיצד התשובה לסעיף הקודם הייתה משתנה אם הינו מגדילים את גודל המדגם? נמקו.
- (4) ידוע שבהצעת חוק מסוימת תומכים 60% מהציבור. נדגמו באקראי 10 אנשים.  
 א. מה ההסתברות שבדיוק 60% מהמדגם תומכים בהצעת החוק?  
 ב. כיצד התשובה הייתה משתנה אם היו דוגמים 20 אנשים?
- (5) שני חוקרים ביצעו מדגם מאותה אוכלוסייה. חוקר א דגם 20 תצפיות והשני דגם 40 תצפיות וכל אחד מהם חישב את ממוצע המדגם:  $\bar{X}_{20}$  ו-  $\bar{X}_{40}$ . ידוע שהתפלגות היא נורמלית ושהתוחלת באוכלוסייה הינה 500. בסעיפים הבאים נמקו אילו הסתברויות מהזוגות המוצגים גדולה יותר או שהן שוות ונמקו.  
 א.  $P(\bar{X}_{40} > 500)$  או  $P(\bar{X}_{20} > 500)$   
 ב.  $P(480 < \bar{X}_{40} < 520)$  או  $P(480 < \bar{X}_{20} < 520)$   
 ג.  $P(\bar{X}_{40} < 490)$  או  $P(\bar{X}_{20} < 490)$
- (6) נתון ש:  $X \sim G(P=0.1)$ . מבצעים מדגם אקראי בגודל  $n$  מההתפלגות הזו ומחשבים את ממוצע המדגם:  $\bar{X}_n$ . הוכיחו:  $\lim_{n \rightarrow \infty} \bar{X}_n = 10$ .

### תשובות סופיות:

- (1) אחד מתוך מדגם של חמישה יהיה מובטל.
- (2) לפחות 3 מתוך 10 יהיו מובטלים.
- (3) א. 0.1151. ב. קטנה.
- (4) א. 0.2508. ב. קטן.
- (5) א.  $P(\bar{X}_{40} > 500) = P(\bar{X}_{20} > 500)$ .
- ב.  $P(480 < \bar{X}_{40} < 520) > P(480 < \bar{X}_{20} < 520)$ .
- ג.  $P(\bar{X}_{40} < 490) < P(\bar{X}_{20} < 490)$ .
- (6) שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 59 - אי שוויונים בהסתברות

תוכן העניינים

275 .....	1. אי שוויון מרקוב
279 .....	2. אי שוויון צ'בישב

## אי-שוויון מרקוב:

### רקע:

אי-שוויון מרקוב רלבנטי לשימוש עבור כל משתנה מקרי אי-שלילי.

הפרמטר  $a$  הינו ממשי וחיובי ואז חייב להתקיים ש:  $P(X \geq a) \leq \frac{E[X]}{a}$ ,

לכן מתקיים גם ש:  $P(X < a) \geq 1 - \frac{E[X]}{a}$ .

### דוגמה:

אורך חיים של מכשיר מתפלג עם תוחלת של 500 שעות. חשבו לפי אי-שוויון מרקוב את ההסתברות שאורך חיים של מכשיר יהיה לפחות 1500 שעות.

$X$  = אורך חיים של מכשיר (רציף).

$$X \geq 0, E[X] = 500$$

$$P(X \geq 1500) \leq \frac{E[X]}{a} = \frac{500}{1500} = \frac{1}{3}$$

לכן,  $0 \leq P(X \geq 1500) \leq \frac{1}{3}$ .

## שאלות:

- (1) ידוע מניסיון העבר כי ציון במבחן הגמר של סטודנט הוא משתנה מקרי שתוחלתו 65. מצאו חסם עליון להסתברות שציון מבחן הגמר של סטודנט יהיה לפחות 75.
- (2) התפלגות מספר הילדים למשפחה במדינה מסוימת היא עם תוחלת של 2 ילדים. נלקחו 5 משפחות אקראיות. העריכו את הסיכוי שבסה"כ בחמשת המשפחות יש יותר מ-15 ילדים.
- (3)  $X$  משתנה מקרי רציף אי-שלילי, שתוחלתו 15. האם ייתכן ש:  $P(X > 65) = 0.3$  ?
- (4)  $X$  משתנה מקרי בדיד, המקיים:  $X > -2$ , ותוחלתו 6. מצאו חסם תחתון ל-  $P(X < 10)$ .
- (5)  $X$  משתנה מקרי המקיים:  $P(X \geq 0) = 1$  ו-  $s > 0$  קבוע ממשי. הוכיחו כי:  $P(X < sE(X)) \geq \frac{s-1}{s}$ .
- (6) נתון ש-  $X_i \sim P(\lambda)$ ,  $i = 1, 2, \dots, n$  הם משתנים בלתי תלויים. מצאו חסמים להסתברות ש-  $\sum_{i=1}^n X_i \geq 3n\lambda$ .
- (7) הוכיחו את אי-שוויון מרקוב. רמז: היעזרו במשתנה אינדיקטור המקבל את הערך 1 כאשר  $X \geq a$ .

(8) בשכונה חדשה בונים  $n$  בתים חדשים הנבנים בחלקת אדמה עגולה. כל בית נצבע בלבן בסיכוי  $p$  ללא תלות בבתיים האחרים. בניין שלא נצבע בלבן נצבע באפור. בית לבן הוא בית בודד אם הוא נמצא בין שני בתים אפורים. נגדיר את  $X$  להיות מספר הבתים הלבנים הבודדים.

א. מצאו את התוחלת של  $X$ .

ב. כעת נניח ש- $p = \frac{1}{4}$ . הראו שהסיכוי שמספר הבתים הלבנים הבודדים

$$\text{יהיה קטן מ-} \frac{n}{4} \text{ הוא לפחות } \frac{7}{16}.$$

(9) הוכיחו את אי-שוויון צ'בישב: אם  $X$  הוא משתנה מקרי שתוחלתו ושונותו הן

$$P\{|X - E(X)| \geq k\} \leq \frac{\text{Var}(X)}{k^2}.$$

רמז: היעזרו באי-שוויון מרקוב.

## תשובות סופיות:

$$. P(X \geq 75) \leq \frac{65}{75} = \frac{13}{15} \quad (1)$$

$$. 0 \leq P\left(\sum_{i=1}^n X_i \geq 3n\lambda\right) \leq \frac{1}{3} \quad (2)$$

(3) לא יתכן.

$$. \frac{1}{3} \quad (4)$$

(5) שאלת הוכחה.

$$. 0 \leq P\left(\sum_{i=1}^n X_i \geq 3n\lambda\right) \leq \frac{1}{3} \quad (6)$$

(7) שאלת הוכחה.

(8) א.  $n \cdot p(1-p)^2$ . ב. שאלת הוכחה.

(9) שאלת הוכחה.

## אי-שוויון צ'בישב:

**רקע:**

אם  $X$  הוא משתנה מקרי שתוחלתו ושונוותו הן סופיות, אז לכל ערך  $a$  חיובי

$$. P\{|X - E(X)| \geq a\} \leq \frac{\text{Var}(X)}{a^2} \quad \text{מתקיים:}$$

$$. P\{|X - E(X)| < a\} \geq 1 - \frac{\text{Var}(X)}{a^2} \quad \text{מכאן גם נובע שמתקיים:}$$

אי-שוויון צ'בישב נותן חסמים להסתברות סימטרית סביב התוחלת ללא צורך בדיעת ההתפלגות של המשתנה המקרי  $X$ .

**דוגמה:**

נתון משתנה מקרי עם סטיית תקן של 3. האם יתכן שההסתברות שהסטייה של המשתנה המקרי מתוחלתו תהיה קטנה מ-5 היא 0.6?

$$\sigma(X) = 3$$

$$: \text{נציב, } P\{|X - E(X)| < a\} \geq 1 - \frac{\text{Var}(X)}{a^2}$$

$$P\{|X - E(X)| < 5\} \geq 1 - \frac{3^2}{5^2} = 1 - \frac{9}{25} = \frac{16}{25} = 0.64$$

. לכן לא יתכן,  $P\{|X - E(X)| < 5\} \neq 0.6$ .

## שאלות:

- (1) מצאו חסמים להסתברויות הבאות עבור משתנה מקרי רציף בעל תוחלת 8 וסטית תקן 3:
- א.  $p(2 < x < 14)$ .
- ב.  $p(|x - 8| \geq 9)$ .
- (2) האם קיים משתנה מקרי  $X$  בעל תוחלת  $\mu$  וסטית תקן  $\sigma$  שעבורו מתקיים  $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.7$ ? הסבירו.
- (3) מספר המטוסים המגיעים לנמל תעופה ב-20 דקות מתפלג התפלגות פואסונית עם תוחלת של 100. היעזרו באי-שוויון צ'בישב כדי למצוא גבול תחתון להסתברות שמספר המטוסים המגיעים בתקופה בת 20 דקות נתונה תהיה בין 80 ל-120.
- (4) מטילים מטבע 120 פעמים. מה ניתן להגיד על הסיכוי שהתוצאה עץ תתקבל בין 50 ל-70 פעמים לפי אי-שוויון צ'בישב?
- (5) מתוך קו יצור של רכיבים שאורכם הממוצע הנו 10 ס"מ ושונוותם 3 סמ"ר יש לקחת מדגם. מהו גודל המדגם שיבטיח שבהסתברות של 0.9 לפחות ימצא ממוצע המדגם בין 9 ל-11 ס"מ?
- (6) אחוז התומכים במפלגה מסוימת הנו 40%. נלקח מדגם מקרי בגודל 200. תנו חסם תחתון לכך שאחוז התומכים במדגם יהיה בין 35% ל-45%.
- (7) בוחרים קוד  $n$  ספרתי באופן מקרי.
- א. עבור  $n = 10$ , העריכו את ההסתברות שממוצע הספרות במספר יסטה מתוחלתו בלפחות 1.
- ב. מה אורך הקוד המינימלי ( $n$ ) שיבטיח שבהסתברות של לפחות 95%, ממוצע הספרות יסטה מתוחלתו בפחות מ-0.75?

(8) בעיר מסוימת ל-5% מהמשפחות אין מכונית, ל-20% יש מכונית אחת, ל-35% יש שתי מכוניות, ל-30% שלוש מכוניות וליתר ארבע מכוניות. נניח שמספר המשפחות בעיר גדול מאד. העריכו את ההסתברות שמספר המכוניות הכולל בעשר משפחות אקראיות יהיה לפחות 17 ולכל היותר ל-27.

(9) הם משתנים מקיים המתפלגים גיאומטרית עם פרמטר  $p$  באופן

$$P\left(\frac{\sum_{i=1}^n X_i}{n} \geq \frac{2}{p}\right) \leq \frac{1-p}{n} : 0 < p < 1. \text{ הוכיחו שמתקיים:}$$

(10) הם משתנים מקיים המתפלגים פואסונית עם פרמטר  $\lambda_i = i$ ,

$$T = \sum_{i=1}^n X_i : \text{ באופן בלתי תלוי זה בזה. נסמן את:}$$

$$P(|2T - n(n+1)| < 2n) \geq \frac{n-1}{2n} : \text{ הוכיחו שמתקיים:}$$

**תשובות סופיות:**

- (1) א. בין  $\frac{3}{4}$  ל-1. ב. בין 0 ל- $\frac{1}{9}$ .
- (2) לא יתכן.
- (3) 0.75
- (4) לפחות 0.7
- (5) לפחות 30
- (6) 0.52
- (7) א. לכל היותר 0.825. ב. 294
- (8) 0.7056
- (9) שאלת הוכחה.
- (10) שאלת הוכחה.

# סטטיסטיקה והסתברות 01 88165

פרק 60 - מושגי יסוד באמידה

תוכן העניינים

1. כללי ..... 283

## מושגי יסוד באמידה:

### רקע:

כזכור מהמפגש הקודם, פרמטר הוא גודל המתאר את האוכלוסייה או התפלגות מסוימת. כמו ממוצע הגבהים בקרב מתגייסים לצה"ל -  $\mu$ . כמו פרופורציית התומכים בממשלה בקרב אזרחי המדינה -  $p$ . בדרך כלל הפרמטרים הם גדלים שאינם ידועים באמת, ולכן מבצעים מדגמים במטרה לאמוד אותם. אין אפשרות לחשב אותם הניסיון הוא בלהעריך כמה הם שווים ככל שניתן.

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

### דוגמה (פתרון בהקלטה):

בכנסת ה-19 קיבלה מפלגת העבודה 15 מנדטים. בערוץ 10 ברגע סגירת הקלפיות העריכו את מספר המנדטים של המפלגה להיות 17 מנדטים וזאת על סמך תוצאות מדגם של הערוץ.

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

**פרמטרים מרכזיים והאומדים שלהם:**

**ממוצע האוכלוסייה  $\mu$ :**

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

$$E(\bar{x}) = \mu \quad \text{לכן } \bar{x} \text{ הינו אומר חסר הטיה ל-} \mu \text{ . כמו כן, טעות תקן: } \sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}} = SE$$

**פרופורציה באוכלוסייה  $p$ :**

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

$$E(\hat{p}) = p, \quad \text{לכן } \hat{p} \text{ הינו אומר חסר הטיה ל-} p \text{ . כמו כן טעות התקן: } \sigma(\hat{p}) = \sqrt{\frac{p \cdot (1-p)}{n}}$$

**שונות האוכלוסייה  $\sigma^2$ :**

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

$$E(S^2) = \sigma^2 \quad \text{ולכן } S^2 \text{ הינו אומד חסר הטיה ל-} \sigma^2 \text{ .}$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$$

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

**דוגמה (פתרון בהקלטה):**

נדגמו 10 משפחות בתל אביב ונבדק עבור כל משפחה מספר הילדים שלה.  
 להלן התוצאות שהתקבלו: 2, 1, 3, 2, 1, 4, 5, 2, 1, 3.  
 אמדו באמצעות אומדים 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. נתון שהסיכוי שטירון יהיה עם שבר הליכה הוא 0.25.
- מהי האוכלוסייה המוצגת בשאלה? מהם הפרמטרים שלה?
  - מהי טעות התקן של האומדן כשהמדגם בגודל 500?
  - מהו האומדן לפרמטר?
  - מהי טעות האמידה?
- (2) לפי נתוני היצרן, מקרר צורך בממוצע 2400 וואט לשעה עם סטיית תקן של 500 וואט לשעה.
- במדגם של 25 מקררים של היצרן התקבל ממוצע של 2342 וואט לשעה.
- מהי האוכלוסייה המוצגת בשאלה? מהם הפרמטרים שלה?
  - מהי טעות התקן של האומדן?
  - מהו האומדן לפרמטר?
  - מהי טעות האמידה?
- (3) נדגמו עשרה מתגייסים לצה"ל. גובהם נמדד בס"מ. להלן התוצאות שהתקבלו: 168, 184, 192, 171, 180, 177, 187, 168, 177 ו-175.
- מצאו אומדן חסר הטיה לגובה הממוצע של מתגייסי צה"ל.
  - מצאו אומדן חסר הטיה לשונות הגבהים של מתגייסי צה"ל.
  - מצאו אומדן חסר הטיה לפרופורציות המתגייסים בגובה של לפחות 180 ס"מ.
- (4) נדגמו 20 שכירים באקראי. עבור כל שכיר נמדד השכר באלפי שקלים.
- להלן התוצאות שהתקבלו:  $\sum_{i=1}^{20} X_i = 162$ ,  $\sum_{i=1}^{20} X_i^2 = 1502.2$ .
- אמדו את השכר הממוצע של השכירים במשק.
  - אמדו את סטיית התקן של שכר השכירים במשק.
- (5) במטרה לאמוד את ממוצע האוכלוסייה, דגמו תצפיות 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 תהייה:

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

7) במדגם מקרי, מתי 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 נורמאלית.
- ד. כאשר מעוניינים באומדן חסר הטיה לשונות האוכלוסייה ממנה הוצא המדגם.
- ה. כאשר מעוניינים לחשב את שונות התפלגות הדגימה של ממוצע המדגם.

8)  $X_1, X_2, \dots, X_{16}$  מדגם מקרי מתוך אוכלוסייה בעלת ממוצע  $\mu$  לא ידוע ושונות:  $\sigma^2 = 64$ . טעות התקן של האומדן ל- $\mu$  היא:

- א. 16.
- ב. 8.
- ג. 4.
- ד. 2.

9) מהו אומדן חסר הטיה?

- א. אומדן שערכו שווה לממוצע התפלגות הדגימה שלו.
- ב. אומדן שערכו שווה לערך הפרמטר באוכלוסייה.
- ג. אומדן שממוצע התפלגות הדגימה שלו שווה לערך הפרמטר באוכלוסייה.
- ד. אומדן שהסיכוי שערכו יהיה גבוה מערך הפרמטר באוכלוסייה שווה לסיכוי שיהיה נמוך ממנו.

### תשובות סופיות:

- (1) א. 0.25    ב. 0.019    ג. 0.24    ד. 0.01
- (2) א. אוכלוסייה: מקררים של יצרן, תוחלת: 2400, סטיית תקן: 500.  
 ב. 100    ג. 2342    ד. 58
- (3) א. 177.9    ב. 64.1    ג. 0.4
- (4) א. 8.1    ב. 3.16
- (5) ד'
- (6) ג'
- (7) ד'
- (8) ד'
- (9) ג'

# סטטיסטיקה והסתברות 01 88165

פרק 61 - אמידה נקודתית

תוכן העניינים

1. אומדן חסר הטייה ..... 288
2. אומדן ניראות מקסימלית ..... 295

## אומד חסר הטייה:

רקע:

$\hat{\theta}$  יהיה אומד חסר הטייה ל- $\theta$ , אם התוחלת של  $\hat{\theta}$  תהיה שווה ל- $\theta$ :  $E(\hat{\theta}) = \theta$ .

דוגמה (פתרון בהקלטה):

המשתנה  $X$  הוא בעל פונקציית ההסתברות הבאה:

3	2	1	$X$
$4\theta$	$1 - 60\theta$	$2\theta$	הסתברות

מעוניינים לאמוד את  $\theta$  על סמך שתי תצפיות מההתפלגות:  $X_1$  ו- $X_2$ .

א. הראו שהאומד:  $T_1 = \frac{2X_1 + X_2}{2}$ , הוא אומד מוטה ל- $\theta$ .

הטיה של אומד היא:  $E(\hat{\theta}) - \theta$ . כמובן שלאומד חסר הטיה אין הטיה.

ב. מהי ההטיה של האומד  $T_1$ ?

ג. תקנו את  $T_1$ , כך שיהיה אומד חסר הטיה.

אם יש שני אומדים חסרי הטיה עדיף זה עם השונות היותר קטנה.

ד. מוצא האומד הבא:  $T_3 = 1.5X_1 - X_2 - 1$ .

האם הוא עדיף על האומד שהצעת בסעיף ג'?

אם  $\hat{\theta}$  אומד חסר הטיה ל- $\theta$ , אז  $g(\hat{\theta})$  יהיה אומד חסר הטיה עבור  $g(\theta)$ , רק אם  $g$  תהיה לינארית.

ה. מצאו אומד חסר הטיה ל:  $P(X = 3)$ .

אומד חסר הטיה לשונות האוכלוסייה  $\sigma^2$ :  $S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$ .

ו. מצאו אומד חסר הטיה לשונות של  $X$ .

**תזכורות חשובות:**

אם:  $Y = aX + b$ , אזי:  $E(Y) = aE(X) + b$ ,  $V(Y) = a^2 \cdot V(X)$ ,  $\sigma_Y = |a|\sigma_X$ .

אם:  $X_1, X_2, \dots, X_n$  משתנים מקריים, אזי:

$$E(T) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

אם:  $X_1, X_2, \dots, X_n$  משתנים מקריים בלתי תלויים בזוגות, אזי:

$$V(T) = V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$$

## שאלות:

- (1) הציון במבחן מסוים של תלמידי כתה ח' הנו משתנה מקרי בעל תוחלת  $\mu$  וסטיית תקן 10. כדי לאמוד את התוחלת  $\mu$ , נלקח מדגם של 5 ציונים:  $X_1, \dots, X_5$ . שלושה חוקרים הציעו אומדים לתוחלת על סמך מדגם זה:

$$T_1 = \frac{X_1 + \dots + X_5}{5} \quad \text{חוקר א' הציע:}$$

$$T_2 = \frac{2X_1 - X_3 + X_4}{2} \quad \text{חוקר ב' הציע:}$$

$$T_3 = \frac{2X_1 + X_3}{2} \quad \text{חוקר ג' הציע:}$$

- איזה מן האומדים הוא חסר הטיה?
- הציעו תיקון לאומד המוטה כך שיהיה חסר הטיה.
- במדגם התקבלו הציונים הבאים: 100, 82, 58, 78, 65. חשבו את האומדנים המתקבלים עבור האומדים חסרי ההטיה.
- איזה מבין שני האומדים חסרי ההטיה עדיף? נמקו.

- (2) כדי לאמוד את המשקל הממוצע של הנשים בארה"ב, נבחר מדגם של  $2n$  נשים. נסמן את שונות הגובה ב- $\sigma^2$ . הוצעו שני אומדים לממוצע המשקל על סמך מדגם

$$T_1 = \frac{1}{n} \sum_{i=1}^n X_i, \quad T_2 = \frac{1}{2n} \sum_{i=1}^{2n} X_i \quad \text{זה:}$$

- בדקו לגבי כל אומד אם הוא בלתי מוטה.
- איזה אומד עדיף? נמקו.

- (3)  $X \sim B(n, p)$ . כלומר,  $X$  הינו משתנה מקרי המתפלג בינומית עם פרמטר  $P$  (סיכוי להצלחה בניסיון בודד) במדגם בגודל  $n$ .

- פתחו אומד חסר הטיה ל- $P$ .
- מהו אומד חסר הטיה לסיכוי לכישלון בניסיון בודד?
- מהו אומד חסר הטיה ל- $E(X)$ ?
- מצאו אומד חסר הטיה ל- $E(X^2)$ .

4) בתיק מניות שתי מניות. מספר המניות שיעלו ביום מסוים הוא משתנה מקרי התלוי בפרמטר לא ידוע:  $\theta$ ,  $0 \leq \theta \leq 2$ .

פונקציית ההסתברות של  $X$  - מספר המניות שיעלו ביום מסוים:

$$P(X=0) = 1 - \frac{\theta}{2}, \quad P(X=1) = \frac{\theta}{3}, \quad P(X=2) = \frac{\theta}{6}$$

א. מצאו אומד בלתי מוטה ל- $\theta$ , שמתבסס על מספר המניות שיעלו ביום מסוים.

ב. מצאו אומד בלתי מוטה ל- $\theta$ , שמתבסס על מספר המניות שעלו ביום,

במשך שלושה ימים -  $X_1, X_2, X_3$  (לכל אחד מהם אותה התפלגות כנ"ל

והם בלתי תלויים).

5) בקרב המטפלות בת"א, מספר התינוקות שבטיפולן הוא משתנה מקרי בעל התפלגות התלויה בפרמטר  $\theta$  באופן הבא:

הסיכוי שמטפלת תטפל בתינוק אחד בלבד הוא  $3\theta$ ,

הסיכוי שמטפלת תטפל ב-2 תינוקות הוא  $1 - 4\theta$ ,

הסיכוי שמטפלת תטפל ב-3 תינוקות הוא  $\theta$ .

במדגם מיקרי של 4 מטפלות מת"א, נמצא כי שתיים מהם מטפלות בתינוק אחד בלבד, אחת מהן בשנים ואחת השלושה תינוקות.

א. מצאו אומד חסר הטיה לפרמטר  $\theta$  על סמך תצפית בודדת.

ב. מצאו אומד חסר הטיה לפרמטר  $\theta$  על סמך 4 תצפיות.

ג. מהו האומדן לפרמטר  $\theta$  על סמך תוצאות המדגם.

ד. מצאו אומד חסר הטיה לסיכוי שלמטפלת בת"א תטפל בתינוק בודד אחד.

ה. מצאו אומדים חסרי הטיה לתוחלת ולשונות של מספר התינוקות בטיפול אצל מטפלת מת"א. חשבו אומדנים.

6) קבעו אילו מהטענות הבאות נכונות:

א. אם  $T$  הוא אומד בלתי מוטה עבור פרמטר  $\theta$ , אז  $5T$  אומד בלתי מוטה

עבור הפרמטר  $5\theta$ .

ב. אם  $T$  הוא אומד בלתי מוטה עבור פרמטר  $\theta$ , אז  $T^2$  אומד בלתי מוטה

עבור הפרמטר  $\theta^2$ .

- (7) במפעל שתי מכונות המייצרות מוצרים. במכונה הראשונה ההסתברות שמכשיר תקין היא  $p$ , ובמכונה השנייה ההסתברות שמכשיר תקין היא  $2p$ . דוגמים 20 מכשירים מהייצור של כל מכונה. נסמן ב- $X$  את מספר המכשירים התקינים שיוצרו על ידי המכונה הראשונה, וב- $Y$  את מספר המכשירים התקינים שיוצרו על ידי המכונה השנייה. איזה מבין האומדים הבאים אינו אומד חסר הטיה ל- $p$ ?

א.  $\frac{X}{20}$ .

ב.  $\frac{Y}{20}$ .

ג.  $\frac{X+Y}{60}$ .

ד.  $\frac{2X+Y}{80}$ .

- (8) יהיו  $T_1$  ו- $T_2$  אומדים חסרי הטיה ובלתי תלויים לפרמטר  $\theta$ .  
 א. מצאו אומד חסר הטיה ל- $\theta^2$ , המתבסס על  $T_1$  ו- $T_2$ .  
 ב. מצאו אומד חסר הטיה ל- $\theta(1-\theta)$ , המתבסס על  $T_1$  ו- $T_2$ .

- (9) נתון ש- $X$  הינו משתנה מקרי עם תוחלת  $\mu$  ושונות  $\sigma^2$ . נדגמו  $n$  תצפיות בלתי תלויים מאותה אוכלוסיה.

א. הראו ש- $\sum_{i=1}^n p_i x_i$  אומד חסר הטיה ל- $\mu$ , כאשר:  $\sum_{i=1}^n p_i = 1$ .

ב. נתבונן במכפלת שתי התצפיות הראשונות:  $X_1 \cdot X_2$ .

הראו שהוא אומד חסרי הטיה ל- $\mu^2$ .

- (10)  $X_i \sim N(\mu, 1)$ , כאשר:  $i = 1, 2, \dots, n$ . נתון שהתצפיות הינן בלתי תלויות זו בזו. מצאו אומד חסר הטיה ל- $\mu^2$ .

**(11)** נתונות  $n$  תצפיות בלתי תלויות מתוך התפלגות בעלת הצפיפות הבאה :

$$f(x) = \begin{cases} \frac{1+\beta x}{2} & -1 < x < 1 \\ 0 & \text{else} \end{cases}$$

- א. הראו כי האומד  $3\bar{X}$  הנו אומד בלתי מוטה ל- $\beta$ .  
 ב. מצאו את השונות של האומד מהסעיף הקודם.

**(12)**  $X_1, X_2, \dots, X_n$  הינם משתנים מקריים רציפים בלתי תלויים בעלי פונקציית

$$f(x) = \begin{cases} X \cdot A & 0 \leq x \leq \theta \\ 0 & \text{אחר ת} \end{cases} \quad \text{הצפיפות הבאה :}$$

- א. בטאו את ערכו של  $A$  באמצעות  $\theta$ , כדי שפונקציית הצפיפות תהיה לגיטימית.  
 ב. מצאו אומד חסר הטיה ל- $\theta$ , על סמך  $n$  התצפיות.

## תשובות סופיות:

$$(1) \quad \text{א. } T_1 \text{ ו- } T_2 \quad \text{ב. } \frac{2}{3} T_3 \quad \text{ג. } T_1 = 76.6, T_2 = 110 \quad \text{ד. } T_1$$

$$(2) \quad \text{א. ראו בוידאו.} \quad \text{ב. } T_2$$

$$(3) \quad \text{א. } \frac{x}{n} \quad \text{ב. } 1 - \frac{x}{n} \quad \text{ג. } X \quad \text{ד. } \theta$$

$$(4) \quad \text{א. } \frac{3x}{2} \quad \text{ב. } \frac{3\bar{x}}{2}$$

$$(5) \quad \text{א. } 1 - \frac{x}{2} \quad \text{ב. } 1 - \frac{1}{2} \bar{x} \quad \text{ג. } 0.125 \quad \text{ד. } 3 \left( 1 - \frac{1}{2} \bar{x} \right)$$

ה. לשונות 0.917.

$$(6) \quad \text{א. נכון.} \quad \text{ב. לא נכון.}$$

(7) ב'.

$$(8) \quad \text{א. } T_1 \cdot T_2 \quad \text{ב. } T_1 - T_1 \cdot T_2$$

(9) א. שאלת הוכחה. ב. שאלת הוכחה.

$$(10) \quad \bar{X}^2 - \frac{1}{n}$$

$$(11) \quad \text{א. שאלת הוכחה.} \quad \text{ב. } V(3\bar{X}) = \frac{3 - \beta^2}{n}$$

$$(12) \quad \text{א. } A = \frac{2}{\theta^2} \quad \text{ב. } \theta = \frac{3 - \bar{X}}{2}$$

## אומד נראות מקסימלית:

### רקע:

להלן נלמד את שיטת הנראות המקסימלית למציאת אומדים. נניח ש- $X$  משתנה מקרי בדיד עם פונקציית הסתברות  $P(x, \theta)$ , כאשר  $\theta$  הפרמטר הבלתי ידוע.

יהיו:  $X_1, X_2, \dots, X_n$  תוצאות מדגם מקרי בגודל  $n$  הנלקח מאוכלוסייה זו.

נבנה את פונקציית ההסתברות המשותפת (פונקציית הדגימה).

אם אנו יודעים את תוצאות המדגם, ולא את הפרמטר, קוראים לפונקציית הנראות שהיא פונקציה של הפרמטר.

נגדיר את פונקציית הנראות:

$$L(\theta) = P(x_1, \theta) \cdot P(x_2, \theta) \cdot \dots \cdot P(x_n, \theta) = \prod_{i=1}^n P(x_i, \theta)$$

פונקציית הנראות היא ההסתברות לקבל את התצפית הראשונה (כפונקציה של  $\theta$ ), כפול ההסתברות לקבל את התצפית השנייה, וכולי. כלומר, המשמעות של פונקציית הנראות היא ההסתברות לקבל את המדגם שהתקבל, כפונקציה של הפרמטר המבוקש  $\theta$ .

אם מדובר במשתנה רציף, נכפיל את פונקציות הצפיפות ולא את פונקציות ההסתברות:

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta) = \prod_{i=1}^n f(x_i, \theta)$$

### דוגמה (פתרון בהקלטה):

הסיכוי של שחקן כדורסל לקלוע לסל הוא  $p$  (לא ידוע). השחקן זורק כדורים לסל עד שהוא קולע בפעם הראשונה. נניח כי הזריקות בלתי תלויות זו בזו. הכדור נכנס לסל לראשונה בניסיון השלישי. השחקן חוזר על התהליך שוב, והפעם הכדור נכנס לסל בניסיון החמישי. מצאו את פונקציית הנראות של  $p$ .

אומד נראות מקסימלית עבור  $\theta$  הוא האומד  $\hat{\theta}$ , שממקסם את פונקציית הנראות  $L(\theta)$ . כלומר, אנו מחפשים את האומד שיגרום לכך שהמדגם המקרי שקיבלנו יהיה כמה שיותר סביר.

#### שלבים למציאת אומד נראות מקסימלית:

- לוקחים את פונקציית ההסתברות המשותפת של המדגם (או צפיפות משותפת אם המשתנה רציף).
- מציבים את תוצאות המדגם ומקבלים את פונקציית הנראות (פונקציה של הפרמטר הנחקר).
- מוצאים מקסימום לפונקציית הנראות (לעיתים כדאי להוסיף  $\ln$  כדי להקל על המלאכה).

#### המשך דוגמה:

חשבו את אומדן הנראות המקסימלית עבור  $p$ .

**משפט:** אם  $\hat{\theta}$  הוא אומד נראות מקסימלית עבור  $\theta$ , אזי  $g(\hat{\theta})$  הוא אומד נראות מקסימלית עבור  $g(\hat{\theta})$ , בהנחה והפונקציה היא חד-חד ערכית (אינווריאנטיות).

#### המשך דוגמה:

מצאו אומדן נראות מקסימלית לסיכוי של שחקן הכדורסל לקלוע לסל פעמיים ברצף.

## שאלות:

- (1) הסיכוי של שחקן לנצח במשחק הוא  $p$  (לא ידוע).  
 השחקן משחק במשחק עד אשר הוא מנצח בפעם הראשונה.  
 נתון שהשחקן ניצח לראשונה רק במשחק השני.  
 א. חשבו את פונקציית הנראות של  $p$ , וציירו גרף שלה.  
 ב. מצאו אומדן נראות מקסימלית עבור  $p$ .  
 ג. מצאו אומדן נראות מקסימלית ל- $p$ , אם ביום אחד הוא נאלץ לשחק 4 פעמים וביום אחר הוא נאלץ לשחק 5 פעמים, עד אשר ניצח.
- (2) מספר הלקוחות שנכנסים לחנות מסוימת, מתפלג פואסונית עם תוחלת של  $\lambda$  לקוחות ביום.  
 א. מצאו אומדן נראות מקסימלית ל- $\lambda$ , על סמך מספר הלקוחות שנכנסים ביום מסוים.  
 ב. מצאו אומדן נראות מקסימלית ל- $\lambda$ , על סמך מספר הלקוחות שנכנסים ב- $n$  ימים מסוימים.
- (3) הזמן שלוקח לאדם לחכות בתור מתפלג מעריכית עם פרמטר  $\lambda$ .  
 דגמו 4 אנשים מקריים שחיכו בתור ומדדו את זמני ההמתנה שלהם.  
 התוצאות שהתקבלו בדקות הן: 3, 5, 7 ו-3.  
 א. פתחו אומדן נראות מקסימלית לפרמטר זה על סמך  $n$  תצפיות כלשהן.  
 ב. מהו האומדן לפרמטר?
- (4) משך זמן הכנת שיעורי הבית (בשעות) של בני נוער, ביום אחד, מתפלג אחיד:  $U(0, q)$ .  
 כדי לאמוד את  $\theta$ , נשאלו ביום מסוים מספר בני נוער כמה שעות הם הכינו שיעורי-בית באותו יום.  
 א. אלעד הכין ביום מסוים שיעורי בית במשך שעה שלמה. חשבו את פונקציית הנראות של  $\theta$  המתבססת על תצפית זו, וציירו את הגרף שלה.  
 ב. מצאו אומדן נראות מקסימלית ל- $\theta$  על סמך התצפית.  
 ג. משכי הכנת שיעורי בית (שעות) של 3 בני נוער היו 1.5, 3, 1.  
 ד. מצאו אומדן נראות מקסימלית ל- $\theta$  על סמך המדגם הזה.  
 מצאו באופן כללי אומדן נראות מקסימלית ל- $\theta$ , על סמך מדגם של  $n$  בני נוער –  $X_1, \dots, X_n$ .

(5) הגובה של אוכלוסייה מסוימת מתפלג נורמלית עם תוחלת ידועה של 170 ס"מ ושונות  $\sigma^2$  לא ידועה.

- א. מצאו אומד נראות מקסימלית עבור השונות על סמך מדגם  $X_1, \dots, X_n$  מתצפיות מהאוכלוסייה.  
 ב. נדגמו 5 אנשים בלתי תלויים בעלי הגבהים: 170, 182, 165, 174, 174. מהו האומדן לשונות הגבהים באוכלוסייה?

(6) פתחו אומד נראות מקסימלית לפרמטר  $p$  בהתפלגות הבינומית, על סמך מדגם בגודל  $n$ , בו  $X$  הוא מספר ההצלחות במדגם.

$$(7) \quad f(x) = \begin{cases} 2\theta x e^{-\theta x^2} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad X \text{ הוא משתנה מקרי בעל פונקציית הצפיפות:}$$

- א. מצאו אומד נראות מקסימלית ל- $\theta$  על סמך  $n$  תצפיות בלתי תלויות:  $X_1, \dots, X_n$ .  
 ב. מצאו אומד נראות מקסימלית ל- $\theta^2$ .

(8) בכד א' 10 כדורים שחורים ו-10 לבנים ובכד ב' 5 כדורים שחורים ו-15 לבנים. דוגמים באקראי כדור, בלי לדעת מאיזה כד.

- א. מצא אומד נראות מקסימלית לכד שממנו הוצא הכדור על סמך הצבע של הכדור.  
 ב. מהו האומדן אם הצבע הוא שחור?

(9) הזמן שלוקח ליוסי לפתור תשבץ מתפלג מעריכית עם תוחלת לא ידועה. נתנו ליוסי לפתור חמישה תשבצים ובממוצע לקח לו 32 דקות לפתור אותם.

- א. מה אומדן הנראות המקסימלית לתוחלת זמן הפתרון של תשבץ על ידי יוסי (אין חובה לפתח).  
 ב. מה אומדן הנראות המקסימלית לסיכוי שייקח לו לפחות חצי שעה לפתור את התשבץ הבא?

10 מספר הלקוחות הממתינים בתור במוקד טלפוני הוא משתנה מקרי  $X$ , בעל התפלגות התלויה בפרמטר  $\theta$ , באופן הבא:

2	1	0	$X$
$1 - 4\theta + 4\theta^2$	$4\theta - 8\theta^2$	$4\theta^2$	$P(X)$

בחמישה זמנים שונים שנבחרו באקראי נמצאו: 0, 1, 0, 0, 0 לקוחות ממתינים בתור.

א. מצאו אומדן בשיטת הנראות המקסימלית עבור הפרמטר  $\theta$ , על-סמך המדגם הנתון.

ב. מצאו אומדן בשיטת הנראות המקסימלית לסיכוי שלא יהיו לקוחות בתור.

11 אדם מחזיק בידו שני מטבעות: מטבע הוגן ומטבע שאינו הוגן – שהסיכוי לקבל בו תוצאה של עץ הוא 0.2. האדם מטיל את אחד המטבעות פעמיים ומודיע לך כמה פעמים הוא קיבל עץ. אתה צריך לנחש איזה מטבע הוא הטיל: את ההוגן או את זה שאינו הוגן.

א. מצא אומדן בשיטת הנראות המקסימלית לסוג המטבע שהוטל.

ב. מהו האומדן אם האדם קיבל פעמיים עץ?

12 מעוניינים לאמוד את אחוז המובטלים באוכלוסייה.

דוגמים 50 אנשים אקראיים ומתקבל ש-4 מהם מובטלים.

א. מצא אומדן נראות מקסימלית לשיעור המובטלים באוכלוסייה.

ב. מצא אומדן לשיעור העובדים באוכלוסייה.

ג. מצא אומדן ליחס בין שיעור העובדים לשיעור המובטלים באוכלוסייה.

13 במשחק מחשב שלוש רמות משחק:

ברמה 1 הסיכוי של יוסי לסיים את המשחק הוא 0.9.

ברמה 2 הסיכוי של יוסי לסיים את המשחק הוא 0.7.

ברמה 3 הסיכוי של יוסי לסיים את המשחק הוא 0.4.

יוסי בחר ברמה מסוימת, אך אינו יודע באיזו רמה הוא בחר.

הוא משחק במשחק ברמה שבחר פעמיים.

א. הציעו א.נ.מ. לרמה של המשחק שיוסי שיחק, על סמך מספר הפעמים שסיים את המשחק.

ב. אם יוסי סיים את שני המשחקים, מה יהיה האומדן לרמה?

ג. מהו א.נ.מ. לסיכוי, שמתוך שני משחקים הוא יצליח בדיוק משחק אחד?

14  $X_1, X_2, \dots, X_n$  מתפלגים אחד בקטע:  $[-\theta, \theta]$ . מצא אומדן נראות מקסימלית לפרמטר  $\theta$ .

15  $X_1, X_2, \dots, X_n$  מתפלגים בדיד לפי פונקציית ההסתברות הבאה:

$$P(X = k) = \frac{\binom{2}{k} \cdot P^k \cdot (1-P)^{2-k}}{1 - (1-P)^2} \quad K = 1, 2$$

הוכח שא.נ.מ ל- $P$ , הינו:  $2 - \frac{2}{X}$ .

16 במכשיר חשמלי יש 2 סוללות שפועלות באופן ב"ת זו בזו, והוא מפסיק לפעול ברגע שאחת הסוללות מפסיקה לעבוד. הסיכוי של סוללה לתפקד לפחות חודש הוא  $P$ . כאשר המכשיר מפסיק לפעול מחליפים את שתי הסוללות שלו. בתחילת הניסוי נלקחו 80 מכשירים כאלה עם סוללות חדשות ולאחר חודש נמצא ש-30 מהם עדיין פועלים.

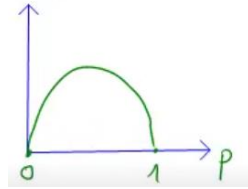
- מצא אומדן נראות מקסימלית עבור  $P$ .
- רשמו את האומדן שבו השתמשתם בחלק א' באופן כללי, עבור מדגם של  $n$  מכשירים שמתוכם נמצאו  $Y$  מכשירים שעדיין פועלים לאחר חודש אחד.
- בהנחה שאורך החיים (בחודשים) של סוללה בודדת הוא מעריכי, עם פי צפיפות:  $f(t) = \theta e^{-\theta t}$  עבור  $t > 0$ . מצא א.נ.מ. עבור  $\theta$ , המבוסס על  $Y$ . מהו האומדן המתאים מן המדגם הנתון?

17 חיוג אוטומטי של מכשיר טלפון משדר אות אחת לשתי דקות. אם לאחר 20 דקות (10 אותות חיוג) המספר שאליו מטלפנים עדיין תפוס, החיוג האוטומטי נפסק.

- רשמו את פונקציית ההסתברות של המשתנה  $X$  – מספר הפעמים שהחייגן האוטומטי מחייג למספר הטלפון המבוקש, אם ההסתברות לקבלת צליל "פנוי" בשידור אחד של אות חיוג הוא  $P$ .
- מתוך 12 ניסיונות חיוג אוטומטי למשרד הרישוי בזמנים שונים במשך 5 ימים, התקבלו התוצאות הבאות: בשני ניסיונות הופסק החיוג האוטומטי ובשאר הניסיונות שבהם הצליח המטלפן להשיג את המספר המבוקש, מספר החיוגים האוטומטיים עד לקבל צליל "פנוי" היו: 1, 6, 2, 7, 3, 8, 2, 2, 1, 5. מצאו אומדן נראות מקסימלית עבור  $P$ , על סמך התוצאות שהתקבלו.

## תשובות סופיות:

להלן גרף:



- (1) א.  $L(p) = (1-p) \cdot p$  .  
 ב. 0.5 .  
 ג.  $\frac{2}{9}$  .
- (2) א.  $\bar{X}$  .  
 ב.  $\bar{X}$  .
- (3) א.  $\frac{1}{\bar{X}}$  .  
 ב.  $\frac{2}{9}$  .
- (4) א. 1 .  
 ב. 1 .  
 ג. 3 .  
 ד.  $\hat{\theta} = \max \{X_1, \dots, X_n\}$  .
- (5) א.  $\frac{\sum_{i=1}^n (x_i - 170)^2}{n}$  .  
 ב. 40.2 .
- (6)  $\frac{x}{n}$  .
- (7) א.  $\sum X_i^2$  .  
 ב.  $\left( \sum X_i^2 \right)^2$  .
- (8) א. ראה סרטון .  
 ב. כד א' .
- (9) א. 32 .  
 ב. 0.3916 .
- (10) א. 0.45 .  
 ב. 0.81 .
- (11) א. ראה סרטון .  
 ב. הוגן .
- (12) א. 0.08 .  
 ב. 0.92 .  
 ג. 11.5 .
- (13) א.  $\hat{\theta} = \begin{cases} 3 & X = 0,1 \\ 1 & X = 2 \end{cases}$  .  
 ב. 1 .  
 ג.  $\hat{p} = \begin{cases} 2 \cdot 0.4 \cdot 0.6 & X = 0,1 \\ 2 \cdot 0.9 \cdot 0.1 & X = 2 \end{cases}$  .
- (14)  $\max |X_i|$  .
- (15) שאלת הוכחה .
- (16) א. 0.6124 .  
 ב.  $\hat{p} = \sqrt{\frac{y}{n}}$  .  
 ג. 0.49 .
- (17) א.  $P(x) = \begin{cases} (1-p)^{x-1} p & 1 \leq x \leq 9 \\ (1-p)^9 & x = 10 \end{cases}$  .  
 ב. 0.1818 .

**נספח:**  
**התפלגויות רציפות**

ההתפלגות	פונקציית הצפיפות	פונקציית ההתפלגות המצטברת	תוחלת	שונות	הערות	אנ"מ
$X \sim U(a, b)$	$f(x) = \frac{1}{b-a}$ $a \leq x \leq b$	$\frac{t-a}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$		$b = \max(X_i)$ $a = \min(X_i)$
$X \sim \exp(\lambda)$	$f(x) = \lambda e^{-\lambda x}$	$1 - e^{-x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	הזמן עד להתרחשות מאורע מסוים. $\lambda$ - הוא ממוצע האירועים ביחידת זמן.	$\hat{\lambda} = \frac{1}{\bar{X}}$
$X \sim N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\Phi(t)$	$\mu$	$\sigma^2$		$\hat{\mu} = \bar{X}$ $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ $Z = \frac{X - \mu}{\sigma}$

**התפלגויות בדידות**

ההתפלגות	פונקציית ההסתברות $P(X = k)$	תוחלת	שונות	הערות	אנ"מ
בינומית $B(n, p)$ $0 \leq p \leq 1$	$\binom{n}{k} p^k (1-p)^{n-k}$ $k = 0, 1, \dots, n$	$np$	$np(1-p)$	(1)	$\hat{p} = \frac{Y}{n}$
גיאומטרית $G(p)$ $0 < p \leq 1$	$(1-p)^{k-1} p$ $k = 1, 2, \dots, \infty$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	(2)	$\hat{p} = \frac{1}{\bar{X}}$
אחידה $U(a, b)$	$\frac{1}{b-a+1}$ $K = a, \dots, b$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2 - 1}{12}$	(3)	$b = \max(X_i)$ $a = \min(X_i)$
פואסונית $P(\lambda)$ $\lambda > 0$	$\frac{e^{-\lambda} \lambda^k}{k!}$ $k = 0, 1, \dots, \infty$	$\lambda$	$\lambda$	(4)	$\hat{\lambda} = \bar{X}$

(1) מספר ההצלחות ב-  $n$  ניסויי ברנולי ב"ת.  $p$  - ההסתברות להצלחה.

(2) מספר הניסויים עד להצלחה הראשונה בסדרת ניסויי ברנולי ב"ת,  $p$  - ההסתברות להצלחה.

(3) בחירה אקראית של מספר בין  $a$  ו- $b$ .

(4) מספר אירועים ביחידת זמן,  $\lambda$  - קצב האירועים.

# סטטיסטיקה והסתברות 01 88165

פרק 62 - רווח סמך לתוחלת (ממוצע)

תוכן העניינים

- 303 ..... 1. רווח סמך כששונות האוכלוסיה ידועה.
- 309 ..... 2. קביעת גודל מדגם.
- 311 ..... 3. רווח סמך כששונות האוכלוסיה לא ידועה.

## רווח סמך כששונות האוכלוסייה ידועה:

### רקע:

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

מה שנהוג לעשות כדי לאמוד את ממוצע האוכלוסייה, זה לבנות רווח סמך.

נבנה מרווח בטחון שהסיכוי שהפרמטר  $\mu$  ייכלל בתוכו הוא:  $1-\alpha$ .

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

A - גבול התחתון של רווח הסמך.

B - הגבול העליון של רווח הסמך.

$L = B - A$  - אורך רווח הסמך.

### דוגמה (פתרון בהקלטה):

חוקר דגם 25 חיילים שנבחנו במבחן הפסיכומטרי. הוא בנה רווח סמך לממוצע הציונים במבחן הפסיכומטרי בקרב אוכלוסיית החיילים וקיבל בין 510 ל-590. רווח הסמך נבנה ברמת סמך של 95%.

1. מהי אוכלוסיית המחקר?

2. מה המשתנה באוכלוסייה?

3. מה הפרמטר שהחוקר רצה לאמוד?

4. מהו רווח הסמך?

5. מה אורך רווח הסמך?

6. מהי רמת הביטחון של רווח הסמך?

בפרק זה נרצה לבנות רווח סמך לתוחלת ( $\mu$ ) במקרה ש- $\sigma^2$  (שוונות האוכלוסייה) ידועה.

פרמטר אותו נרצה לאמוד:  $\mu$ .

אומד נקודתי:  $\bar{x}$ .

תנאים לבניית רווח הסמך:  $X \sim N$  או  $n \geq 30$ .

$\sigma^2$  (שוונות האוכלוסייה) ידועה.

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

### דוגמה (פתרון בהקלטה):

על פי נתוני היצרן אורך חיי סוללה מתפלג נורמאלית עם סטיית תקן של 1 שעה. מעוניינים לאמוד את תוחלת חיי סוללה. נדגמו באקראי 4 סוללות, אורך החיים הממוצע שהתקבל הוא 13.5 שעות. בנו רווח סמך ברמת סמך של 95% לתוחלת אורך חיי סוללה.

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

$\varepsilon$  - נותן את שגיאת האמידה המקסימלית, דבר שנקרא גם טעות סטטיסטית, טעות דגימה.

### דוגמה (פתרון בהקלטה):

בהמשך לשאלה עם הסוללות. מה ניתן להגיד בביטחון של 95% על שגיאת האמידה?

קשרים מתמטיים ברווח הסמך:

• אורך רווח הסמך הוא פעמיים שגיאת האמידה המקסימלית:  $L = 2\varepsilon$ .

• ממוצע המדגם נופל תמיד באמצע רווח הסמך:  $\bar{X} = \frac{A+B}{2}$ .

• ככל שמספר התצפיות ( $n$ ) גבוה יותר, כך יש יותר אינפורמציה ולכן האומד יותר מדויק, ולכן נקבל רווח סמך יותר קצר.

• ככל שרמת הביטחון ( $1-\alpha$ ) גבוהה יותר, כך:  $z_{1-\frac{\alpha}{2}}$  יותר גבוה, ורווח הסמך יותר ארוך.

## שאלות:

- 1) חוקר התעניין לאמוד את השכר הממוצע במשק. על סמך מדגם הוא קבע שבביטחון של 95% כי השכר הממוצע במשק נע בין 9200 ל-9800 ₪.
- מי האוכלוסייה במחקר?
  - מה המשתנה הנחקר?
  - מה הפרמטר שאותו רוצים לאמוד?
  - מה רווח הסמך לפרמטר?
  - מהי רמת הסמך לפרמטר?
  - מה אורך רווח הסמך?
  - מה הסיכוי שטעות הדגימה תעלה על 300 ₪?
- 2) מעוניינים לאמוד את התפוקה היומית הממוצעת של מפעל מסוים ברמת סמך של 95%. במדגם אקראי של 100 ימים התקבלה תפוקה ממוצעת 4950 מוצרים ביום. לצורך פתרון הנח שסטיית התקן האמתית ידועה ושווה 150 מוצרים ביום. בנו את רווח הסמך.
- 3) מעוניינים לאמוד את ממוצע אורך החיים של מכשיר. מנתוני היצרן ידוע שאורך החיים מתפלג נורמאלי עם סטיית תקן של 20 שעות. נדגמו 25 מכשירים ונמצא כי ממוצע אורך החיים שלהם היה 230 שעות.
- בנו רווח סמך ברמת סמך של 90% לאורך החיים הממוצע של מכשיר.
  - בנו רווח סמך ברמת סמך של 95% לאורך החיים הממוצע של מכשיר.
  - הסבירו כיצד ומדוע השתנה רווח הסמך.
- 4) דגמו 200 עובדים מהמשק הישראלי. השכר הממוצע שלהם היה 9700 ₪. נניח שסטיית התקן של השכר במשק היא 3000 ₪.
- בנו רווח סמך ברמת סמך של 95% לתוחלת השכר במשק.
  - מה ניתן לומר בביטחון של 95% על הסטייה המרבית בין ממוצע המדגם לתוחלת השכר?
  - מה היה צריך להיות גודל המדגם אם הינו רוצים להקטין את רווח הסמך ב-50%?
  - אם היינו מגדילים את גודל המדגם ובונים רווח סמך באותה רמת סמך האם היה ניתן לטעון בביטחון רב יותר שרווח הסמך מכיל את הפרמטר?

- (5) בנו רווח סמך לממוצע הציונים של מבחן אינטליגנציה. ידוע שסטיית התקן היא 15 והמדגם מתבסס על 100 תצפיות. רווח הסמך שהתקבל הוא (105,99). שחזרו את:
- ממוצע המדגם.
  - שגיאת האמידה המקסימאלית.
  - רמת הסמך.
- (6) זמן החלמה מאנגינה מתפלג עם סטיית תקן 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 רווח סמך לתוחלת זמן ההחלמה תחת האנטיביוטיקה החדשה ברמת סמך של 90%.
  - מה היה קורה לאורך רווח הסמך אם היה תקציב להגדלת גודל המדגם פי 4? הסבירו.
  - מה היה קורה לאורך רווח הסמך אם היינו בונים את רווח הסמך ברמת סמך גדולה יותר? הסבירו.
- (7) חוקר בנה רווח סמך לממוצע וקיבל את רווח הסמך הבא:  $82 < \mu < 92$ . נתון שסטיית התקן בהתפלגות שווה ל-10 ושהמדגם מתבסס על 16 תצפיות. התפלגות המשתנה היא נורמאלית.
- מהו ממוצע המדגם?
  - מהי רמת הסמך של רווח הסמך שנבנה?
  - מה הסיכוי ששגיאת האמידה באמידת ממוצע האוכלוסייה תעלה על 5?
- (8) חוקר בנה רווח סמך לתוחלת כאשר השונות בהתפלגות ידועה ברמת סמך של 95%. אם החוקר כעת יבנה על סמך אותם נתונים רווח סמך ברמת סמך קטנה מ-95%, איזה מהמשפטים הבאים לא יהיה נכון.
- אורך רווח הסמך החדש יהיה קטן יותר.
  - גודל המדגם יהיה כעת קטן יותר.
  - המרחק בין ממוצע המדגם לקצות רווח הסמך יהיו קטנים יותר ברווח הסמך החדש.
  - רמת הביטחון לבנות רווח הסמך החדש תהיה קטנה יותר.

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

א.  $\mu = 51$ .

ב.  $\bar{X} = 6$ .

ג.  $\bar{X} = 51$ .

ד. אורך רווח הסמך הינו 3.

10) איזה מהגורמים הבאים אינו משפיע על גודלו של רווח בר סמך, כאשר שונות האוכלוסייה ידועה (בחרו בתשובה הנכונה):

א. רמת הביטחון.

ב. סטיית התקן באוכלוסייה.

ג. מספר המשתתפים.

ד. סטיית התקן במדגם.

11) חוקר בנה רווח סמך לממוצע וקיבל את רווח הסמך הבא:  $63 < \mu < 83$ . נתון שסטיית התקן בהתפלגות הייתה ידועה לו ושהמדגם התבסס על 40 תצפיות.

א. אם החוקר היה רוצה לבנות רווח סמך באורך 10.

כמה תצפיות עליו היה לדגום?

ב. רווח הסמך שנבנה על ידי החוקר היה ברמת סמך של 95%.

בנו את רווח הסמך שהיה מתקבל ברמת סמך של 98%.

12) נתון משתנה מקרי רציף מתפלג אחיד:  $X_i \sim U(\mu - 0.5, \mu + 0.5)$ . נרצה לאמוד את  $\mu$ . מצאו רווח סמך ל- $\mu$  ברמת-ביטחון של 0.95 אם במדגם של 45 תצפיות התקבל:  $\bar{x} = 74$ .

(תזכורת על השונות בהתפלגות אחידה רציפה:  $Var(X_i) = \frac{(b-a)^2}{12}$ .)

## תשובות סופיות:

- (1) א. העובדים במשק. ב. שכר ב-ש. ג.  $\mu$ . ד.  $9200 < \mu < 9800$ .  
 ה. 0.95. ו. 600. ז. 0.05.
- (2)  $4920.6 < \mu < 4979.4$
- (3) א.  $223.42 < \mu < 236.58$ . ב.  $222.16 < \mu < 237.84$ .  
 ג. ראה סרטון.
- (4) א.  $10,116 < \mu < 9284$ . ב. הסטיה המירבית בין  $\bar{x}$  ל- $\mu$  היא 416 שם בביטחון של 95%.  
 ג. 800. ד. לא.
- (5) א. 102. ב. 3. ג. 0.9544.
- (6) א.  $4.42 < \mu < 83.5$ . ב. יקטן פי 2. ג. גדל.
- (7) א. 87. ב. 5. ג. 0.9544.
- (8) ב'.
- (9) ג'.
- (10) ד'.
- (11) א. 160. ב.  $61.13 < \mu < 84.87$ .
- (12)  $0.74 \pm 0.084$

## קביעת גודל מדגם:

### רקע:

אם מעוניינים לאמוד את ממוצע האוכלוסייה כאשר סטיית התקן של האוכלוסייה ידועה:  $\sigma$  ברמת סמך של  $1-\alpha$  ושגיאת אמידה שלא תעלה על  $\varepsilon$  מסוים, נציב

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

בנוסחה הבאה:

כדי להציב בנוסחה צריך שהמשתנה הנחקר יתפלג נורמלית או שהמדגם 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 נורמאלית עם סטיית תקן של 2 ק"ג. כמה נוסעים יש לדגום אם מעוניינים שבביטחון של 98% הסטייה המרבית בין ממוצע המדגם לממוצע האמיתי לא יעלה על 0.5 ק"ג? (תשובה: 87).

## שאלות:

- (1) משתנה מקרי מתפלג נורמאלית עם סטיית תקן ידועה 12. מה צריך להיות גודל המדגם כדי לבנות רווח סמך ברמת סמך של 98% שאורכו לא יעלה על 2?
- (2) מעוניינים לאמוד את הדופק הממוצע של מתגייסים לצבא. מעוניינים שבביטחון של 95% שגיאת האמידה המרבית תהיה 0.5. נניח שהדופק מתפלג נורמאלית על סטיית תקן של 3 פעימות לדקה.  
 א. כמה מתגייסים יש לדגום?  
 ב. אם ניקח מדגם הגדול פי 4 מהמדגם של סעיף א ונאמוד את הממוצע באותה רמת סמך כיצד הדבר ישפיע על שגיאת האמידה?
- (3) יהי  $X$  משתנה מקרי עם ממוצע  $\mu$  וסטיית תקן  $\sigma$ . חוקר רוצה לבנות רווח בר סמך ל- $\mu$  ברמת ביטחון של 0.95, כך שהאורך של הרווח יהיה  $0.5\sigma$ . מהו גודל המדגם הנדרש?

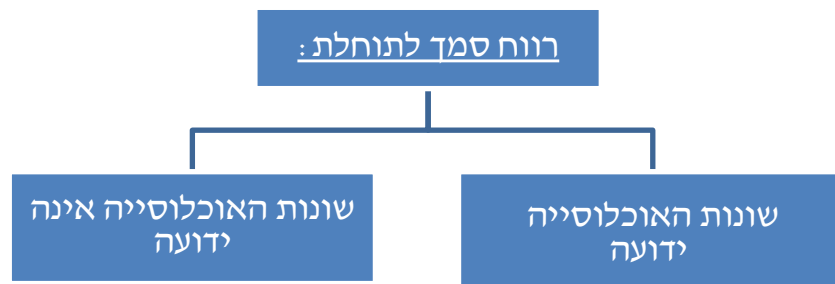
## תשובות סופיות:

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

## רווח סמך כששונות האוכלוסייה לא ידועה:

רקע:

בבואנו לבנות רווח סמך לתוחלת אנו צריכים להתמקד בשני המצבים הבאים:

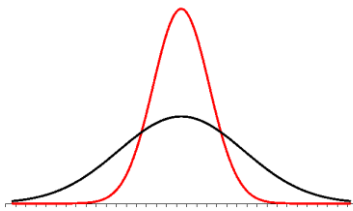


בפרק זה נעסוק במקרה ששונות האוכלוסייה  $(\sigma^2)$  אינה ידועה לנו.

מקרה יותר פרקטי.

**התנאי:**  $X \sim N$  או שהמדגם גדול.

**רווח 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 פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה

להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים.

התפלגות T תלויה במושג שנקרא דרגות חופש. דרגות החופש הן:  $df = n-1$ .

ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה.

כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

**דוגמה (פתרון בהקלטה):**

הזמן שלוקח לפתור שאלה מסוימת בחשבון מתפלג אצל תלמידי כיתות ח' נורמאלית.

במטרה לאמוד את תוחלת זמן הפתרון נדגמו 4 תלמידים בכיתה ח'. להלן התוצאות

שהתקבלו בדקות: 4.7, 5.2, 4.6, 5.3.

בנו רווח סמך ברמת סמך של 95% למוצע זמן הפתרון לשאלה בקרב 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, 88, 84, 79, 89. הערה: לצורך פתרון הנח שקצב פעימות הלב מתפלג נורמאלית בקירוב.
- א. בנו רווח סמך ברמת סמך של 95% לתוחלת הדופק של נוטלי התרופה הנ"ל.  
 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 רווח הסמך ברמת ביטחון של 99%, כיצד הדבר היה משפיע על רווח הסמך?
- (2) במדגם שנעשה על 25 מתגייסים לצבא האמריקאי התקבל כי גובה ממוצע של חייל הינו 178 ס"מ עם סטיית תקן:  $S = 13$  ס"מ. בנו רווח 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 הנסיעה בהם בדקות הוא: 30, 40, 32, 34, 27.
- א. ברמת ביטחון של 95% אמוד את זמן הנסיעה הממוצע. מהי ההנחה הדרושה לצורך פתרון?  
 ב. איך גודל רווח הסמך היה משתנה אם היו דוגמים עוד ימים?
- (4) ציוני מבחן אינטליגנציה מתפלגים נורמאלית. נדגמו 25 מבחנים והתקבל ממוצע ציונים 102 וסטיית תקן מדגמית 13.
- א. בנו רווח 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 הסעיפים הנ"ל.
- (5) נשקלו 60 תינוקות אשר נולדו בשבוע ה-40 של ההיריון. המשקל נמדד בקילוגרמים. להלן התוצאות שהתקבלו:  $\sum_{i=1}^{60} X_i = 195$ ,  $\sum_{i=1}^{60} X_i^2 = 643.19$ .
- בנו רווח סמך ברמת סמך של 95% לתוחלת משקל תינוק ביום היוולדו.

- (6) נדגמו 120 אנשים אקראיים מעל גיל 50. עבור כל אדם נבדק מספר שנות השכלתו. להלן התוצאות שהתקבלו:  $\bar{x} = 13.8$ ,  $S = 2$ . בנו רווח סמך ברמת סמך של 96% לממוצע ההשכלה של אזרחים מעל גיל 50.
- (7) שני סטטיסטיקאים בנו רווח בר-סמך לאותו פרמטר  $\mu$ . לכל אחד מהסטטיסטיקאים מדגם אחר, אך באותו גודל 10. שניהם קבעו אותה רמת סמך. סטטיסטיקאי א': הניח  $\sigma = 20$ . סטטיסטיקאי ב': חישב לפי המדגם וקיבל  $S = 20$ . למי משני הסטטיסטיקאים יהיה רווח סמך ארוך יותר?  
 א. סטטיסטיקאי א'.  
 ב. סטטיסטיקאי ב'.  
 ג. אותו אורך רווח סמך לשני הסטטיסטיקאים.  
 ד. תלוי בתוצאות המדגם של כל סטטיסטיקאי.
- (8) נתון ש:  $X \sim N(\mu, \sigma^2)$  ביצעו מדגם בגודל 16 וקיבלו סטיית תקן מדגמית 10. אורך רווח הסמך שהתקבל הוא: 8.765. מהי רמת הביטחון של רווח הסמך?

### תשובות סופיות:

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

# סטטיסטיקה והסתברות 01 88165

פרק 63 - רווח סמך להפרש תוחלות (ממוצעים) במדגמים בלתי תלויים

תוכן העניינים

1. כששוניות האוכלוסיה ידועות.....314
2. כששוניות האוכלוסיה לא ידועות ובהנחת שוויון שוניות.....316

## כששונויות האוכלוסייה ידועות:

### רקע:

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

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

התנאים לבניית רווח הסמך:

1.  $\sigma_1^2, \sigma_2^2$  ידועות.

2.  $X_1, X_2 \sim N$  או  $n_1, n_2 > 30$ .

3. שני מדגמים בלתי תלויים.

רווח סמך:  $(\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

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

### דוגמה (פתרון בהקלטה):

נדגמו 100 תושבים מאזור A והמשכורת הממוצעת הייתה שם 9200 ₪. כמו כן נדגמו 120 תושבים מאזור B וממוצע המשכורות שהתקבל שם 8700 ₪. לצורך פתרון נניח שסטיית התקן של המשכורות באוכלוסיית שני האזורים היא 1800 ₪. אמדו ברמת סמך של 90% את הפרש השכר הממוצע בין אזור A לאזור B.

## שאלות:

- (1) מעוניינים לבדוק האם קיים הבדל בין ממוצע ציוני הפסיכומטרי של חיילים לממוצע ציוני הפסיכומטרי של תלמידי תיכון. ידוע שציוני הפסיכומטרי מתפלגים נורמאלית עם סטיית תקן 100. במדגם של 16 נבחנים חיילים התקבל ממוצע 543. במדגם של 20 תלמידי תיכון התקבל ממוצע 508. בנו רווח סמך לפער תוחלות הציונים בין חיילים לתלמידי תיכון ברמת סמך של 90%. מה ניתן להסיק מרווח סמך זה?
- (2) ציוני IQ מתוכננים כך שיתפלגו נורמאלית עם סטיית תקן של 15. במדגם של 20 נבחנים ישראלים התקבל ממוצע ציונים 104. במדגם של 23 נבחנים אמריקאיים התקבל ממוצע ציונים 99.  
א. בנו רווח סמך ברמת סמך של 95% לפער בין ישראל לארה"ב בממוצע הציונים במבחן ה-IQ.  
ב. האם קיים הבדל בין ישראלים לאמריקאים מבחינת ממוצע הציונים?
- (3) חברה להנדסת בניין מעוניינת להשוות ברמת הקשיות של שני סוגי ברגים. ידוע שרמת הקשיות של ברגים מתפלגת נורמלית עם סטיית תקן של 4 יחידות. במדגם של 15 ברגים מסוג א' התקבל רמת קשיות ממוצעת של 28 יחידות ובמדגם של 12 ברגים מסוג ב' התקבל רמת קשיות ממוצעת של 25. עבור אילו רמות בטחון יקבע שאין הבדל בין שני סוגי הברגים מבחינת ממוצע רמת הקשיות שלהם?

## תשובות סופיות:

- (1)  $(-20, 90)$ .
- (2) א.  $-3.99 < \mu_1 - \mu_2 < 13.99$ .  
ב. לא נוכל לטעון בביטחון של 95% שקיים הבדל בין ישראל לארה"ב.  
ג. רמות בטחון הגבוהות מ-0.9476.
- (3)

## כששונויות האוכלוסייה לא ידועות ובהנחת שוויון שונויות:

### רקע:

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

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

התנאים לבניית רווח הסמך:

$$1. \sigma_1^2 = \sigma_2^2$$

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

3. מדגמים בלתי תלויים.

**השונויות המשוקללת:** כיוון שאנו מניחים שבין שתי האוכלוסיות השונויות שוות אנו אומדים את השונויות הזו על ידי שקלול שתי השונויות של שני המדגמים על ידי

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

הנוסחה הבאה:

$$d.f = n_1 + n_2 - 2$$

דרגות החופש:

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

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

### דוגמה (פתרון בהקלטה):

מחקר מעוניין לבדוק האם קיים הבדל בין תל אביב לבאר שבע מבחינת ההכנסה הממוצעת של אקדמאים. להלן תוצאות המדגם שנעשה:

באר שבע	תל אביב	
10	20	מספר האקדמאים
9500	11,000	ממוצע הכנסות של אקדמאים
250	200	סטיית התקן של הכנסות אקדמאים

בנו רווח סמך ברמת ביטחון של 90% להפרש תוחלות ההכנסה בשני האזורים. הניחו שהשכר מתפלג נורמלית עם אותה שונות בכל אחד מהאזורים.

## שאלות:

- (1) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו למבחן IQ. להלן תוצאות המדגם:

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

מצאו רווח סמך ברמת סמך של 95% לסטייה בין ממוצע הציונים בישראל לממוצע הציונים בארה"ב. רשמו את כל ההנחות הדרושות לצורך פתרון התרגיל.

- (2) להלן 4 תצפיות 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 רווח סמך ל- $\mu_y - \mu_x$  ברמת הסמך 90%, בהנחה ששני המדגמים בלתי תלויים.

## תשובות סופיות:

- (1) הנחות:
- השונויות שווה.
  - שהציונים מתפלגים נורמלית.
  - המדגמים אינם תלויים זה בזה.
- $$-5.52 < \mu_1 - \mu_2 < 17.52$$
- (2)
- $$-9.6 < \mu_y - \mu_x < 1.6$$

# סטטיסטיקה והסתברות 01 88165

פרק 64 - רווח סמך לשונות וסטיית תקן

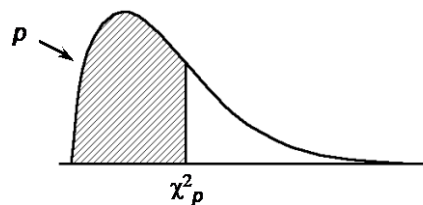
תוכן העניינים

1. רווח סמך לשונות וסטיית תקן..... 318

## רווח סמך לשונות וסטיית תקן:

### רקע:

בפרק זה נדון על בניית רווח סמך לשונות האוכלוסייה. התנאי לבניית רווח הסמך: המשתנה הנחקר מתפלג נורמלית, למרות שנהוג לא לדרוש את התנאי הזה אם המדגם מספיק גדול. רווח הסמך יתבסס על התפלגות הנקראת חי בריבוע. התפלגות זו היא התפלגות אסימטרית חיובית המתחילה מהערך אפס ותלויה בדרגות חופש. דרגות החופש במקרה זה יהיו:  $n-1$ .



$$\text{רווח הסמך לשונות: } \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}}$$

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

אם נרצה לבנות רווח סמך לסטיית תקן אז נוציא שורש לרווח סמך לשונות.

### דוגמה:

זמן התגובה מתפלג נורמלית. במטרה לאמוד את שונות זמן התגובה נדגמו 4 תצפיות. להלן התוצאות בשניות: 4.7, 5.2, 4.6, 5.3. בנו רווח סמך, ברמת סמך של 95%, לשונות זמן התגובה באוכלוסייה.

פתרון:

פרמטר:  $\sigma^2$ .

$$X \sim N(\mu, \sigma^2) = \text{זמן תגובה (בשניות)}$$

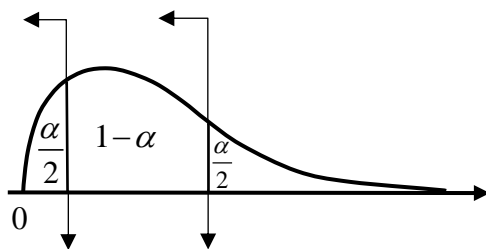
תוצאות מדגם:  $n = 4$ .

$$\bar{X} = \frac{4.7+5.2+4.6+5.3}{4} = 4.95$$

$$d.f = n-1 = 4-1 = 3$$

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

$$S^2 = \frac{4.7^2 + 5.2^2 + \dots - 4 \cdot 4.95^2}{4-1} = 0.123$$



$$X^2_{0.025} = 0.216 \quad X^2_{0.975} = 9.35$$

$$1 - \alpha = 0.95$$

$$\alpha = 0.05$$

$$\frac{\alpha}{2} = 0.025$$

(טבלת התפלגות חי-בריבוע מופיעה בעמוד האחרון).

$$\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}} \quad \text{נציב:}$$

$$\frac{(4-1) \cdot 0.123}{9.35} < \sigma^2 < \frac{(4-1) \cdot 0.123}{0.216}$$

$$0.039 < \sigma^2 < 1.708$$

## שאלות:

(1) חמישה מטופלים קבלו תרופה מסוימת. בדקו לכל מטופל את זמני התגובה שלו. להלן הזמנים שהתקבלו בדקות: 18, 17, 21, 26, 28. בהנחה וזמני התגובה מתפלגים נורמאלית, בנו רווח סמך ברמת סמך של 95% לשונות זמן התגובה.

(2) נדגמו 20 ימים אקראיים מחודשי יולי-אוגוסט ונמדדה בהם הטמפי' במעלות צלזיוס בת"א. במדגם התקבל טמפי' ממוצעת 30.8 וסטיית תקן מדגמית 1.1. בהנחה והטמפי' מתפלגת נורמאלית:  
 א. בנו רווח סמך לתוחלת הטמפי' בחודשים אלה בת"א ברמת סמך של 95%.  
 ב. בנו רווח סמך לסטיית התקן של הטמפי' בחודשים אלה בת"א ברמת סמך של 95%.

(3) ציוני IQ בארה"ב מתפלגים נורמאלית עם ממוצע 100 וסטיית תקן 15. נבחנו 20 נבחנים ישראלים במבחן ה-IQ.

$$\sum_{i=1}^{20} X_i = 2080, \quad \sum_{i=1}^{20} X_i^2 = 218,220$$

להלן התוצאות שהתקבלו:

נניח שגם בישראל הציונים מתפלגים נורמאלית.

- א. מצאו אומדנים לממוצע הציונים בישראל ולשונות הציונים בישראל באמצעות אומדנים חסרי הטיה.  
 ב. אמדו ברמת ביטחון של 95% את תוחלת הציונים של נבחנים בישראל.  
 ג. אמדו ברמת סמך של 90% את סטיית התקן של הציונים של נבחנים ישראלים.  
 ד. על סמך הסעיפים הקודמים, האם בישראל ממוצע הציונים וסטיית התקן של הציונים שונה מבארה"ב? הסבירו.

(4) באוכלוסייה מסוימת נדגמו 10 תצפיות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{10} X_i = 750, \quad \sum_{i=1}^{10} (X_i - \bar{X})^2 = 900$$

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

נתון ש:

- א. בנו רווח סמך ל- $\mu$  ברמת סמך של 95%.  
 ב. בנו רווח סמך ל- $\sigma^2$  ברמת סמך של 95%.

### תשובות סופיות:

$$(1) \quad .8.4 < \sigma^2 < 194.2$$

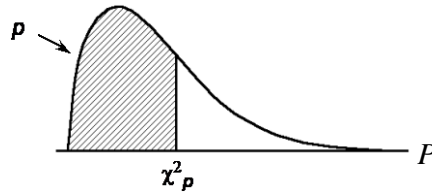
$$(2) \quad \text{א. } .30.285 < \mu < 31.315 \quad \text{ב. } .0.836 < \sigma < 1.606$$

$$(3) \quad \text{א. ממוצע: } 104, \text{ שונות: } 100. \quad \text{ב. } .99.32 \leq \mu \leq 108.68 \quad \text{ג. } .7.94 < \sigma < 13.7$$

ד. בביטחון של 95% ממוצע הציונים איננו שונה, ובביטחון של 90% סטיית התקן שונה.

$$(4) \quad \text{א. } .68.75 < \mu < 82.15 \quad \text{ב. } .47.4 < \sigma^2 < 333.3$$

נספח - טבלת התפלגות חי-בריבוע – ערכי החלוקה  $\chi^2_p$  :



df	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995
1	0.004393	0.005399	0.008391	0.013445	0.020001	0.031793	0.053993	0.101350	0.157055	0.214770	0.270009	0.317706	0.374544
2	0.010000	0.020000	0.050000	0.103000	0.211000	0.575000	1.390000	2.770000	4.610000	5.990000	7.380000	9.210000	10.590000
3	0.071717	0.115000	0.216000	0.352000	0.584000	1.210000	2.370000	4.110000	6.250000	7.810000	9.350000	11.300000	12.830000
4	0.207000	0.297000	0.484000	0.711000	1.060000	1.920000	3.360000	5.390000	7.780000	9.490000	11.100000	13.300000	14.900000
5	0.412000	0.554000	0.831000	1.150000	1.610000	2.670000	4.350000	6.630000	9.240000	11.100000	12.800000	15.100000	16.700000
6	0.676000	0.872000	1.240000	1.640000	2.200000	3.450000	5.350000	7.840000	10.600000	12.600000	14.400000	16.800000	18.500000
7	0.989000	1.240000	1.690000	2.170000	2.830000	4.250000	6.350000	9.040000	12.000000	14.100000	16.000000	18.500000	20.300000
8	1.340000	1.650000	2.180000	2.730000	3.490000	5.070000	7.340000	10.200000	13.400000	15.500000	17.500000	20.100000	22.000000
9	1.730000	2.090000	2.700000	3.330000	4.170000	5.900000	8.340000	11.400000	14.700000	16.900000	19.000000	21.700000	23.600000
10	2.160000	2.560000	3.250000	3.940000	4.870000	6.740000	9.340000	12.500000	16.000000	18.300000	20.500000	23.200000	25.200000
11	2.600000	3.050000	3.820000	4.570000	5.580000	7.580000	10.300000	13.700000	17.300000	19.700000	21.900000	24.700000	26.800000
12	3.070000	3.570000	4.400000	5.230000	6.300000	8.440000	11.300000	14.800000	18.500000	21.000000	23.300000	26.200000	28.300000
13	3.570000	4.110000	5.010000	5.890000	7.040000	9.300000	12.300000	16.000000	19.800000	22.400000	24.700000	27.700000	29.800000
14	4.070000	4.660000	5.630000	6.570000	7.790000	10.200000	13.300000	17.100000	21.100000	23.700000	26.100000	29.100000	31.300000
15	4.600000	5.230000	6.260000	7.260000	8.550000	11.000000	14.300000	18.200000	22.300000	25.000000	27.500000	30.600000	32.800000
16	5.140000	5.810000	6.910000	7.960000	9.310000	11.900000	15.300000	19.400000	23.500000	26.300000	28.800000	32.000000	34.300000
17	5.700000	6.410000	7.560000	8.670000	10.100000	12.800000	16.300000	20.500000	24.800000	27.600000	30.200000	33.400000	35.700000
18	6.260000	7.010000	8.230000	9.390000	10.900000	13.700000	17.300000	21.600000	26.000000	28.900000	31.500000	34.800000	37.200000
19	6.840000	7.630000	8.910000	10.100000	11.700000	14.600000	18.300000	22.700000	27.200000	30.100000	32.900000	36.200000	38.600000
20	7.430000	8.260000	9.590000	10.900000	12.400000	15.500000	19.300000	23.800000	28.400000	31.400000	34.200000	37.600000	40.000000
21	8.030000	8.900000	10.300000	11.600000	13.200000	16.300000	20.300000	24.900000	29.600000	32.700000	35.500000	38.900000	41.400000
22	8.640000	9.540000	11.000000	12.300000	14.000000	17.200000	21.300000	26.000000	30.800000	33.900000	36.800000	40.300000	42.800000
23	9.260000	10.200000	11.700000	13.100000	14.800000	18.100000	22.300000	27.100000	32.000000	35.200000	38.100000	41.600000	44.200000
24	9.890000	10.900000	12.400000	13.800000	15.700000	19.000000	23.300000	28.200000	33.200000	36.400000	39.400000	43.000000	45.600000
25	10.500000	11.500000	13.100000	14.600000	16.500000	19.900000	24.300000	29.300000	34.400000	37.700000	40.600000	44.300000	46.900000
26	11.200000	12.200000	13.800000	15.400000	17.300000	20.800000	25.300000	30.400000	35.600000	38.900000	41.900000	45.600000	48.300000
27	11.800000	12.900000	14.600000	16.200000	18.100000	21.700000	26.300000	31.500000	36.700000	40.100000	43.200000	47.000000	49.600000
28	12.500000	13.600000	15.300000	16.900000	18.900000	22.700000	27.300000	32.600000	37.900000	41.300000	44.500000	48.300000	51.000000
29	13.100000	14.300000	16.000000	17.700000	19.800000	23.600000	28.300000	33.700000	39.100000	42.600000	45.700000	49.600000	52.300000
30	13.800000	15.000000	16.800000	18.500000	20.600000	24.500000	29.300000	34.800000	40.300000	43.800000	47.000000	50.900000	53.700000

# סטטיסטיקה והסתברות 01 88165

פרק 65 - בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן)

תוכן העניינים

1. בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן)..... 323

## בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן):

### רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטיקה. בתהליך זה ישנן שתי השערות שנבדקות:

1. השערת האפס: המסומנות ב- $H_0$ .
2. השערה אלטרנטיבית (השערת המחקר): המסומנת ב- $H_1$ .

בדרך כלל השערת האפס מסמנת את אשר היה מקובל עד עכשיו, את השגרה הנורמה ואילו ההשערה האלטרנטיבית את החדשנות בעצם ההשערה האלטרנטיבית מדברת על הסיבה שהמחקר נעשה.

### דוגמה:

ישנה תרופה קיימת למחלה A אשר גורמת ל-10% מהמשתמשים בה לתופעות לוואי. חברת תרופות טוענת שפיתחה תרופה שיעילה באותה מידה, אך מקטינה את הסיכוי לתופעות הלוואי. לכן יש לבצע מחקר שעל סמך תוצאותיו ננסה להכריע איזה השערה נקבל:

$H_0$ : התרופה החדשה הנה קונבנציונאלית וגורמת ל-10% תופעות לוואי.

$H_1$ : התרופה החדשה מקטינה את אחוז הסובלים מתופעות לוואי מתחת ל-10%.

בתהליך של בדיקת השערות יוצרים כלל שנקרא כלל הכרעה. הכלל יוצר אזורים:

1. אזור דחייה: דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.
2. אזור קבלה: קבלה של השערת האפס ודחייה של האלטרנטיבה.

כלל ההכרעה מתבסס על איזשהו סטטיסטי. בתהליך יש ללכת לתוצאות המדגם ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה וכך להגיע למסקנה. המסקנה היא בעירבון מוגבל כיוון שהיא תלויה בכלל ההכרעה ובתוצאות המדגם. אם נשנה את כלל ההכרעה אז אנחנו יכולים לקבל מסקנה אחרת, אם נבצע מדגם חדש אז אנחנו עלולים לקבל תוצאה אחרת.

לכן יתכנו טעויות במסקנות שלנו :

		הכרעה	
		$H_0$	$H_1$
מציאות	$H_0$	אין טעות	טעות מסוג 1
	$H_1$	טעות מסוג 2	אין טעות

### הגדרת הטעויות:

טעות מסוג ראשון: להכריע לדחות את  $H_0$  למרות שבמציאות  $H_0$  נכונה.

טעות מסוג שני: להכריע לקבל את  $H_0$  למרות שבמציאות  $H_1$  נכונה.

### הגדרת הסתברויות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):

$$\alpha = P(H_0 \text{ לדחות את } H_0 \mid H_0 \text{ נכונה}) = P_{H_0}(H_0)$$

הסיכוי לבצע טעות מסוג 2:

$$\beta = P(H_0 \text{ לקבל את } H_0 \mid H_1 \text{ נכונה}) = P_{H_1}(H_0)$$

רמת בטחון:

$$(\alpha - 1) = P(H_0 \text{ לקבל את } H_0 \mid H_0 \text{ נכונה}) = P_{H_0}(H_0)$$

עוצמה:

$$(\beta - 1) = \pi = P(H_0 \text{ לדחות את } H_0 \mid H_1 \text{ נכונה}) = P_{H_1}(H_0)$$

### דוגמה (פתרון בהקלטה):

בכד יש 10 כדורים. יתכן ש-5 מהם לבנים והיתר שחורים (כד א' – השערת האפס)

או ש-7 מהם לבנים והיתר שחורים (כד ב' – השערה אלטרנטיבית).

כדי להחליט איזה מהכדים ברשותנו, הוחלט להוציא כדור ולהשתמש בכלל

ההחלטה הבא: אם הכדור שהוצא הוא לבן שזהו כד ב'  $H_1$ .

א. חשבו את רמת המובהקות ואת רמת הביטחון של המבחן המוצע.

ב. חשבו את הסיכוי לטעות מסוג שני והעוצמה של המבחן המוצע.

## שאלות:

- (1) אדם חשוד בביצוע פשע. מהן הטעויות האפשריות בהכרעת הדין?
- (2) ילד קנה שקית סוכריות אטומה שבה ציפה ל-10 סוכריות תות ו-5 לימון. ישנה שקית אחרת אותה הוא לא רצה בה 6 סוכריות תות ו-9 לימון. הוא החליט להוציא באקראי סוכרייה, אם היא תהיה לימון הוא יחזיר את השקית לחנות. מה הסיכויים לכל סוג של טעות בהכרעתו?
- (3) יהי  $X$  מספר שלם הנבחר באקראי מבין המספרים השלמים. הסיכוי ש- $X$  יקבל ערך כלשהו נתון על ידי הנוסחה:  $p(X = k) = \frac{1}{n}$  עבור:  $k = 1, 2, \dots, n$ . נתונות ההשערות הבאות לגבי התפלגות של  $X$ :  $H_0: n = 4$ ,  $H_1: n = 6$ . כמו כן נתון כלל ההכרעה הבא: נדחה את השערת האפס אם:  $X > 3$ . חשבו את הסיכוי לטעות מסוג ראשון וטעות מסוג שני ואת העוצמה?
- (4) איכות של מוצר מסווגת ל-4 רמות איכות: מצוין, טוב, בינוני וירוד. להלן התפלגות טיב המוצר בשני מפעלים:

מפעל / איכות	מצוין	טוב	בינוני	ירוד
"היוצר"	0.6	0.2	0.2	0
"שמשון"	0.1	0.2	0.3	0.4

- בוחרים ממשלוח מוצר באקראי, אך לא יודעים מאיזה מפעל המשלוח הגיע. על סמך בדיקת האיכות מנסים להכריע האם מדובר במפעל "היוצר" (השערת האפס) או במפעל "שמשון" (השערה אלטרנטיבית).
- א. להלן כלל החלטה: אם מדובר במוצר שטיבו "טוב" נכריע שהמוצר בא ממפעל "שמשון", מהן ההסתברויות לסוגי הטעויות השונים?
- ב. להלן כלל החלטה: אם מדובר במוצר שטיבו "בינוני" או גרוע מכך נכריע שהמוצר בא ממפעל "שמשון", מהן ההסתברויות לסוגי הטעויות השונים?
- ג. איזה כלל החלטה עדיף? נמקו!
- (5) במטרה לבדוק האם מטבע תקין הטילו אותו 8 פעמים. הוחלט שאם מספר העצים יהיה בין 1 ל-7 כולל יוחלט שהמטבע תקין, אחרת נחליט שהמטבע מזויף.
- א. רשמו את השערות המחקר.
- ב. מה ההסתברות לטעות מסוג ראשון?
- ג. מהי עצמת המבחן אם במציאות אכן המטבע אינו תקין כי הסיכוי לעץ בו הוא 20%.

6) להלן השערות:

$H_0: X \sim t(5)$  - התפלגות T עם חמש דרגות חופש.

$H_1: X \sim Z$  - התפלגות נורמלית.

כלל החלטה: נדחה את השערת האפס אם  $X$  גדול מ-2.015.

א. מהי רמת המובהקות של כלל החלטה?

ב. מהי העוצמה של כלל החלטה?

7)

במפעל מסוים נפלטים לאוויר חומרים רעילים. במצב שיגרה העוצמה הממוצעת של החומר הרעיל אמורה להיות 6,000 יחידות עם סטיית תקן 900. במצב חירום העוצמה הממוצעת היא 7,000 עם סטיית תקן 900. במפעל מערכת התראה נתמכת על ידי 9 חיישנים. אם ממוצע העוצמה של החומר הרעיל לפי תשעת החיישנים עולה על 6,600 יחידות מופעלת מערכת ההתראה. נתון שעוצמת הזיהום מתפלגת נורמאלית.

א. מה הסיכוי להתראת שווא? (באיזה סוג טעות מדובר)?

ב. מה הסיכוי שבמצב חירום מערכת ההתראה לא תפעל? (באיזה סוג טעות מדובר)?

ג. מה ההסתברות שאם המצב הוא מצב חירום מערכת ההתראה תפעל? (איך קוראים להסתברות זו)?

ד. בסעיפים הבאים נשנה בכל סעיף נתון מסוים. כל סעיף עומד בפני עצמו, כיצד השינוי ישנה את הסיכוי לטעות מסוג ראשון ושני?

i. המפעל יקנה עוד 4 חיישנים.

ii. מצב חרום מוגדר כעת בתוחלת של 7,500 יחידות.

iii. מערכת ההתראה תופעל אם ממוצע של תשעת החיישנים יהיה מעל 6,700.

8)

במטרה לבדוק האם במקום עבודה מסוים פרופורציית הבנים נמוכה מפרופורציית הבנות נדגמו באקראי 10 עובדים. הוחלט שאם מספר הבנים במדגם יהיה לכל היותר 2 תתקבל הטענה שפרופורציית הבנים נמוכה מפרופורציית הבנות.

א. מה רמת המובהקות של כלל ההכרעה הני"ל?

ב. מהי העוצמה בהנחה ובחברה 30% בנים?

- 9) זמן ההשפעה של משכך הכאבים "אופטלנוס" מתפלג נורמאלית עם תוחלת של 40 דקות וסטיית תקן של 12 דקות. חברת התרופות המייצרת את התרופה מנסה לשפר את התרופה כך שתוחלת הזמן עד להשפעה תתקצר. לצורך כך, דגמו 25 מטופלים שיקבלו את התרופה "אופטלנוס פורטה", ממוצע זמן התגובה של המטופלים היה 34.5 דקות. חברת התרופות החליטה מראש שאם ממוצע הזמן עד להשפעה יהיה נמוך מ-35 דקות, היא תמשיך בתהליך שיווק "אופטלנוס פורטה".
- א. מהי רמת המובהקות של המבחן המוצע?  
 ב. על סמך תוצאות המדגם, מהי המסקנה ומהי הטעות האפשרית במסקנה?  
 ג. מהי עצמת המבחן המוצע אם במציאות התרופה "אופטלנוס פורטה" מפחיתה את התוחלת לכדי 32 דקות?  
 ד. כיצד תשתנה התשובה לסעיף ג' אם החברה הייתה מחליטה שהיא תמשיך בתהליך שיווק התרופה החדשה כאשר ממוצע המדגם יהיה נמוך מ-36 דקות?
- 10) ציוני פסיכומטרי מתפלגים נורמלית עם סטיית תקן 120. מכון טוען שלימודים אצלו מעלים את ממוצע הציונים ביותר מ-30 נקודות. נלקחו 20 שלמדו במכון ו-20 שניגשו לבחינה בלמידה עצמית. הוחלט במשרד פרסום לקבל את טענת המכון רק אם במדגם ממוצע הציונים של אלה שלמדו במכון יהיה גבוהה בלפחות 50 נקודות מאלה שלא היו.
- א. מהי רמת המובהקות של המחקר?  
 ב. מה הסיכוי לעשות טעות מסוג שני II בהנחה שהמכון מעלה את ממוצע הציונים ב-60 נקודות?  
 ג. כיצד התשובות לסעיף א ו ב' היו משתנות אם מסתבר שסטיית התקן בציוני הפסיכומטרי הינה 100. הסבירו ללא חישוב.
- 11) קו ייצור נחשב תקין אם יש בו לכל היותר 4% פגומים, ונחשב שאינו תקין אחרת. מנהל האיכות דוגם בכל יום מקו הייצור 500 מוצרים. אם במדגם יהיה לפחות 30 מוצרים פגומים יפסיקו באותו היום את קו הייצור.
- א. מה ההסתברות להפסיק את קו הייצור כשהוא תקין. איך קוראים להסתברות זאת?  
 ב. מה ההסתברות להמשיך ביום מסוים את קו הייצור למרות שאינו תקין כי היו 8% פגומים בקו הייצור. איך קוראים להסתברות זאת?
- 12) מעוניינים לבדוק האם בפקולטה מסוימת ישנה העדפה לגברים. הוחלט לדגום 200 מתקבלים ועל סמך מספר הבנים לקבוע אם טענת המחקר מתקבלת. חוקר א' קבע רמת מובהקות של 5% וחוקר ב' החליט לקבל את טענת המחקר אם במדגם יהיו לפחות 120 בנים. למי מבין החוקרים רמת מובהקות גדולה יותר?

**13** מספר המכוניות הנכנסות לחניון "עזרים" מתפלג פואסונית. בשנה שעברה המכוניות נכנסו לחניון בקצב של 2 מכוניות לדקה. בעקבות תלונות על עומס יתר בכניסה לחניון מעוניין מנהל החניון לבדוק האם קצב כניסת המכוניות לחניון גדל השנה. מנהל החניון החליט לספור את מספר המכוניות שיכנסו לחניון בדקה אקראית. אם מספר המכוניות שיספרו יהיה לפחות 4 יפתח מנהל החניון שער נוסף לחניון.

- א. רשמו את השערות מנהל החניון ואת כלל ההחלטה שלו. האם כלל ההכרעה הגיוני?
- ב. מהי רמת המובהקות של כלל ההכרעה?
- ג. מהי העוצמה של כלל ההחלטה, אם כיום קצב כניסת המכוניות לחניון גדל ל-4 מכוניות בדקה?

**14** עובד עובד במפעל שבו מתחילים לעבוד בשעה 8:00. עובד בדרך כלל מאחר לעבודה והמנהל החליט לרשום את שעת הגעתו. המנהל טוען שמשך האיחור של עובד (בדקות),  $X$ , הוא משתנה אחיד  $U(0, 60)$ . עובד טוען שהוא לא מגיע באיחור כה גדול, אלא שהתפלגות  $X$  היא בעלת התפלגות מעריכית עם תוחלת איחור של 20 דקות.

לבדיקת טענת המנהל ( $H_0$ ) כנגד טענת עובד ( $H_1$ ), המבוסס על משך האיחור של חגי ביום אחד. מוצאים שני ככלי הכרעה:

- כלל 1: דחה את השערת האפס אם משך האיחור יהיה לפחות 40 דקות.
- כלל 2: דחה את השערת האפס אם משך האיחור יהיה לכל היותר 20 דקות.

חשבו את הסיכוי לטעות מסוג ראשון ושני לכל אחת מכללי ההכרעה. מי עדיף?

## תשובות סופיות:

- (1) ראה סרטון וידאו.
- (2)  $\beta = \frac{2}{5}$ ,  $\alpha = \frac{1}{3}$
- (3)  $\beta = 0.5$ ,  $\alpha = 0.25$
- (4) א.  $\beta = 0.8$ ,  $\alpha = 0.2$
- (5) א. השערות:  $H_0$  - מטבע תקין.  
 $H_1$  - מטבע לא תקין.
- (6) א. 0.05. ב. 0.022.
- (7) א. 0.0228. ב. 0.0918. ג. 0.9082. ד. i.  $\alpha, \beta$  יקטנו.  
 ii.  $\alpha$  לא משתנה,  $\beta$  קטנה.  
 iii.  $\alpha$  קטנה,  $\beta$  גדלה.
- (8) א. 0.055. ב. 0.383.
- (9) א. 0.0188. ב. טעות מסוג I. ג. 0.8944. ד. העוצמה תגדל.
- (10) א. 0.2981. ב. 0.3974. ג. קטן.
- (11) א. 0.0113. ב. 0.0495.
- (12) חוקר א'.
- (13) א. ראה סרטון וידאו. ב. 0.1428. ג. 0.566.
- (14) להלן טבלת טעויות, ממנה ניתן להסיק שכלל 2 עדיף.

$\beta$	$\alpha$	כלל
0.865	$\frac{1}{3}$	1
0.368	$\frac{1}{3}$	2

# סטטיסטיקה והסתברות 01 88165

פרק 66 - מבוא לבדיקת השערות על פרמטרים

תוכן העניינים

330	.....	1. הקדמה
334	.....	2. סוגי טעויות

## הקדמה:

### רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטיקה. בבדיקת השערות על פרמטרים נעבוד לפי השלבים הבאים:

**שלב א:** נוהה את הפרמטר הנחקר.

**שלב ב:** נרשום את השערות המחקר.

השערת האפס המסומנות ב- $H_0$ .

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

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

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

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

**שלב ד:** נרשום את כלל ההכרעה. בתהליך של בדיקת השערות יוצרים כלל שנקרא כלל הכרעה. הכלל יוצר אזורי שנקראים:

1. **אזור דחייה:**

דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.

2. **אזור קבלה:**

קבלה של השערת האפס ודחייה של האלטרנטיבה. כלל ההכרעה מתבסס על איזשהו סטטיסטי. אזור הדחייה מוכתב על ידי סיכון שלוקח החוקר מראש

שנקרא רמת מובהקות ומסומן ב- $\alpha$ .

**שלב ה:** בתהליך יש ללכת לתוצאות המדגם ולחשב את הסטטיסטי המתאים ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה.

**שלב ו:** להסיק מסקנה בהתאם לתוצאות המדגם.

**דוגמה (פתרון בהקלטה):**

משרד הבריאות פרסם שמשקל ממוצע של תינוקות ביום לידתם בישראל 3300 גרם. משרד הבריאות רוצה לחקור את הטענה שנשים מעשנות בזמן ההיריון יולדות תינוקות במשקל נמוך מהממוצע. במחקר השתתפו 20 נשים מעשנות בהריון. להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:

$$n = 20, \bar{X} = 3120, S = 280$$

- א. מהי אוכלוסיית המחקר?
- ב. מה המשתנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

## שאלות:

בשאלות הבאות, ענו על הסעיפים הבאים:

- א. מהי אוכלוסיית המחקר?
- ב. מה המשתנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

- (1) ממוצע הציונים בבחינת 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 אקראיים. ממוצע הציונים של אותם תלמידים לאחר שלמדו בשיטתו היה 75.5.
- (2) לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 סמ"ק במדגם בגודל 25.
- (3) במשך שנים אחוז המועמדים שהתקבל לפקולטה למשפטים היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. מחקר מעוניין לבדוק האם השנה מקשים על הקבלה לפקולטה למשפטים.
- (4) בחודש ינואר השנה פורסם שאחוז האבטלה במשק הוא 8% במדגם עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים. רוצים לבדוק ברמת מובהקות של 5% האם כיום אחוז האבטלה הוא כמו בתחילת השנה.

### תשובות סופיות:

- (1) א. נבחנים בבגרות באנגלית.  
 ב. ציון.  
 ג. ממוצע הציונים בשיטת לימוד חדשה.  
 ד.  $H_0: \mu = 72$   
 $H_1: \mu > 72$
- (2) א. משקאות בבקבוק של חברה מסוימת.  
 ב. נפח משקה בסמ"ק.  
 ג. ממוצע נפח המשקה בבקבוק.  
 ד.  $H_0: \mu = 500$   
 $H_1: \mu < 500$
- (3) א. מועמדים לפקולטה למשפטים.  
 ב. משתנה דיכוטומי (התקבל, לא התקבל).  
 ג. אחוז הקבלה.  
 ד.  $H_0: p = 0.25$   
 $H_1: p < 0.25$
- (4) א. אזרחים בוגרים במשק.  
 ב. משתנה דיכוטומי (מובטל, עובד).  
 ג. אחוז האבטלה כיום.  
 ד.  $H_0: p = 0.08$   
 $H_1: p \neq 0.08$

## סוגי טעויות:

### רקע:

בתהליך של בדיקת השערות יוצרים כלל שניקרא כלל הכרעה. הכלל יוצר אזורים שנקראים:

1. אזור דחייה – דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.
2. אזור קבלה – קבלה של השערת האפס ודחייה של האלטרנטיבה.

כלל ההכרעה מתבסס על איזשהו סטטיסטי. בתהליך יש ללכת לתוצאות המדגם ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה וכך להגיע למסקנה – המסקנה היא בעירבון מוגבל כיוון שהיא תלויה בכלל ההכרעה ובתוצאות המדגם. אם נשנה את כלל ההכרעה אז אנחנו יכולים לקבל מסקנה אחרת. אם נבצע מדגם חדש אז אנחנו עלולים לקבל תוצאה אחרת. לכן יתכנו טעויות במסקנות שלנו:

		הכרעה	
		$H_0$	$H_1$
מציאות	$H_0$	אין טעות	טעות מסוג 1
	$H_1$	טעות מסוג 2	אין טעות

### הגדרת הטעויות:

טעות מסוג ראשון: להכריע לדחות את  $H_0$  למרות שבמציאות  $H_0$  נכונה.

טעות מסוג שני: להכריע לקבל את  $H_0$  למרות שבמציאות  $H_1$  נכונה.

### דוגמה (פתרון בהקלטה):

אדם חשוד בביצוע עבירה ונתבע בבית המשפט. אילו סוגי טעויות אפשריות בהכרעת הדין?

## שאלות:

- (1) לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 סמ"ק במדגם בגודל 25. בסופו של דבר הוחלט להכריע לטובת חברת המשקאות.
- א. רשמו את השערות המחקר.  
 ב. מה מסקנת המחקר?  
 ג. איזו סוג טעות יתכן וביצעו במחקר?
- (2) במחקר על פרמטר מסוים הוחלט בסופו של דבר לדחות את השערת האפס.
- א. האם ניתן לדעת אם בוצע טעות במחקר?  
 ב. 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 שכיום ממוצע מספר הילדים במשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה. על סמך תוצאות המדגם נקבע שלא ניתן לקבוע שבאופן מובהק תוחלת מספר הילדים למשפחה קטנה כיום.
- א. מהי אוכלוסיית המחקר?  
 ב. מה המשתנה הנחקר?  
 ג. מה הפרמטר הנחקר?  
 ד. מה השערות המחקר?  
 ה. מה מסקנת המחקר?  
 ו. מהי סוג הטעות האפשרית במחקר?

## תשובות סופיות:

- (1) א.  $H_0: \mu = 500$   
 ב. לא דחינו את  $H_0$ .  
 ג. טעות מסוג שני.
- (2) א. לא ניתן לדעת.  
 ב. טעות מסוג ראשון.  
 (3) א. משפחות כיום.  
 ב. מס' הילדים.  
 ג. תוחלת מספר הילדים למשפחה כיום.  
 ה. לא לדחות את  $H_0$ . ו. טעות מסוג שני.
- ד.  $H_0: \mu = 2.3$   
 $H_1: \mu < 2.3$

# סטטיסטיקה והסתברות 01 88165

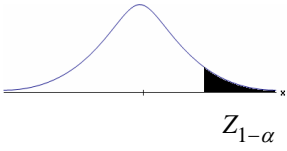
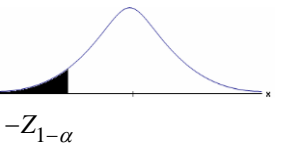
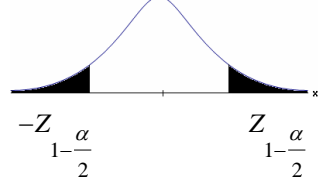
פרק 67 - בדיקת השערות על תוחלת (ממוצע)

תוכן העניינים

1. בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסיה ידועה ..... 336
2. סיכוי לטעויות ועוצמה (ששונות האוכלוסיה ידועה) ..... 340
3. קביעת גודל מדגם (ששונות האוכלוסיה ידועה) ..... 346
4. מובהקות תוצאה - אלפא מינימלית (ששונות האוכלוסיה ידועה) ..... 349
5. בדיקת השערות על תוחלת ( ממוצע) כששונות האוכלוסיה לא ידועה ..... 354
6. מובהקות תוצאה - אלפא מינימלית (ששונות האוכלוסיה לא ידועה) ..... 358
7. הקשר בין רווח סמך לבדיקת השערות על תוחלת (ממוצע) ..... 361

## בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסייה ידועה:

רקע:

$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	השערת האפס: השערה אלטרנטיבה:
1. $\sigma$ ידועה 2. $X \sim N$ או מדגם מספיק גדול			תנאים:
$Z_{\bar{x}} > Z_{1-\alpha}$	$Z_{\bar{x}} < -Z_{1-\alpha}$	$Z_{\bar{x}} < -Z_{1-\frac{\alpha}{2}}$ או $Z_{\bar{x}} > Z_{1-\frac{\alpha}{2}}$	כלל ההכרעה: אזור הדחייה של $H_0$ :
			
דוחים את $H_0$ ■	דוחים את $H_0$ ■	דוחים את $H_0$ ■	

$$Z_{\bar{x}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

סטטיסטי המבחן:

חלופה אחרת לכלל הכרעה:

$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$	נדחה $H_0$ אם מתקיים:
--	--	--	-----------------------

**דוגמה:**

יבול העגבניות מתפלג נורמלית עם תוחלת של 10 טון לדונם וסטיית תקן של 2.5 טון לדונם בעונה. משערים ששיטת זיבול חדשה תעלה את תוחלת היבול לעונה מבלי לשנות את סטיית התקן. נדגמו 4 חלקות שזובלו בשיטה החדשה. היבול הממוצע שהתקבל היה 12.5 טון לדונם. בדקו את ההשערה ברמת מובהקות של 1%.

**פיתרון:**

אוכלוסייה: עגבניות.

המשתנה:  $X =$  יבול העגבניות בטון לעונה.

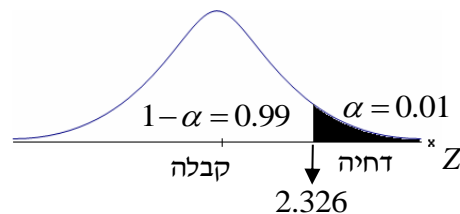
הפרמטר:  $\mu =$  תוחלת היבול בשיטת הזיבול החדשה.

השערות:  
 $H_0: \mu = 10$   
 $H_1: \mu > 10$

**תנאים:**

1.  $X \sim N$

2.  $\sigma = 2.5$

**כלל הכרעה:**

נדחה את  $H_0$  אם  $Z_{\bar{x}} > 2.326$

תוצאות:  $n = 4$ ,  $\bar{x} = 12.5$

סטטיסטי המבחן:  $Z_{\bar{x}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

נציב:  $Z_{\bar{x}} = \frac{1.25 - 10}{\frac{2.5}{\sqrt{4}}} = 2 < 2.326$

**מסקנה:**

לא נדחה  $H_0$  (נקבל  $H_0$ ).

ברמת מובהקות של 1% לא נוכל לקבל את הטענה ששיטת הזיבול החדשה מעלה את תוחלת היבול של העגבניות.

## שאלות:

- (1) ממוצע הציונים בבחינת 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 אקראיים. ממוצע הציונים של אותם תלמידים לאחר שלמדו 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 ס"מ<sup>3</sup> וסטיית תקן 20 ס"מ<sup>3</sup>. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 ס"מ<sup>3</sup> במדגם בגודל 25. א. מה מסקנתכם ברמת מובהקות של 2.5%? ב. האם ניתן לדעת מה תהיה המסקנה עבור רמת מובהקות הגבוהה מ-5%?
- (3) מהנדס האיכות מעוניין לבדוק אם מכונה מכיילת (מאופסת). המכונה כוונה לחתוך מוטות באורך 50 ס"מ. לפי נתוני היצרן סטיית התקן בחיתוך המוטות היא 0.5 ס"מ. במדגם של 50 מוטות התקבל ממוצע אורך המוט 50.93 ס"מ. מה מסקנתכם ברמת מובהקות של 5%?
- (4) המשקל הממוצע 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 שצריכה להביא להורדת המשקל. לשם בדיקת יעילות הדיאטה נלקח מדגם מקרי של 50 ספורטאים ובתום שנה של שימוש בדיאטה התברר שהמשקל הממוצע במדגם 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. במדגם של 25 ברגים העובי הממוצע 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.
- א. הגדלת רמת המובהקות לא תשנה את מסקנת המחקר.  
 ב. הגדלת רמת המובהקות תשנה את מסקנת המחקר.  
 ג. הקטנת רמת המובהקות לא תשנה את מסקנת המחקר.  
 ד. הקטנת רמת המובהקות תשנה את מסקנת המחקר.

(7) חוקר ערך מבחן דו צדדי ברמת מובהקות של  $\alpha$  והחליט לדחות את השערת האפס.

אם החוקר היה עורך מבחן צדדי ברמת מובהקות של  $\frac{\alpha}{2}$  אזי בהכרח:

- א. השערת האפס הייתה נדחית.
- ב. השערת האפס הייתה לא נדחית.
- ג. לא ניתן לדעת מה תהיה מסקנתו במקרה זה.

(8) שני סטטיסטיקאים בדקו השערות:  $H_0: \mu = \mu_0$  כנגד  $H_1: \mu > \mu_0$ ,

עבור שונות ידועה ובאותה רמת מובהקות.

שני החוקרים קבלו אותו ממוצע במדגם אך לחוקר א' היה מדגם בגודל 100 ולחוקר ב' מדגם בגודל 200.

- א. אם חוקר א' החליט לדחות את  $H_0$ , מה יחליט חוקר ב'? נמקו.
- ב. אם חוקר א' יחליט לא לדחות את  $H_0$ , מה יחליט חוקר ב'? נמקו.

### תשובות סופיות:

(1) נקבל  $H_0$ , בר"מ של 5% לא נקבל את הטענה של המורה ששיטת הלימוד שלו מעלה את ממוצע הציונים.

(2) א. נדחה  $H_0$ , בר"מ של 2.5% נקבל את תלונת אגודת הצרכנים בדבר הפחתת נפח המשקה בבקבוק.

ב. הגדלנו את רמת המובהקות לכן אנחנו נשארים בדחייה של  $H_0$  והמסקנה לא תשתנה.

(3) נדחה  $H_0$ , בר"מ של 5% נקבע שהמכונה לא מאופסת.

(4) נדחה  $H_0$ , בר"מ של 0.1 נקבל את הטענה שהדיאטה יעילה ומפחיתה את המשקל הממוצע.

(5) נקבל  $H_0$ , בר"מ של 0.05 נכריע שתוחלת עובי הבורג מתים למפרט.

(6) א'.

(7) ג'.

(8) א. לדחות. ב. לא ניתן לדעת.

## סיכוי לטעויות ועוצמה (ששונות האוכלוסייה ידועה):

רקע:

		הכרעה	
		$H_0$	$H_1$
מציאות	$H_0$	אין טעות	טעות מסוג 1
	$H_1$	טעות מסוג 2	אין טעות

הגדרת הסתברויות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):

$$\alpha = P(H_0 \text{ לדחות את } H_0 \mid H_0 \text{ נכונה}) = P_{H_0}(H_0)$$

הסיכוי לבצע טעות מסוג 2:

$$\beta = P(H_0 \text{ לקבל את } H_0 \mid H_1 \text{ נכונה}) = P_{H_1}(H_0)$$

רמת בטחון:

$$(\alpha - 1) = P(H_0 \text{ לקבל את } H_0 \mid H_0 \text{ נכונה}) = P_{H_0}(H_0)$$

עוצמה:

$$(\beta - 1) = \pi = P(H_0 \text{ לדחות את } H_1 \mid H_1 \text{ נכונה}) = P_{H_1}(H_0)$$

**התהליך לחישוב סיכוי לטעות מסוג שני:**

השערת האפס: השערת האלטרנטיבה:	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$
תנאים:	1. $\sigma$ ידועה 2. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של $H_0$ :	$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
חישוב $\beta$ :	$P_{\mu_1} \left( \bar{X} < \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right)$	$P_{\mu_1} \left( \bar{X} > \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right)$	$P_{\mu_1} \left( \mu_0 - Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$

התפלגות ממוצע המדגם:  $\bar{X} \sim N \left( \mu, \frac{\sigma^2}{n} \right)$

התקנון:  $Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

**דוגמה:**

בתחילת השנה חשבון הטלפון הסלולארי הממוצע לאדם היה 200₪ עם סטיית תקן של 80₪ לחודש. בעקבות כניסתן של חברות טלפון סלולארית חדשות מעוניינים לבדוק האם כיום ממוצע חשבון הטלפון הסלולארי פחת. לצורך בדיקה דגמו באקראי 36 אנשים וחשבון הטלפון הסלולארי שלהם היה 150₪ בממוצע לחודש.

- רשמו את השערות המחקר ובנו כלל הכרעה במונחי חשבון ממוצע מדגמי ברמת מובהקות של 5%.
- מה מסקנתכם? איזה סוג טעות אפשרית במסקנה?
- נניח שבמציאות כיום החשבון הממוצע הוא 160₪. מה הסיכוי לבצע טעות מסוג שני?
- אם נקטין את רמת המובהקות מסעיף א', כיצד הדבר ישפיע על התשובה מסעיף ג'?

## פתרון:

א. אוכלוסייה: משלמי חשבון טלפון סלולאר כיום.

המשתנה:  $X =$  חשבון הטלפון החודשי בשקלים.

הפרמטר:  $\mu$ .

השערות:  
 $H_0: \mu = 200$   
 $H_1: \mu < 200$

תנאים:

1.  $\mu = 200$ .

2.  $n = 36$ .

$$\bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, K = \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\alpha = 0.05$$

$$Z_{1-\alpha} = Z_{0.95} = 1.645$$

$$\text{נציב: שקלים } K = 200 - 1.645 \cdot \frac{80}{\sqrt{36}} = 178.07$$

כלל ההכרעה: דחה את  $H_0$  אם שקלים  $\bar{X} < 178.07$ .

ב. ברמת מובהקות של 5% נכריע שאכן ממוצע חשבון הטלפון הסלולרי פחת מתחילת השנה.

ג. השערות:  
 $H_0: \mu_0 = 200$   
 $H_1: \mu < 200$

כלל ההכרעה: נדחה את  $H_0$  אם  $\bar{X} < 178.07$ .

$$H_1: \bar{X} \sim N\left(160, \frac{80^2}{36}\right)$$

$$Z = \frac{178.07 - 160}{\frac{80}{\sqrt{36}}} = 1.36$$

$$\beta = P_{H_1}(\text{לקבל את } H_0) = P_{H_1}(\bar{X} > 178.07) = 1 - \phi(1.36) = 1 - 0.9131 = 0.0869$$

ד. הקטנת  $\alpha$  מגדילה את  $\beta$ .

## שאלות:

$$(1) \text{ נתון ש: } X \sim N(\mu, \sigma^2 = 1)$$

להלן השערות של חוקר לגבי הפרמטר  $\mu$ :  $H_0: \mu = 5$ ,  $H_1: \mu = 7$ . מעוניינים ליצור כלל הכרעה המתבסס על הסמך תצפית בודדת כך שרמת המובהקות תהיה 5%.

- עבור אילו ערכים של  $X$  שידגם נדחית השערת  $H_0$ ?
- מה הסיכוי לבצע טעות מסוג שני?
- 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 סטיית תקן 0.4. מעוניינים לבדוק אם כיום ממוצע מספר הילדים למשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה.

- רשמו כלל הכרעה במונחי ממוצע מדגם קריטי ברמת מובהקות של 5%.
- בהמשך לסעיף א' מה תהיה המסקנה ומהי הטעות האפשרית במסקנה?
- אם באמת ממוצע מספר הילדים במשפחה פחת לכדי 2.1 מהי העצמה של הכלל מסעיף א'?

(3) להלן נתונים על תהליך של בדיקת השערות על תוחלת:

$$H_0: \mu = 200, H_1: \mu \neq 200, \sigma = 30, n = 225$$

- רשמו כלל הכרעה במונחי ממוצע מדגם קריטי וברמת מובהקות של 10%.
- בהמשך לסעיף א', מהי העצמה אם התוחלת שווה ל-195?
- הסבירו, ללא חישוב, איך העצמה תשתנה אם רמת המובהקות תהיה 5%?

(4) מפעל לייצור צינורות מייצר צינור שקוטרו מתפלג נורמלית עם תוחלת של 50 מ"מ וסטית תקן של 6 מ"מ. במחלקת ביקורת האיכות דוגמים בכל יום 81 צינורות ומודדים את קוטרים, בכדי לבדוק, בעזרת מבחן סטטיסטי, האם מכונת הייצור מכוילת כנדרש או שקוטר הצינורות קטן מהדרוש.

- רשמו את ההשערות ואת כלל ההכרעה ברמת מובהקות של 5%.
- אם ביום כלשהו מכונת הייצור התקלקלה והיא מייצרת את הצינורות בקוטר שתוחלתו 48 מ"מ בלבד (סטית התקן לא השתנתה), מה ההסתברות שהתקלה לא תתגלה בביקורת האיכות? כיצד נקראת הסתברות זו?
- הסבירו ללא חישוב כיצד התשובה לסעיף ב' תשתנה אם רמת המובהקות תגדל.
- הסבירו ללא חישוב כיצד התשובה לסעיף ב' תשתנה 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 תצפיות. ידוע שסטיית התקן של ההתפלגות הינה 20.
- בנו כלל הכרעה שהסיכוי לטעות מסוג שני בו הוא 10%. מהי רמת המובהקות?
  - כיצד הייתה משתנה רמת המובהקות אם (כל סעיף בפני עצמו)?
    - סטיית התקן הייתה יותר גדולה.
    - הסיכוי לטעות מסוג שני גדול יותר.

השאלות שלהלן הן שאלות רב-ברירה, בחרו בתשובה הנכונה ביותר:

- 6) אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו אזי:
- הסיכוי לטעות מסוג ראשון גדל.
  - העוצמה של המבחן גדלה.
  - הסיכוי לטעות מסוג שני גדל.
  - תשובות א' ו-ב' נכונות.

- 7) חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן:
- השערת האפס נכונה.
  - השערת האפס נדחתה.
  - השערת האפס לא נדחתה.
  - אף אחת מהתשובות לא נכונה בהכרח.

- 8) מה המצב הרצוי לחוקר המבצע בדיקת השערה:

$1 - \beta$	$\alpha$
גדולה	א. גדולה
קטנה	ב. גדולה
גדולה	ג. קטנה
קטנה	ד. קטנה

- 9) נערך שינוי בכלל ההחלטה של בדיקת השערה מסוימת ובעקבותיו אזור דחיית  $H_0$  קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה מכך:
- הן  $\alpha$ , והן  $1 - \beta$ , יקטנו.
  - $\alpha$  יישאר ללא שינוי ואילו  $1 - \beta$  יגדל.
  - $\alpha$  יגדל ואילו  $1 - \beta$  יקטן.
  - הן  $\alpha$  והן  $1 - \beta$  יגדלו.

**10** ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניח שלחץ הדם בקרב עיתונאים גבוה יותר מהממוצע באוכלוסייה. הוא לקח מדגם של 60 עיתונאים וקיבל ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק שלחץ הדם בקרב העיתונאים אינו גבוה יותר. מה הטעות האפשרית שהרופא עושה?

- א. טעות מסוג ראשון.
- ב. טעות מסוג שני.
- ג. טעות מסוג שלישי.
- ד. אין טעות במסקנתו.

### תשובות סופיות:

- 1) א. מעל 6.645. ב. 0.3594.  
ג. דחינו את  $H_0$ , תתכן טעות מסוג ראשון.
- 2) א. נדחה  $H_0$  אם  $\bar{X} < 2.24$ . ב. נדחה  $H_0$  ג. 1.
- 3) א. נדחה  $H_0$  אם  $\bar{X} > 203.29$  או  $\bar{X} < 196.71$ . ב. 0.8051. ג. תקטן.
- 4) א. נדחה  $H_0$  אם  $\bar{X} < 48.9$ . ב. 0.0885. ג. תקטן. ד. תקטן.
- 5) א. 0.0033. ב. i. רמת המובהקות הייתה קטנה.  
ii. רמת המובהקות הייתה גדלה.
- 6) ד'
- 7) ג'
- 8) ג'
- 9) א'
- 10) ב'

## קביעת גודל מדגם (ששונות האוכלוסייה ידועה):

### רקע:

השערות המחקר הן:  $H_0: \mu = \mu_0$ ,  $H_1: \mu = \mu_1$ .

סטיית התקן של האוכלוסייה ידועה  $\sigma$  ומעוניינים לבצע מחקר שרמת המובהקות לא תעלה על  $\alpha$  והסיכוי לטעות מסוג שני לא יעלה על  $\beta$ .

$$n \geq \left( \frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

הנוסחה הבאה נותנת את גודל המדגם הרצוי:

### דוגמה:

משרד החינוך מפעיל בגן חובה שיטת חינוך שפותחה בשנת 1995. לפי שיטת חינוך זו תוחלת הציון במבחן אוצר מילים לגיל הרך הוא 70. אנשי חינוך החליטו לבדוק שיטת חינוך שפותחה בהולנד הנותנת שם תוחלת ציון אוצר מילים של 80. נניח שציוני מבחן זה מתפלגים נורמאלית עם  $\sigma = 17$ . כדי לבדוק האם גם בישראל הפעלת שיטת החינוך ההולנדית תעבוד בגנים, רוצים לבנות מחקר ברמת מובהקות של 5%. כמו כן, מעוניינים שאם בהפעלת 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 במבחן אוצר מילים.

הפרמטר:  $\mu$ .

השערות:  
 $H_0: \mu = 70$   
 $H_1: \mu = 80$

$$X \sim N(\mu, \sigma^2 = 17^2)$$

אם בהפעלת השיטה ההולנדית התוחלת תעלה ל-80, נגלה זאת בסיכוי 90%.

$$n \geq \left( \frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

$$\alpha = 0.05$$

$$1 - \beta = 0.9$$

$$\mu_0 = 70$$

$$\mu_1 = 80$$

$$\sigma = 17$$

$$Z_{1-\alpha} = Z_{0.95} = 1.645$$

$$Z_{1-\beta} = Z_{0.9} = 1.282$$

$$n \geq \left( \frac{(1.645 + 1.282) \times 17}{70 - 80} \right)^2 = 24.76 \quad \text{נציב:}$$

$$\text{לכן, } n_{\min} = 25$$

## שאלות:

(1) במבחן אינטליגנציה הציונים מתפלגים נורמאלית עם סטיית תקן 8 וממוצע 100. פסיכולוג מעוניין לבדוק את הטענה שבאוכלוסיות במצב סוציו אקונומי נמוך תוחלת הציונים היא 95. אם מעוניינים לגלות את הטענה בהסתברות של לפחות 99% כשרמת 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 החודשי 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 הטלפון הסלולרית נכונה המחקר יגלה זאת בסיכוי של 90% (איך קוראים להסתברות זאת?).

(3) השערות המחקר הן:  $H_0: \mu = \mu_0$ ,  $H_1: \mu = \mu_1$ .

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

יעלה על  $\beta$ . הוכיחו שגודל המדגם הרצוי לכך יהיה: 
$$n \geq \left( \frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

## תשובות סופיות:

(1) 41.

(2) 78.

(3) שאלת הוכחה.

## מובהקות תוצאה – אלפא מינימלית (ששונות האוכלוסייה ידועה):

### רקע:

דרך נוספת להגיע 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 יוכל לחשב רק אחרי שיהיו לו את התוצאות.

המסקנה של המחקר תקבע לפי העיקרון הבא: אם  $p_v \leq \alpha$ , דוחים את  $H_0$ . מובהקות התוצאה זה הסיכוי לקבלת תוצאות המדגם וקיצוני מתוצאות אלה בהנחת השערת האפס.

(לקבל את תוצאות המדגם וקיצוני)  $p_v = P_{H_0}$ .

אם ההשערה היא דו צדדית:

(לקבל את תוצאות המדגם וקיצוני)  $p_v = 2P_{H_0}$ .

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	השערת האפס: השערה אלטרנטיבה:
1. $\sigma$ ידועה			תנאים:
2. $X \sim N$ או מדגם מספיק גדול			
$P_{H_0}(\bar{X} \geq \bar{x})$	$P_{H_0}(\bar{X} \leq \bar{x})$	אם $2 \cdot P_{H_0}(\bar{X} \geq \bar{x}) \leftarrow \bar{x} > \mu_0$ אם $2 \cdot P_{H_0}(\bar{X} \leq \bar{x}) \leftarrow \bar{x} < \mu_0$	p-value

כאשר בהנחת 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 לצבא לפני 20 שנה היה 65 ק"ג. מחקר מעוניין לבדוק האם כיום המשקל הממוצע של מתגייסים גבוה יותר. נניח שמשקל המתגייסים מתפלג נורמאלית עם סטיית תקן של 12 ק"ג. במדגם של 16 מתגייסים התקבל משקל ממוצע של 71 ק"ג.

א. מהי מובהקות התוצאה?

ב. מה המסקנה אם רמת המובהקות היא 5% ואם רמת המובהקות היא 1%?

**פתרון:**

א. אוכלוסייה: המתגייסים לצבא כיום.

משתנה:  $X =$  משקל בק"ג.

פרמטר:  $\mu$ .

$$H_0: \mu = 65$$

השערות:  $H_1: \mu > 65$

תנאים:

$$1. X \sim N$$

$$2. \sigma = 12$$

תוצאות מדגם:

$$n = 16$$

$$\bar{X} = 71$$

$$P_V = P_{H_0} \left( \begin{array}{c} \text{לתוצאות} \\ \text{המדגם} \\ \text{וקיצוני} \end{array} \right) = P_{H_0} (\bar{X} \geq 71) = 1 - \phi(2) = 1 - 0.9772 = 0.0228$$

$$Z_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{71 - 65}{\frac{12}{\sqrt{16}}} = 2$$

$$\alpha_{\min} = 0.0228$$

## שאלות:

- (1) להלן השערות של מחקר:  $H_0: \mu = 70$ ,  $H_1: \mu > 70$ . המשתנה הנחקר מתפלג נורמלית עם סטיית תקן 20. במדגם מאותה אוכלוסייה התקבלו התוצאות הבאות:  $n = 100$ ,  $\bar{x} = 74$ . מהי מובהקות התוצאה?
- (2) השכר הממוצע במשק בשנת 2012 היה 8800 ₪ עם סטיית תקן 2000. במדגם שנעשה אתמול על 100 עובדים התקבל שכר ממוצע 9500 ₪. מטרת המחקר היא לבדוק האם כיום חלה עליה בשכר. עבור אילו רמות מובהקות שיבחר החוקר יוחלט שחלה עליה בשכר הממוצע במשק?
- (3) אדם חושד שחברת ממתקים לא עומדת בהתחייבויותיה, ומשקלו של חטיף מסוים אותו הוא קונה מדי בוקר נמוך מ-100 גרם. חברת הממתקים טוענת מצידה שהיא אכן עומדת בהתחייבויותיה. ידוע כי סטיית התקן של משקל החטיף היא 12 גרם. האדם מתכוון לשקול 100 חפיסות חטיפים ולאחר מכן 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 נקבל את השערת האפס?  
 ד. מה המסקנה ברמת מובהקות של 5?
- (4) מכונה לחיתוך מוטות במפעל חותכת מוטות באורך שמתפלג נורמלית עם תוחלת אליה כוונה המכונה וסטיית תקן 2 ס"מ. ביום מסוים כוונה המכונה לחתוך מוטות באורך 80 ס"מ. אחראי האיכות מעוניין לבדוק האם המכונה מכוילת. לצורך כך נדגמו מקו הייצור 16 מוטות שנחתכו אורכן הממוצע היה 81.7 ס"מ.
- א. מהי רמת המובהקות המינימלית עבורה נכריע שהמכונה לא מכוילת?  
 ב. אם נוסיף עוד תצפית שערכה יהיה 82 ס"מ, כיצד הדבר ישפיע על התשובה של הסעיף הקודם?  
 ג. הכרע ברמת מובהקות של 5% האם המכונה מכוילת.
- (5) אם מקבלים בחישובים אלפא מינימלית (P value) קטנה מאד, סביר להניח כי החוקר ידחה את השערת האפס בקלות. נכון/לא נכון? נמק.

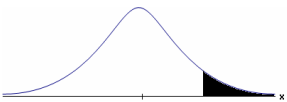

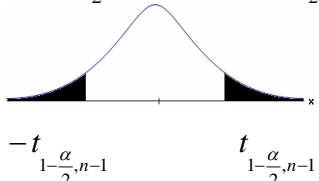
- 6) בבדיקת השערות התקבל שה-  $p\text{-value} = 0.02$ .  
 מה תהיה מסקנת חוקר המשתמש ברמת מובהקות 1%?  
 בחרו בתשובה הנכונה.
- א. יקבל את השערת האפס בכל מקרה.
  - ב. ידחה את השערת האפס מקרה.
  - ג. ידחה את השערת האפס רק אם המבחן הנו דו צדדי.
  - ד. לא ניתן לדעת כי אין מספיק נתונים.
- 7) מובהקות התוצאה (PV) היא גם (בחרו בתשובה הנכונה):
- א. רמת המובהקות המינימאלית לדחות השערת האפס.
  - ב. רמת המובהקות המקסימאלית לדחיית השערת האפס.
  - ג. רמת המובהקות שנקבעת מראש על ידי החוקר שטרם קיבל את תוצאות המחקר.
  - ד. רמת המובהקות המינימאלית לאי דחיית השערת האפס.
- 8) בבדיקת השערות מסוימת התקבל:  $p\text{ value} = 0.0254$  לכן (בחרו בתשובה הנכונה):
- א. ברמת מובהקות של 0.01 אך לא של 0.05 נדחה את  $H_0$ .
  - ב. ברמת מובהקות של 0.01 ושל 0.05 לא נדחה את  $H_0$ .
  - ג. ברמת מובהקות של 0.05 אך לא של 0.01 נדחה את  $H_0$ .
  - ד. ברמת מובהקות של 0.01 ושל 0.05 נדחה את  $H_0$ .

### תשובות סופיות:

- (1) 0.0228.
- (2) עבור כל רמת מובהקות סבירה.
- (3) א.  $H_0: \mu = 100$   
 $H_1: \mu < 100$   
 ב. 0.1056. ג. 0.1056.
- ד. נכריע שיש עמידה בהתחייבות של החברה.
- (4) א. 0.0006. ב. יקטן. ג. נכריע שאין כיול.
- (5) נכון.
- (6) א'.
- (7) א'.
- (8) ג'.

## בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסייה לא ידועה:

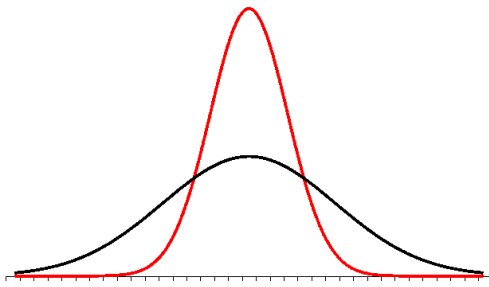
רקע:

$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	השערת האפס: השערה אלטרנטיבה:
1. $\sigma$ אינה ידועה 2. $X \sim N$ או מדגם מספיק גדול			תנאים:
$t_{\bar{x}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha, n-1}$ - דוחים את $H_0$	$t_{\bar{x}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha, n-1}$ - דוחים את $H_0$	$t_{\bar{x}} < -t_{1-\frac{\alpha}{2}}^{(n-1)}$ או $t_{\bar{x}} > t_{1-\frac{\alpha}{2}}^{(n-1)}$  $-t_{1-\frac{\alpha}{2}, n-1}$ $t_{1-\frac{\alpha}{2}, n-1}$ - דוחים את $H_0$	כלל ההכרעה: אזור הדחייה של $H_0$ :
$\bar{X} > \mu_0 + t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} < \mu_0 - t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$ או $\bar{X} < \mu_0 - t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$	חלופה לכלל הכרעה: נדחה $H_0$ אם מתקיים:

$$t_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

סטטיסטי המבחן:

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



### התפלגות T:

הינה התפלגות סימטרית פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים. התפלגות T תלויה במושג שנקרא דרגות חופש.

דרגות החופש הן:  $df = n - 1$ .

ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה. כשדרגות החופש שואפות 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 נדגמו 10 משטחים ונמצא שהעובי הממוצע הוא 0.104 עם אומדן לסטיית תקן 0.002 ס"מ.

א. מהן השערות המחקר?

ב. מה ההנחה הדרושה לצורך פתרון?

ג. בדוק ברמת מובהקות של 5%.

## שאלות:

(1) משך זמן ההחלמה בלקיחת אנטיביוטיקה מסוימת 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 האחרת התקבלו זמני ההחלמה הבאים: 90, 95, 100, 80, 125 שעות. מה מסקנתכם ברמת מובהקות של 5%. מהי ההנחה הדרושה לצורך הפתרון?

(2) משרד הבריאות פרסם שמשקל ממוצע של תינוקות ביום היוולדם בישראל 3300 גר'. משרד הבריאות רוצה לחקור את הטענה ששנים מעשנות בזמן ההיריון יולדות תינוקות במשקל נמוך מהממוצע. 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 אינטליגנציה מתפלגים נורמלית. בארה"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 מארה"ב? הסיקו ברמת מובהקות של 5%.

(4) באוכלוסייה מסוימת נדגמו 10 תצפיות 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 התבססו 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 יגיעו לאותה מסקנה.
  - לא ניתן לדעת על היחס בין דחיית השערת האפס של שני החוקרים.

- (6) נתון ש:  $X \sim N(\mu, \sigma^2)$  כמו כן נתונות ההשערות הבאות:  $H_0: \mu = \mu_0$  ו-  $H_1: \mu < \mu_0$ . חוקר בדק את ההשערות הללו על סמך מדגם שכלל 10 תצפיות.  $\sigma^2$  לא הייתה ידועה לחוקר. החוקר החליט לדחות את השערת האפס ברמת מובהקות של 5% לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תצפיות ושקלל את תוצאות אלה גם למדגם כך שכלל עכשיו 15 תצפיות. בחר בתשובה הנכונה:
- כעת בברור הוא ידחה את השערת האפס.
  - כעת הוא דווקא יקבל את השערת האפס.
  - כעת לא ניתן לדעת מה תהיה מסקנתו.

### תשובות סופיות:

- (1) נדחה  $H_0$ .
- (2) נדחה  $H_0$ .
- (3) נקבל  $H_0$ .
- (4) נקבל  $H_0$ .
- (5) ב'.
- (6) ג'.

## מובהקות תוצאה – אלפא מינימלית (ששונות האוכלוסייה לא ידועה):

### רקע:

נוכיר שהמסקנה של המחקר תיקבע לפי העיקרון הבא: אם  $p_v \leq \alpha$  דוחים את  $H_0$ . מובהקות התוצאה היא הסיכוי לקבלת תוצאות המדגם וקיצוני מתוצאות אלה בהנחת השערת האפס.

•  $p_v = P_{H_0}$  (לקבל את תוצאות המדגם וקיצוני)

אם ההשערה היא דו צדדית:

•  $p_v = 2P_{H_0}$  (לקבל את תוצאות המדגם וקיצוני)

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	השערת האפס: השערה אלטרנטיבה:
1. אינה ידועה או 2. מדגם מספיק גדול $X \sim N$			תנאים:
$P_{H_0}(\bar{X} \geq \bar{x})$	$P_{H_0}(\bar{X} \leq \bar{x})$	אם $2 \cdot P_{H_0}(\bar{X} \geq \bar{x}) \leftarrow \bar{x} > \mu_0$ אם $2 \cdot P_{H_0}(\bar{X} \leq \bar{x}) \leftarrow \bar{x} < \mu_0$	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 לעבודה 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 בדקות הם: 34, 40, 30, 32, 27. הניחו שזמן הנסיעה מתפלג נורמלית.

- א. רשמו את השערות המחקר.
- ב. מצאו חסמים למובהקות התוצאה.
- ג. מה המסקנה ברמת מובהקות של 5%?

**פתרון:**

אוכלוסייה: כלל הנסיעות לעבודה בדרך החלופית.

משתנה:  $X =$  זמן נסיעה בדקות.

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% נכריע שהדרך החלופית מהירה יותר.

## שאלות:

- (1) קו ייצור אריזות סוכר נארזות כך שהמשקל הממוצע של אריזות הסוכר צריך להיות אחד קילוגרם. בכל יום דוגמים מקו הייצור 5 אריזות במטרה לבדוק האם קו הייצור תקין. בבדיקה דגמו 5 אריזות סוכר ולהלן משקלן בגרמים: 1024, 1008, 1005, 996, 997.
- א. רשמו את השערות המחקר.  
 ב. מהי מובהקות התוצאה? הצג חסמים.  
 ג. מה המסקנה ברמת מובהקות של 5%?
- (2) חוקר בדק את הטענה כי פועלים העובדים במשמרת לילה איטיים יותר מפועלים העובדים ביום. ידוע כי משך הזמן הממוצע הדרוש לייצר מוצר מסוים ביום הוא 6 שעות. במדגם מיקרי של 25 פועלים שעבדו במשמרת לילה נמצא כי הזמן הממוצע לייצר אותו מוצר הוא 7 שעות עם סטית תקן של 3 שעות.
- מהי ה- $\alpha$  המינימלית שלפיה ניתן להחליט שאכן העובדים במשמרת לילה איטיים יותר?
- (3) הגובה של מתגייסים לצה"ל מתפלג נורמלית. במדגם של 25 מתגייסים מדדו את הגבהים שלהם בס"מ והתקבלו התוצאות הבאות:
- $$\bar{x} = 176.2, \sum (x_i - \bar{x})^2 = 2832$$
- מטרת המחקר היא לבדוק האם תוחלת הגבהים של המתגייסים גבוה מ-174 ס"מ באופן מובהק.
- מהי בקרוב מובהקות התוצאה ועל פיה מה תהיה המסקנה ברמת מובהקות של 6%?

## תשובות סופיות:

- (1) א.  $H_0: \mu = 1000$   
 ב.  $20\% \leq P_v \leq 50\%$   
 ג. ברמת מובהקות של 5% לא נוכל לקבוע שקו הייצור אינו תקין.
- (2) 10%
- (3) 1.01, נקבל את  $H_0$ .

## הקשר בין רווח סמך לבדיקת השערות על תוחלת (ממוצע):

### רקע:

ניתן לבצע בדיקת השערות דו צדדית ברמת מובהקות  $\alpha$  על  $\mu$  :

$$H_0 : \mu = \mu_0 , H_1 : \mu \neq \mu_0$$

על ידי בניית רווח סמך ברמת סמך של  $1-\alpha$  ל- $\mu$  :

אם  $\mu_0$  נופל ברווח  $\leftarrow$  נקבל 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 רווח סמך ברמה של 90% וקיבל:  $79 < \mu < 84$  .

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

### פתרון (פתרון מלא בהקלטה):

רווח הסמך ברמת סמך של 90% מכיל "80".

ברמת סמך של 95% רווח הסמך יגדל ויכיל "80".

לכן, ברמת מובהקות של 5% נקבל  $H_0$  .

## שאלות:

- (1) חוקר רצה לבדוק את ההשערות הבאות:  $H_0: \mu = 90$ ,  $H_1: \mu \neq 90$ . החוקר בנה רווח סמך לתוחלת ברמת סמך של 95% וקיבל את רווח הסמך הבא: (87, 97). אם החוקר מעוניין לבצע בדיקת השערות ברמת מובהקות של 1% האם ניתן להגיע למסקנה ע"ס רווח הסמך? נמקו.
- (2) חוקר מעוניין לבדוק השפעת דיאטה חדשה על רמת הסוכר בדם. ידוע כי מספר מיליגרם 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 המיליגרם סוכר היה 115.5 מ"ג לסמ"ק.
- א. בנה רווח סמך ברמת סמך 95% לתוחלת רמת הסוכר בדם אצל הניזונים מדיאטה זו.
- ב. ידוע שתוחלת רמת הסוכר בדם באוכלוסיה 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 הפנצילין היא 200 מ"ג לקפסולה. משרד הבריאות ביצע מדגם של 8 קפסולות אקראיות מקו הייצור ומצא שבממוצע יש 196 מ"ג פנצילין לקפסולה עם סטיית תקן מדגמית של 5 מ"ג. בהנחה וכמות הפנצילין בקפסולה מתפלגת נורמלית.
- א. בנו רווח סמך ברמת סמך של 95% לממוצע כמות הפנצילין לקפסולה המיוצרת על ידי יצרן האנטיביוטיקה.
- ב. בדקו ברמת מובהקות של 5% האם יש אמת באינפורמציה המסופקת על ידי היצרן.

## תשובות סופיות:

- (1) נקבל השערת.
- (2) א.  $112.87 \leq \mu \leq 118.13$ .
- ב. נכריע שהדיאטה משפיעה על תוחלת רמת הסוכר בדם.
- (3) א.  $191.8 \leq \mu \leq 200.2$ . ב. נכריע שיש אמת בפרסום.

# סטטיסטיקה והסתברות 01 88165

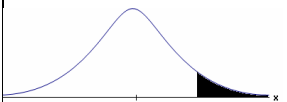


פרק 68 - בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

תוכן העניינים

1. כששושונות האוכלוסייה אינן ידועות והמדגמים גדולים ..... 363

## בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

כשהשנויות של האוכלוסייה לא ידועות והמדגמים גדולים – רקע

			השערת האפס: השערה אלטרנטיבית:
$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 > c$	$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 < c$	$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 \neq c$	
1. מדגמים בלתי תלויים 2. $\sigma_1, \sigma_2$ לא ידועות 3. מדגמים מספיק גדולים			תנאים:
$Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\alpha}$  $Z_{1-\alpha}$ דוחים את $H_0$	$Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ דוחים את $H_0$	או $Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\frac{\alpha}{2}}$ $Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}}, Z_{1-\frac{\alpha}{2}}$ דוחים את $H_0$	כלל ההכרעה: אזור הדחייה של $H_0$ :

$$Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - c}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad \text{סטטיסטי המבחן:}$$

## חלופה אחרת לכלל הכרעה:

נדחה $H_0$ אם מתקיים:	
$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ <p style="text-align: center;">או</p> $\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$
$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$	

## דוגמה (פתרון בהקלטה):

נרצה לבדוק האם קיים הבדל בין ממוצע ציוני הפסיכומטרי של חיילים לממוצע ציוני הפסיכומטרי של תלמידי תיכון. במדגם של 46 נבחנים חיילים התקבל ממוצע 543 וסטיית תקן 123. במדגם של 50 תלמידי תיכון התקבל ממוצע 508 וסטיית תקן 178. מה המסקנה ברמת מובהקות 5%?

## שאלות

(1) חברה להנדסת בניין מעוניינת להשוות ברמת הקשיות של שני סוגי ברגים. במדגם של 35 ברגים מסוג א' התקבל רמת קשיות ממוצעת של 28 יחידות וסטיית תקן 4, ובמדגם של 45 ברגים מסוג ב' התקבל רמת קשיות ממוצעת של 25 וסטיית תקן 6. האם על סמך תוצאות המדגם יש הבדל בין סוגי הברגים מבחינת רמת הקשיות שלהם? בדקו ברמת מובהקות של 5%.

(2) כדי לבדוק האם נהגים השותים תחת השפעת אלכוהול נוהגים מהר יותר מאלו שאינם שותים בוצע מדגם שבו בדקו את המהירות המקסימאלית של כל נהג בקמ"ש. להלן התוצאות:

S	$\bar{X}$	גדול מדגם	
20	80	70	נהגים השותים אלכוהול
15	60	100	נהגים שאינם שותים אלכוהול

א. מהי מובהקות התוצאה?

ב. מה המסקנה ברמת מובהקות של 5%?

## תשובות סופיות

(1) נדחה את  $H_0$ .

(2) א. 0. ב. נדחה את  $H_0$ .

# סטטיסטיקה והסתברות 01 88165

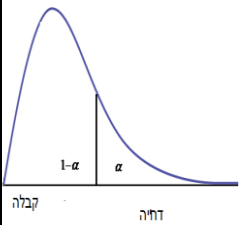
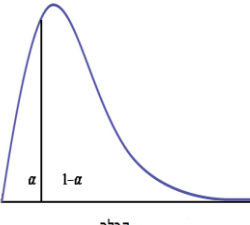
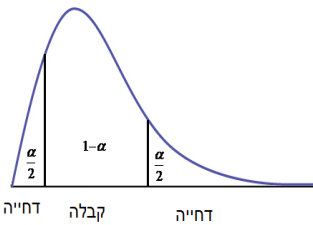
פרק 69 - בדיקת השערות על שוניות

תוכן העניינים

1. בדיקת השערות על שתי שוניות..... 366

## בדיקת השערות על שתי שוניות:

רקע:

השערת האפס : השערה אלטרנטיבית :	$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$	$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$	$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$
	$H_1 : \frac{\sigma_1^2}{\sigma_2^2} > 1$	$H_1 : \frac{\sigma_1^2}{\sigma_2^2} < 1$	$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1$
תנאים :	1. מדגמים בלתי תלויים 2. $X_1, X_2 \sim N$		
נדחה את השערת האפס אם :	 <p>קבלה דחייה</p> $F \geq f_{1-\alpha}^{(n_1-1, n_2-1)}$	 <p>דחייה קבלה</p> $F \leq \frac{1}{f_{1-\alpha}^{(n_2-1, n_1-1)}}$	 <p>דחייה קבלה דחייה</p> $F \geq f_{1-\alpha/2}^{(n_1-1, n_2-1)}$ <p>או</p> $F \leq \frac{1}{f_{1-\alpha/2}^{(n_2-1, n_1-1)}}$

$$F = \frac{S_1^2}{S_2^2} : \text{סטטיסטי המבחן}$$

התפלגות F:

אם  $X_1 \sim N(\mu_1, \sigma^2)$  ו-  $X_2 \sim N(\mu_2, \sigma^2)$  אזי:  $\frac{S_1^2}{S_2^2} \sim F(n_1-1, n_2-1)$

התפלגות F הינה התפלגות אסימטרית חיובית התלויה בדרגות חופש של המונה ושל המכנה.

כמו כן בהתפלגות F מתקיימת התכונה הבאה:  $F_{\alpha}(n_1-1, n_2-1) = \frac{1}{F_{1-\alpha}(n_2-1, n_1-1)}$



$$df_1 = n_1 - 1$$

$$df_2 = n_2 - 1$$

**דוגמה:**

מעוניינים להשוות בין נשים וגברים מבחינת השונות בזמנים שלהם לבצע משימה מסוימת. במדגם של 10 גברים התקבלו התוצאות הבאות לגבי זמן ביצוע המשימה:  $\sum (y_i - \bar{y})^2 = 204$ .

במדגם של 13 נשים התקבלו התוצאות הבאות:  $\sum (x_i - \bar{x})^2 = 200$ . בדקו ברמת מובהקות של 2% האם קיים הבדל בין השונות? מה יש להניח?

**פתרון:**

האוכלוסיות: נשים מול גברים.

משתנה:  $y =$  זמן ביצוע משימה של גבר,  $x =$  זמן ביצוע משימה של אישה

$$\frac{\sigma_x^2}{\sigma_y^2} : \text{פרמטר}$$

$$H_0 : \frac{\sigma_x^2}{\sigma_y^2} = 1$$

$$H_1 : \frac{\sigma_x^2}{\sigma_y^2} \neq 1$$

השערות:

הנחות: 1. מדגימים ב"ת 2.  $x, y \sim N$

כלל הכרעה:

$$\alpha = 2\%$$

$$n_1 = 13, d.f_1 = n_1 - 1 = 12$$

$$n_2 = 10, d.f_2 = n_2 - 1 = 9$$

נדחה את  $H_0$  אם  $F > 5.11$  או  $F < 0.23$

$$S_y^2 = \frac{204}{10-1} = 22\frac{2}{3} : \text{תוצאות המדגם}$$

$$S_x^2 = \frac{200}{13-1} = 16\frac{2}{3}$$

$$F = \frac{S_1^2}{S_2^2} = \frac{16\frac{2}{3}}{22\frac{2}{3}} : \text{סטטיסטי המבחן}$$

מסקנה: ברמת מובהקות של 2% נקבל את  $H_0$ .

לא קיים הבדל מובהק בין גברים לנשים מבחינת השונות שלהם.

$\alpha = 0.05$ טבלת ערכים קריטיים לפי התפלגות F																	
ד"ח מנהל"ח מנבה	1	2	3	4	5	6	7	8	9	10	12	16	20	24	60	120	$\infty$
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	246.46	248.01	249.05	252.20	253.25	254.31
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.69	8.66	8.64	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.84	5.80	5.77	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.60	4.56	4.53	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.92	3.87	3.84	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.49	3.44	3.41	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.20	3.15	3.12	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	2.99	2.94	2.90	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.83	2.77	2.74	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.70	2.65	2.61	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.60	2.54	2.51	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.51	2.46	2.42	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.44	2.39	2.35	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.38	2.33	2.29	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.33	2.28	2.24	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.29	2.23	2.19	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.25	2.19	2.15	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.21	2.16	2.11	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.18	2.12	2.08	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.16	2.10	2.05	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.13	2.07	2.03	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.11	2.05	2.01	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.09	2.03	1.98	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.07	2.01	1.96	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.05	1.99	1.95	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.04	1.97	1.93	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.02	1.96	1.91	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.01	1.94	1.90	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	1.99	1.93	1.89	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.90	1.84	1.79	1.64	1.58	1.51
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.95	1.85	1.78	1.74	1.58	1.51	1.44
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.82	1.75	1.70	1.53	1.47	1.39
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94	1.86	1.76	1.69	1.64	1.46	1.39	1.30
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.73	1.66	1.61	1.43	1.35	1.25
$\infty$	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.64	1.57	1.52	1.32	1.22	1.00

טבלת ערכים קריטיים לפי התפלגות F $\alpha = 0.01$ ראה איור מטה.																	
ד"ח מונה/ד"ח מכנה	1	2	3	4	5	6	7	8	9	10	12	16	20	24	60	120	$\infty$
1	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47	6055.85	6106.32	6170.10	6208.73	6234.63	6313.03	6339.39	6365.86
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.44	99.45	99.46	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.83	26.69	26.60	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.15	14.02	13.93	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.68	9.55	9.47	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.52	7.40	7.31	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.28	6.16	6.07	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.48	5.36	5.28	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.92	4.81	4.73	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.52	4.41	4.33	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.21	4.10	4.02	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	3.97	3.86	3.78	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.78	3.66	3.59	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.62	3.51	3.43	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.49	3.37	3.29	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.37	3.26	3.18	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.27	3.16	3.08	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.19	3.08	3.00	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.12	3.00	2.92	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.05	2.94	2.86	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	2.99	2.88	2.80	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.94	2.83	2.75	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.89	2.78	2.70	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.85	2.74	2.66	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.81	2.70	2.62	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.78	2.66	2.58	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.75	2.63	2.55	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.72	2.60	2.52	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.69	2.57	2.49	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.66	2.55	2.47	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.48	2.37	2.29	2.02	1.92	1.80
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.56	2.38	2.27	2.18	1.91	1.80	1.68
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.31	2.20	2.12	1.84	1.73	1.60
90	6.93	4.85	4.01	3.53	3.23	3.01	2.84	2.72	2.61	2.52	2.39	2.21	2.09	2.00	1.72	1.60	1.46
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.15	2.03	1.95	1.66	1.53	1.38
$\infty$	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.00	1.88	1.79	1.47	1.32	1.00

## שאלות:

- 1) להלן נתונים על שטחי דירות במ"ר עבור דירות חדשות שנבנו בשנת 2012 ובשנת 2013:

120	94	90	130	95	112	120	2012
	69	74	105	91	82	100	2013

- א. בדקו ברמת מובהקות של 10% את ההשערה ששונות שטחי הדירות החדשות בשנת 2012 ובשנת 2013 שוות. מה הן ההנחות הדרושות לביצוע הבדיקה?  
 ב. האם וכיצד הייתה משתנה המסקנה מהסעיף הקודם אם מסתבר שחלה טעות ברישום ויש להפחית 10 מ"ר מכל הדירות שמופיעות במדגם?

- 2) בתחום הבינוי משתמשים בשני סוגי מתכות: מתכת A ומתכת B. מחקר מעוניין לבדוק האם קיים הבדל בין שני סוגי המתכות מבחינת החוזק שלהן. דגמו מספר

B	A	סוג המתכת
10	8	$n$
30	16	$\sum X_i$
198	60	$\sum X_i^2$

- יחידות מתכת מכל סוג והתקבלו התוצאות הבאות:  
 יש להניח שרמת החוזק של המתכות מתפלגת נורמאלית.  
 א. האם קיים הבדל בין שונות החוזק של מתכות?  
 ב. האם קיים הבדל בין תוחלות החוזק של מתכות?  
 בכל סעיף רמת מובהקות של 10%.

- 3) מחקר סוציולוגי מעוניין לחקור את הרגלי הבילויים בקבוצות גיל שונות. ידוע כי בקרב האוכלוסייה הבוגרת (מעל 18) ההוצאה החודשית על בילויים מתפלגת נורמאלית עם תוחלת של 500₪ וסטיות תקן של 300₪. במדגם שנעשה על סטודנטים בגילאי 21-26 התקבל אומד חוסר הטיה לשונות ההוצאה החודשית על בילויים 10,000₪. כמות הסטודנטים שנדגמה 16. במדגם שנעשה על 11 מבוגרים בשנות השלושים התקבל אומד חסר הטיה לשונות ההוצאה החודשית על בילויים 490,000₪.

- א. בדקו ברמת מובהקות של 5% האם שונות ההוצאה על בילויים בקרב סטודנטים בקבוצת גילאי 21-26 נמוכה מהשונות אצל כלל המבוגרים.  
 ב. בדקו ברמת מובהקות של 1% האם הפיזור של ההוצאה החודשית לבילויים גדולה יותר בקבוצת גיל ה-30 מאשר בקבוצת גיל 21-26.

- 4) נתון  $X_i \sim N(\mu_x, \sigma^2)$ , וכמו כן  $Y_i \sim N(\mu_y, \sigma^2)$ . מאוכלוסייה X נדגמו 7 תצפיות ומאוכלוסייה Y נדגמו 13 תצפיות.

א. כיצד  $\frac{S_x^2}{S_y^2}$  מתפלג?

- ב. מה ההסתברות ש- $S_x^2$  גדולה ביותר מפי 3 מאשר  $S_y^2$ ?

**תשובות סופיות:**

- (1) א. לא נדחה את  $H_0$ .  
ב. מסקנה לא תשתנה.
- (2) א. לא נדחה את  $H_0$ .  
ב. לא נדחה את  $H_0$ .
- (3) א. נדחה את  $H_0$ .  
ב. נדחה את  $H_0$ .
- (4) א.  $F(6,12)$ .  
ב. 5%