

סטטיסטיקה



תוכן העניינים

1.	סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה	1
5.	סטטיסטיקה תיאורית- הצגה של נתונים	5
16.	סטטיסטיקה תיאורית- סכימה	16
20.	סטטיסטיקה תיאורית - מדדי מיקום מרכזי	20
29.	סטטיסטיקה תיאורית - מדדי פיזור - הטווח, השונות וסטיית התקן	29
32.	סטטיסטיקה תיאורית - מדדי פיזור - טווח בין רבעוני	32
34.	סטטיסטיקה תיאורית- מדדי מיקום יחסי-ציון תקן	34
36.	סטטיסטיקה תיאורית-מדדי מיקום יחסי-אחוזונים במחלקות	36
39.	סטטיסטיקה תיאורית - טרנספורמציה לינארית	39
42.	סטטיסטיקה תיאורית - שאלות מסכמות	42
49.	מקדם המתאם (מדד קשר) הלינארי ומובהקותו	49
65.	יסודות ההסתברות	65
69.	פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים	69
78.	קומבינטוריקה -כלל המכפלה	78
82.	קומבינטוריקה- תמורה - סידור עצמים בשורה	82
85.	קומבינטוריקה -דגימה סידורית ללא החזרה ועם החזרה	85
87.	קומבינטוריקה - דגימה ללא סדר וללא החזרה	87
90.	קומבינטוריקה - שאלות מסכמות	90
97.	הסתברות מותנית-במרחב מדגם אחיד	97
100.	הסתברות מותנית - מרחב לא אחיד	100
104.	דיאגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה	104
109.	תלות ואי תלות בין מאורעות	109
113.	שאלות מסכמות בהסתברות	113

תוכן העניינים

24.	המשתנה המקרי הבדיד - פונקציית ההסתברות.	118
25.	המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן.	122
26.	המשתנה המקרי הבדיד- טרנספורמציה לינארית.	126
27.	התפלגויות בדידות מיוחדות -התפלגות בינומית.	129
28.	התפלגויות בדידות מיוחדות -התפלגות גיאומטרית.	133
29.	המשתנה המקרי הבדיד - שאלות מסכמות.	136
30.	משתנה דו-מימדי בדיד - פונקציית הסתברות משותפת.	140
31.	משתנה דו מימדי בדיד - מתאם בין משתנים.	146
32.	המשתנה המקרי הדו ממדי - קומבינציות ליניאריות.	153

סטטיסטיקה

פרק 1 - סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה

תוכן העניינים

1. כללי 1

סטטיסטיקה תיאורית – סיווג משתנים וסולמות מדידה:

רקע:

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

מיון משתנים לפי סולמות המדידה:

1. סולם שמי (נומינאלי) – משתנה שלערכיו יש משמעות רק מבחינת הזהות ואין עניין של יותר או פחות. לדוגמה: מצב משפחתי (רווק/נשוי/אלמן/גרש), אזור מגורים. משתנה דיכוטומי (הינו מסולם שמי) אותם משתנים שיש להם רק שני ערכים אפשריות זכר/נקבה. מעשן/לא מעשן.
2. סולם סדר (אורדינאלי) – כאשר לערכים של המשתנה בנוסף לשם ישנה גם משמעות לסדר אבל אין משמעות לגודל ההפרש. למשל, דרגה בצבא.
3. סולם רווחים (אינטרוואלי) – משתנה שלערכים שלו בנוסף לשם ולסדר בניהם יש משמעות לרווחים בין הערכים אבל אין משמעות ליחס בין הערכים. למשל, קומה בבניין. סולם לא כל כך פופולרי.

סולם מנה/יחס:

משתנה שלערכיו בנוסף לשם, לסדר ולרווח יש משמעות גם ליחס בין הערכים. למשל, מספר מכוניות למשפחה, משקל אדם בק"ג. הדרך הקלה ביותר כדי לזהות עם הסולם הוא סולם מנה היא על ידי מבחן האפס. בסולם מנה האפס הוא מוחלט, אבסולוטי, ומייצג אין.

סוגי משתנים:

נבצע סיווג של המשתנים:

משתנה איכותי

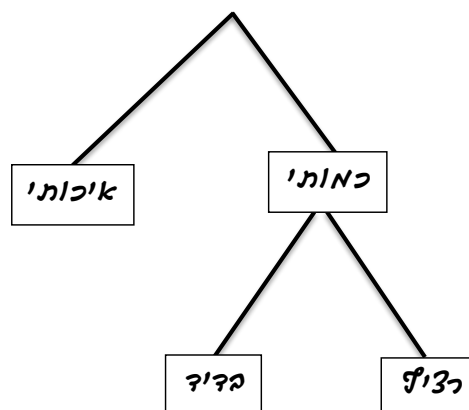
משתנה שלערכיו אין משמעות של יותר או פחות, אין עניין כמותי לערכים המתקבלים. כמו: מקום מגורים של אדם (רעננה, תל אביב, אשדוד...), מין האדם (זכר, נקבה), מצב משפחתי (רווק, נשוי, גרוש, אלמן).

משתנה כמותי

משתנה שערכיו הם מספרים להם יש משמעות כמותית כמו: גובה אדם בס"מ, ציון בבחינה וכדומה. את המשתנה הכמותי נסווג לשני סוגים:

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

משתנה רציף: משתנה שערכיו מתקבלים מתוך אינסוף ערכים בתחום מסוים, הערכים מתקבלים ברצף – ללא קפיצות של ערכים. דוגמאות: גובה בס"מ – אם הגובה הנמוך ביותר הוא 150 ס"מ ועד 190 ס"מ – הגבהים בקבוצה הם ברצף. גם בין 160 ל-161 ס"מ יש רצף אינסופי של ערכים אפשריים (כמו 160.233 ס"מ, למשל).



שאלות:

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

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

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

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

בחברה 200 עובדים.

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

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

- א. שכר ב-ש.
- ב. ציון בחינת בגרות.
- ג. תוצאה של הטלת קובייה.
- ד. מהירות ריצה בתחרות.
- ה. שיעור התמיכה בממשלה.

תשובות סופיות:

- | | | | |
|----|-------------------|--------------------------|----------|
| (1 | א. מנה. | ב. מנה. | ג. סדר. |
| | ד. סדר. | ה. מנה/ סדר. | ו. שמי. |
| | ז. שמי. | ח. שמי. | ט. סדר. |
| (2 | א. מספר האיחורים. | ב. כמותי בדיד בסולם מנה. | |
| (3 | א. רציף. | ב. בדיד. | ג. בדיד. |
| | ד. רציף. | ה. רציף. | |

סטטיסטיקה

פרק 2 - סטטיסטיקה תיאורית- הצגה של נתונים

תוכן העניינים

1. כללי 5

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: 3, 4, 3, 5, 4.

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	$\frac{f_k}{N} \cdot 100$
סה"כ	$N = \sum_{i=1}^k f_i$	100%

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

דוגמה:

להלן התפלגות הציונים בכיתה מסוימת:

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

שכיחות מצטברת – צבירה של השכיחויות.

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

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

 $\frac{f_i}{n}$ – איזה חלק מהתצפיות בקבוצה שוות לערך.**טבלת שכיחויות במחלקות:**

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

דוגמה:

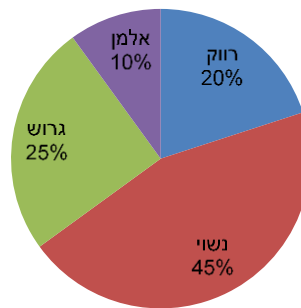
נתנו לקבוצת ילדים לבצע משימה, בדקו את התפלגות זמן הביצוע, בדקות.
להלן ההתפלגות שהתקבלה:

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:

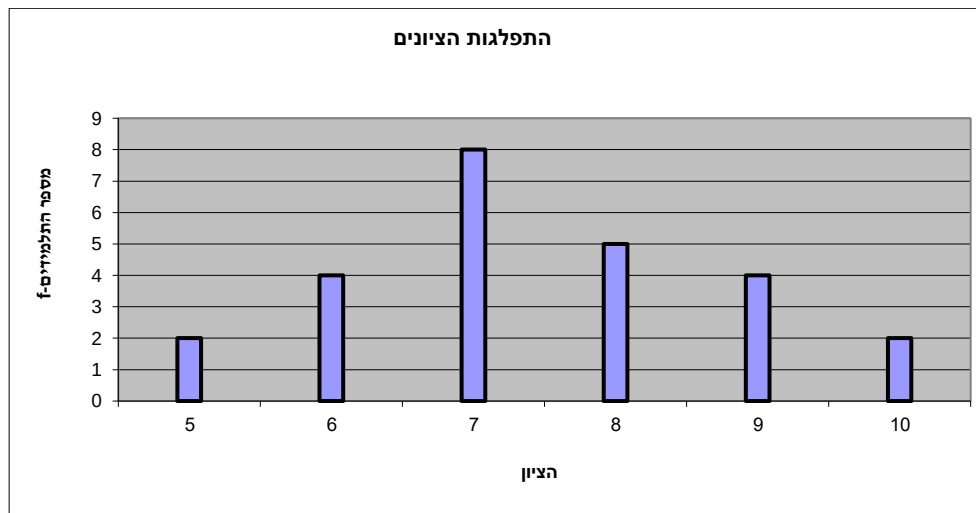
זהו התיאור הגרפי של משתנה איכותי. בדיאגרמת עוגה כל ערך במשתנה מקבל "נתח", שהוא פרופורציונלי לשכיחות היחסית של ערך המשתנה בנתונים.

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.

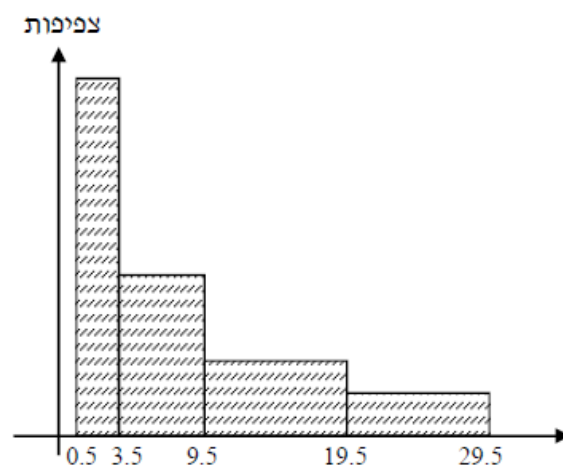


היסטוגרמה:

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



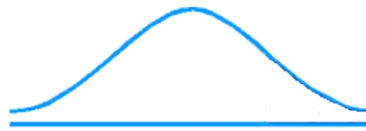
פוליגון – מצולעון:

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.

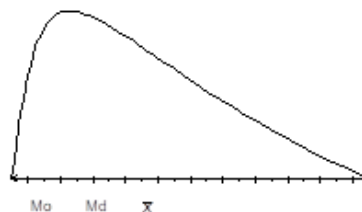


ישנן התפלגויות סימטריות שאינן פעמוניות, כגון:

התפלגות אסימטרית ימנית (חיובית)

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

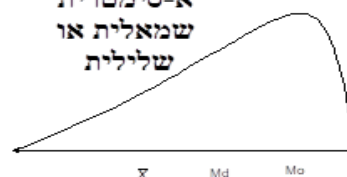
התפלגות א-סימטרית
ימנית או חיובית



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":

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

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

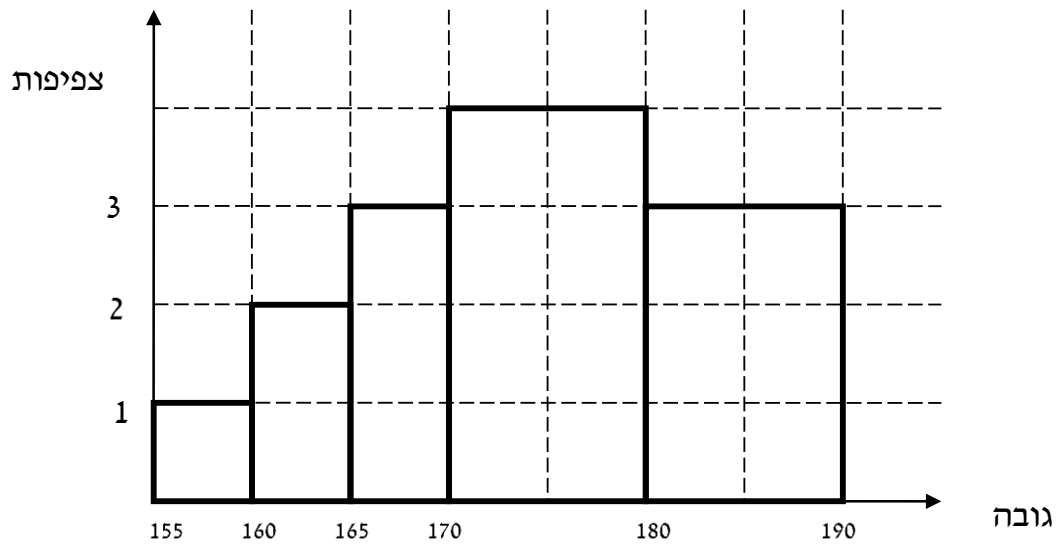
- (3) להלן התפלגות ההשכלה במקום עבודה מסוים:

מספר העובדים	השכלה
60	נמוכה
120	תיכונית
20	אקדמאית

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

- (4) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
- 7, 6, 8, 9, 10, 6, 4, 5, 8, 7, 6, 7, 6, 8, 9, 6, 7, 8, 5, 6.
- א. מהו המשתנה? האם הוא בדיד או רציף?
- ב. 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:



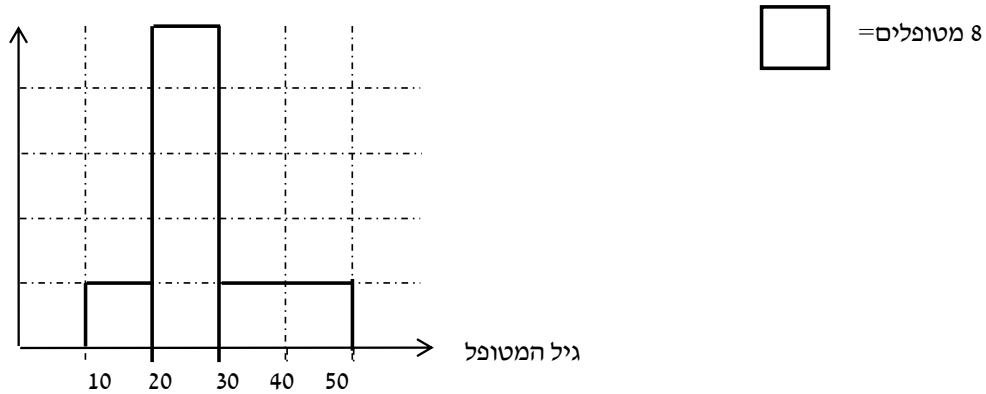
- מהו המשתנה הנחקר? האם הוא בדיד או רציף?
- תארו את הנתונים בטבלת שכיחויות במחלקות.
- הוסיפו שכיחות יחסית לטבלה.
- הוסיפו את הצפיפות של כל מחלקה לטבלה.
- מהי צורת ההתפלגות של הגבהים?

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

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

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

7) להלן גיל המטופלים של ד"ר 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:

(1) א. להלן טבלה: ב. עיין גרף מלא בסרטון הוידאו.

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

- (2) א. מקצוע מועדף. ב. 19.6%.
- (3) א. משתנה נחקר: השכלה, סוג: סדר. ב. עיין גרף מלא בסרטון הוידאו.

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

ב+ג. להלן טבלה :

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

5) א. גובה בס"מ, רציף.

ב+ג+ד. להלן טבלה : ה. אסימטרית.

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

- 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

פרק 3 - סטטיסטיקה תיאורית- סכימה

תוכן העניינים

1. כללי.....16

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:

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

א. $\sum_{i=1}^3 X_i$

ב. $\sum_{i=1}^5 Y_i$

ג. $\sum_{i=1}^4 X_i$

ד. $\left(\sum_{i=1}^4 X_i \right)^2$

ה. $\sum X_i$

ו. $\sum X_i Y_i$

ז. $\sum (X_i) \sum (Y_i)$

- (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

א. $\sum_{i=1}^4 y_i$

ב. $\sum_{i=1}^6 a$

ג. $\sum_{i=1}^6 x_i y_i$

ד. $\sum_{i=1}^6 (x_i + y_i)$

ה. $\sum_{i=1}^6 x_i + a$

- (3) קבעו לכל זהות האם היא נכונה:

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

ב. $\sum_{i=1}^n a = a \cdot n$

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

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

חשבו: $\sum_{i=1}^{10} (X_i - 4)^2$

תשובות סופיות:

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

סטטיסטיקה

פרק 4 - סטטיסטיקה תיאורית - מדדי מיקום מרכזי

תוכן העניינים

1. כללי 20

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.

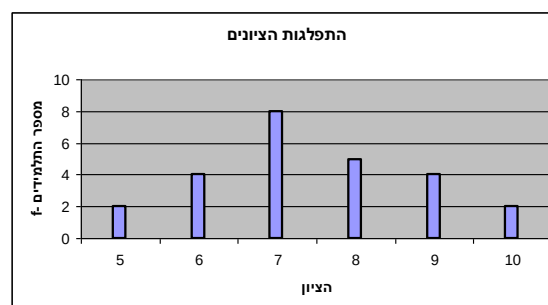
בטבלת שכיחויות בדידה

הערך שהשכיחות שלו היא הגבוהה ביותר.

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

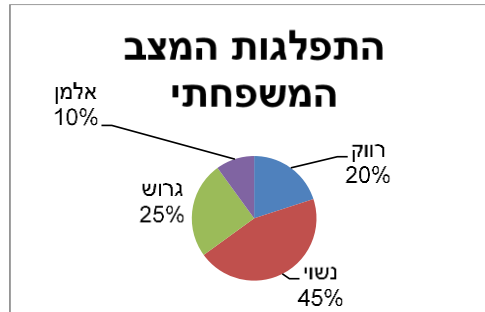
בדיאגרמת מקלות

שיעור ה- X של המקל הגבוה ביותר.



בעוגה

הערך של הפלח הגדול ביותר.



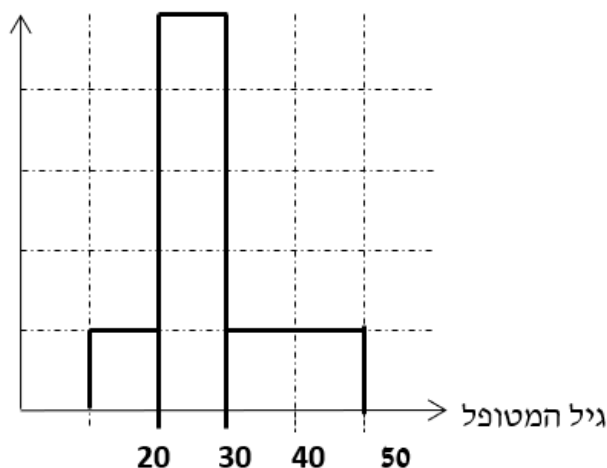
בטבלת שכיחויות במחלקות

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.

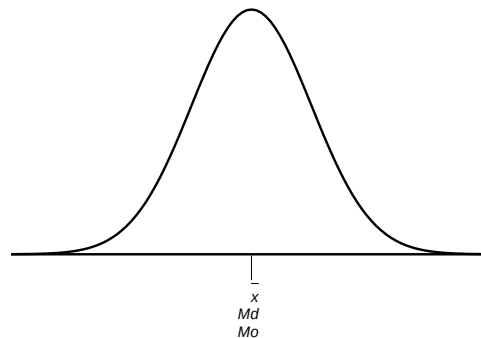
כללי

הממוצע רלבנטי רק למשתנה כמותי.

מדדי המיקום המרכזי בהתפלגויות המיוחדות:

בהתפלגות סימטרית פעמונית כל מדדי המרכז שווים זה לזה:

התפלגות סימטרית

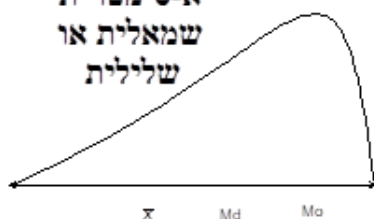


בהתפלגות סימטרית השכיח לא חייב להיות במרכז:

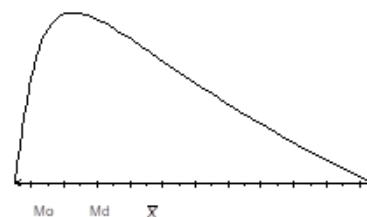
התפלגות U



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, 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 הטלויזיה שנספרו עבור כל משפחה ביישוב מסוים:

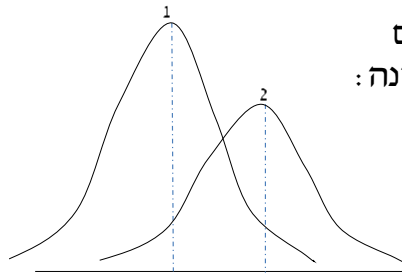
מספר משפחות	מספר מקלטים
0	22
1	28
2	18
3	22
4	10

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 טלויזיה היו רוכשים מקלט אחד.

(4) להלן התפלגות מספר המכוניות למשפחה ביישוב "הגורן":

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

א. כמה משפחות יש ביישוב?
ב. מה אחוז המשפחות ביישוב 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:

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

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

מספר טלוויזיות	מספר משפחות
0	28
1	62
2	
3	

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

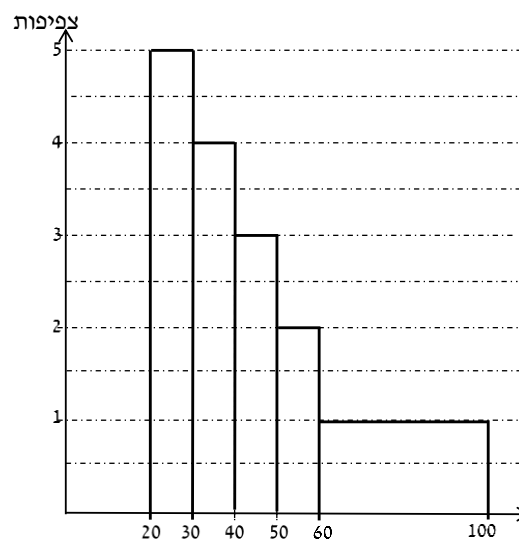
7) להלן התפלגות המשקל של קבוצה מסוימת בק"ג. מה הממוצע והחציון של ההתפלגות?

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.

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

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



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

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

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

תוכן העניינים

1. כללי 29

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
סה"כ		1430

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

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

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

- (2) להלן התפלגות מספר המכוניות למשפחה ביישוב "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.

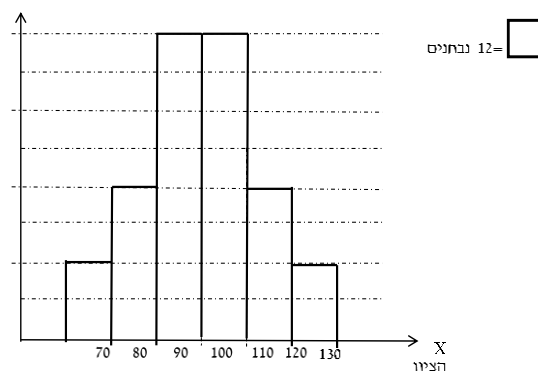
מספר חדרים	פרופורציה
1	0.1
2	0.2
3	0.4
4	0.15
5	

- 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

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

תוכן העניינים

1. כללי 32

סטטיסטיקה תיאורית – מדדי פיזור – טווח בין רבעוני:

רקע:

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

שלבים במציאת טווח בין-רבעוני במחלקות:

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

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:

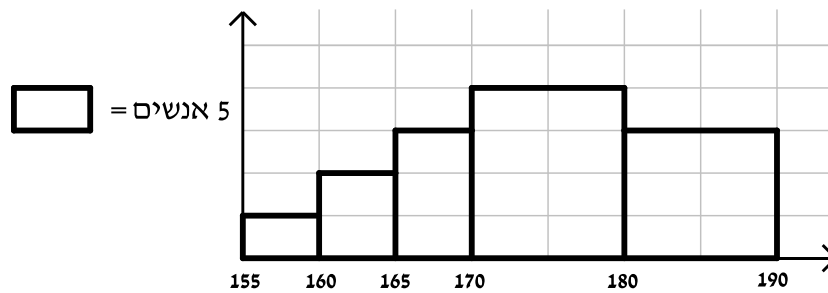
שאלות:

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

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

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

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

תוכן העניינים

1. כללי 34

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:

ממוצע	תפוקה	מספר פועלים
ממוצע	48	15
סטיית תקן	10	2

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

- (3) הגובה הממוצע של המתגייסים לצבא הוא 175 סנטימטר עם סטיית תקן של 10 סנטימטר. המשקל הממוצע הוא 66 ק"ג עם סטיית תקן של 8 ק"ג. 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

פרק 8 - סטטיסטיקה תיאורית-מדדי מיקום יחסי-אחוזונים במחלקות

תוכן העניינים

1. כללי 36

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 שמשכרים מתחת ל-5,000 ₪?

שאלות:

1) להלן התפלגות השכר (באלפי שקלים) בחברה:

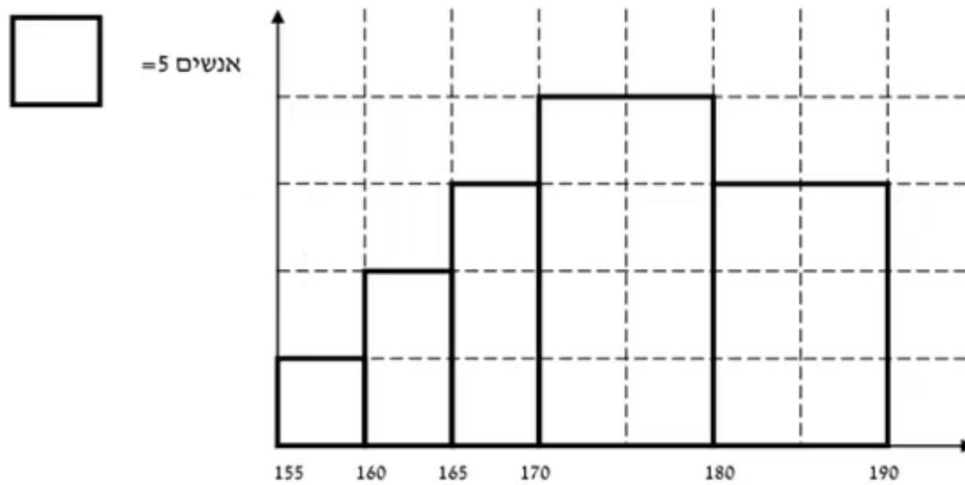
שכר - X	שכיחות מצטברת
6-10	48
10-15	100
15-20	120
20-30	132
30-60	136

- חשבו את המאון ה-60.
- מהו העשירון העליון?
- 20% מהמשכורות הגבוהות ביותר הן משכורות של הבכירים, מהי המשכורת המינימאלית לבכיר?
- מה אחוז האנשים שמשכרים מתחת ל-7,000 ₪?
- איזה אחוז מהעובדים משכרים מעל ל-25,000 ₪?
- איזה אחוז מהעובדים משכרים 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:

- (1) א. 13.23 ב. 22 ג. 17.2 ד. 8.82% ה. 7.36%

ו. 83.82%

(2) להלן טבלה:

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

- (3) א. 162.5 ב. 170 ג. 183.33 ד. 3% ה. 15%

ו. 55%

סטטיסטיקה

פרק 9 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

תוכן העניינים

1. כללי 39

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:

$$\begin{aligned} MR_y &= a \cdot MR + b \\ MO_y &= a \cdot MO + b \\ \bar{y} &= a \cdot \bar{x} + b \\ Md_y &= a \cdot Md_x + b \end{aligned} \quad \text{מדדי המרכז:}$$

$$\begin{aligned} R_y &= |a| R_x \\ S_y &= |a| S_x \\ S_y^2 &= a^2 S_x^2 \end{aligned} \quad \text{מדדי הפיזור:}$$

$$\begin{aligned} Y_p &= a \cdot X_p + b \\ Z_Y &= \frac{a}{|a|} Z_x \end{aligned} \quad \text{מדדי המיקום היחסי:}$$

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 ₪ לשעה עם סטיית תקן של 5 ₪ לשעה. הוחלט להעלות את כל המשכורות ב-10%, אך זה לא סיפק את העובדים ולכן הם קיבלו לאחר מכן תוספת של 2 ₪ לשעה. מה הממוצע ומהי השונות של השכר לשעה לאחר כל השינויים.
- (3) במבחן מסוים הציון החציוני היה 73, טווח הציונים היה 40 נקודות והעשירון העליון היה הציון 87. כיוון שהציונים בבחינה היו נמוכים, המורה החליט לתת פקטור של 4 נק' לכל התלמידים. חשבו את המדדים לאחר הפקטור.
- (4) דגמו מקו ייצור 50 קופסאות של גפרורים. בדקו בכל קופסא בה יש 40 גפרורים את כמות הגפרורים הפגומים. התקבל שבממוצע יש 3 גפרורים פגומים בקופסא, עם סטיית תקן של 1.5 גפרורים. מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקינים בקופסא?
- (5) חברת בזק הציעה את ההצעה הבאה: שלושים שקלים דמי מנוי חודשיים קבועים וכן 10 אגורות לכל דקה של שיחה יוצאת. אדם בדק במשך שנה את דקות השיחות היוצאות שלו, וקיבל שבממוצע חודשי יש לו 600 דקות שיחות יוצאות עם שונות של 2500 דקות רבועות, כמו כן בחודש ינואר ציון התקן היה 2. חשבו את המדדים הללו עבור חשבון הטלפון החודשי של אותו אדם בשקלים אם היה משתמש בחבילה המוצעת לו על ידי בזק.
- (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

פרק 10 - סטטיסטיקה תיאורית - שאלות מסכמות

תוכן העניינים

1. כללי 42

סטטיסטיקה תיאורית – שאלות מסכמות:

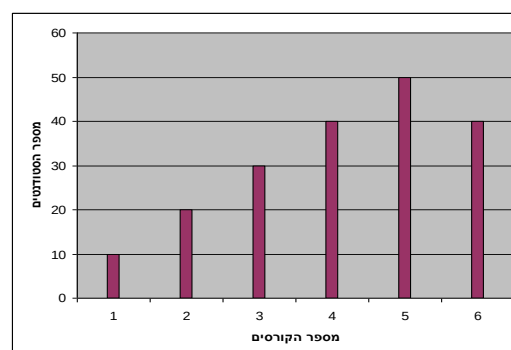
שאלות:

(1) בדקו עבור 5 תלמידים את המשקל שלהם:

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

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

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



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

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

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

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

א. השלימו את השכיחויות החסרות בטבלה.

ב. חשבו את הציון החציוני, השכיח.

ג. חשב שונות וסטיית תקן להתפלגות הציונים.

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

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

א. מה אחוז האנשים עם רמת שביעות רצון נמוכה?

ב. מה המשתנה הנחקר ומאיזה סוג הוא?

ג. מהי הדרך הגרפית המתאימה ביותר לתיאור הנתונים?

i. היסטוגרמה.

ii. דיאגרמת מקלות.

iii. דיאגרמת עוגה.

ד. חשבו את המדדים הבאים:

i. טווח.

ii. שכיח.

iii. 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		

- השלימו את הטבלה.
- חשבו את החציון, השכיח והממוצע של התפלגות מס' שעות העבודה בחברה.
- מה סטיית התקן של מספר שעות העבודה?
- מה העשירון העליון של ההתפלגות?
- איזה אחוז מהעובדים עובדים מעל 45 שעות בשבוע?
- מה ציון התקן של רינה, שעובדת 30 שעות בשבוע?
- 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

- מה אחוז האנשים ששלחו פחות מ-80 מסרונים בחודש?
- מה אחוז האנשים ששלחו בין 50 ל-120 מסרונים?
- הוחלט להעניק מתנה עבור $\frac{1}{4}$ מהלקוחות שמשתמשים במספר הרב ביותר של מסרונים בחודש. החל מאיזה כמות של מסרונים תחולק המתנה?
- 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

- שרטטו היסטוגרמה לתיאור התפלגות זמן ביצוע המשימה.
- מתוך ההיסטוגרמה שבנית בסעיף א', מהי צורת ההתפלגות?
- חשבו את השכיח והחציון של ההתפלגות.
- 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.

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

- השלימו את הטבלה.
- מהו הממוצע והחציון של ההתפלגות?
- מהו הציון ש-40% מהתלמידים קיבלו מעליו? באיזה אחוזון מדובר?
- 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.

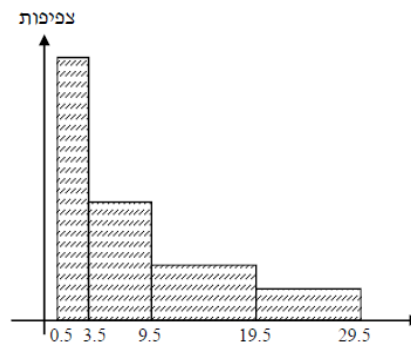
$f(x)$	x
10	1
20	2
30	3
40	4
50	5
40	6
190	סה"כ

- (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

פרק 11 - מקדם המתאם (מדד קשר) הלינארי ומובהקותו

תוכן העניינים

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

מקדם המתאם (מדד קשר) הלינארי ומובהקותו

מדד הקשר הלינארי (פירסון) – מבוא

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

דוגמה (פתרון בהקלטה) :

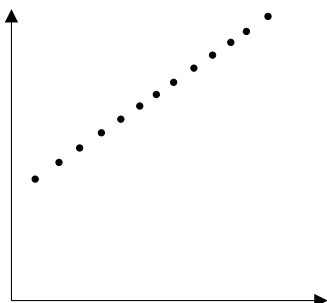
בבניין 8 דירות בדקו לכל דירה את מספר החדרים שלה וכמו כן את מספר הנפשות הגרות בדירה. להלן התוצאות שהתקבלו :

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

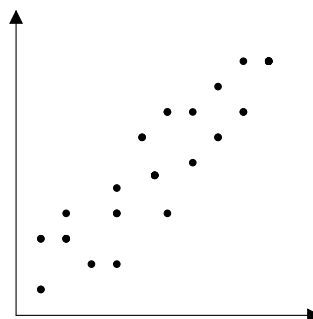
- (1) כמה תצפיות ישנן בדוגמה?
- (2) כמה משתנים ישנם בדוגמה, מי הם?
- (3) שרטטו לנתונים דיאגרמת פיזור.
- (4) מי המשתנה התלוי ומיהו המשתנה הבלתי תלוי?

דיאגרמות פיזור לקשר בין משתנים וניתוחם

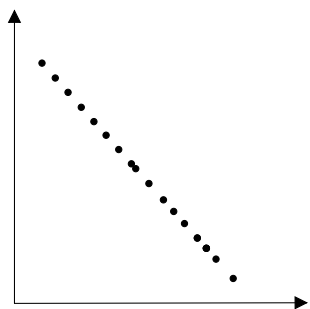
קשר לינארי חיובי מלא



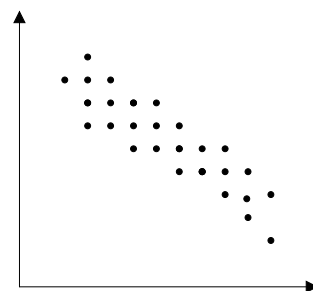
קשר לינארי חיובי חלקי



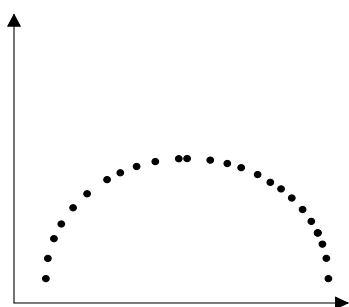
קשר לינארי שלילי מלא



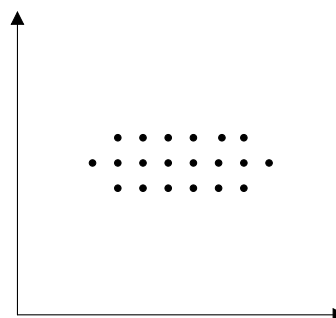
קשר לינארי שלילי חלקי



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-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 שמשתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט ואילו מתאם שלילי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לרדת אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט. ככל שמקדם המתאם קרוב לאפס עוצמת הקשר יותר חלשה וככל שהמדד רחוק יותר מהאפס העוצמה יותר חזקה. לסיכום, מקדם המתאם בודק את עוצמת הקשר 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: ρ - פרמטר המאפיין את עוצמת הקשר הלינארי באוכלוסייה וכיוונו בין שני המשתנים הנחקרים. כאשר:

r - מדד הקשר הלינארי במדגם שמהווה אומד לפרמטר ρ .

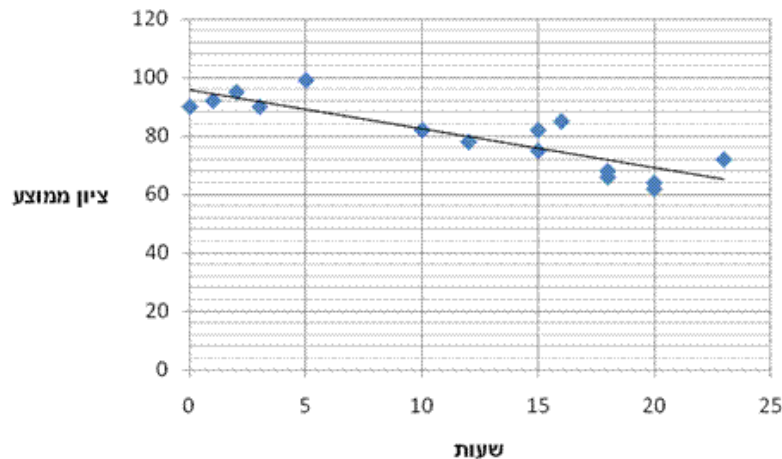
קיומו של מתאם בין שני משתנים אינו מצביע על סיבתיות בהכרח. למשל, אם נמצא מתאם חיובי בין כמות הסוכרזית שאדם אוכל לבין במשקל שלו אין זה אומר שהסיבה להשמנה היא הסוכרזית. מדד הקשר של פירסון הוא מדד קשר סימטרי, כלומר אם נחליף את X ב- Y התוצאה תהיה זהה.

דוגמה (פתרון בהקלטה):

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

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:

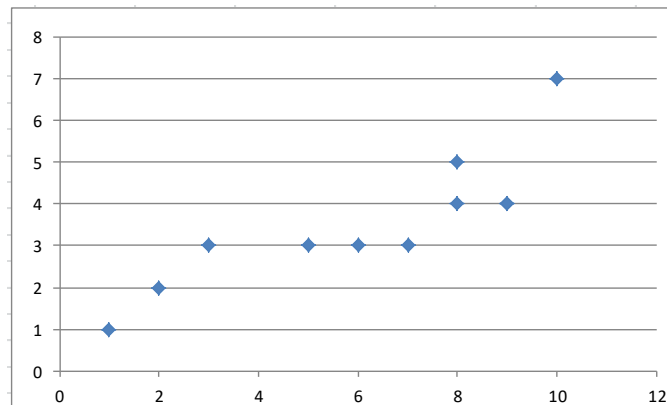
בחודש ינואר ככל שהטמפרטורה היומית נוטה לרדת, כך כמות המשקעים נוטה _____.

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:

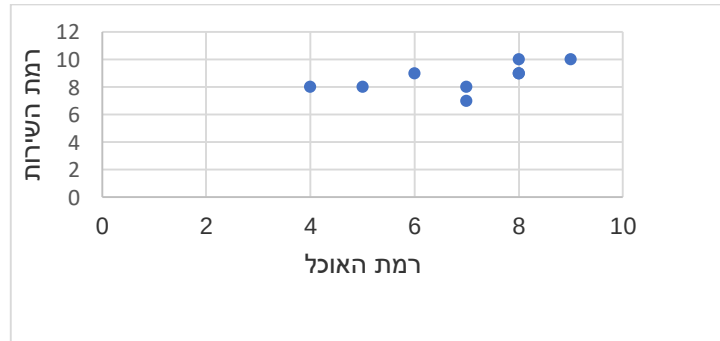
א. אין קשר בין הטמפרטורה בטורונטו לבין הטמפרטורה בסידני בימים שנדגמו.

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.

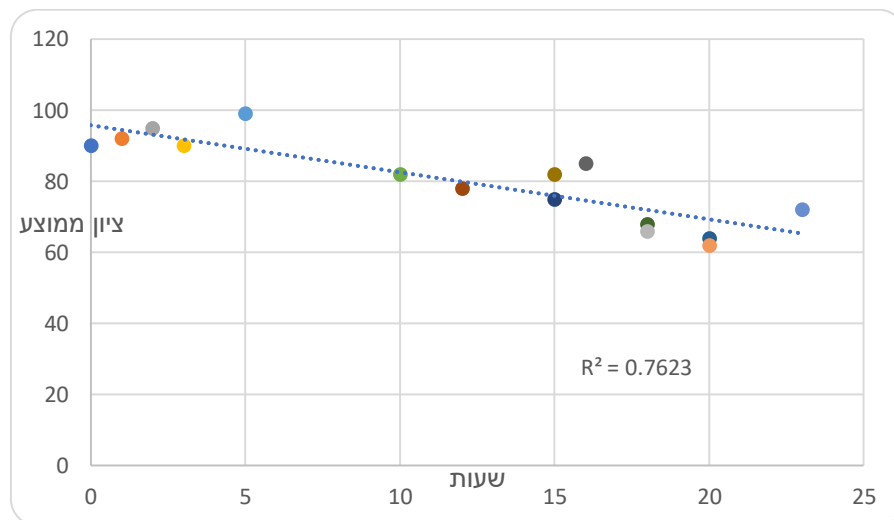
- 7) בסקר שביעות רצון שנערך בבית הקפה "פת לחם" התבקשו הלקוחות לדרג את מידת שביעות הרצון שלהם (בסולם 1-10) בשני נושאים: רמת האוכל ורמת השירות.



מה יהיה ערכו של מקדם המתאם (r)?

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

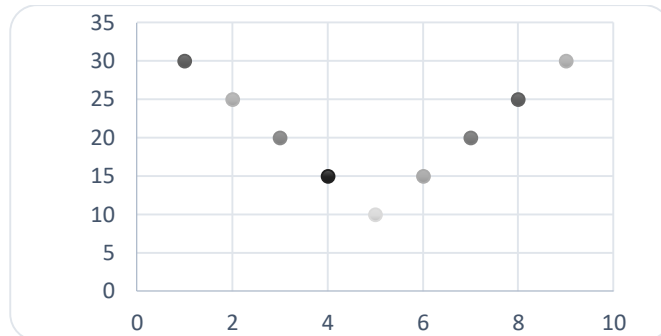
- 8) חוקר רצה לאפיין את הקשר בין מספר השעות בשבוע שסטודנט מקדיש לבילויים לבין הציון הממוצע שלו בסוף הסמסטר. לשם כך הוא אסף נתונים של 15 סטודנטים ויצר דיאגרמת פיזור.



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

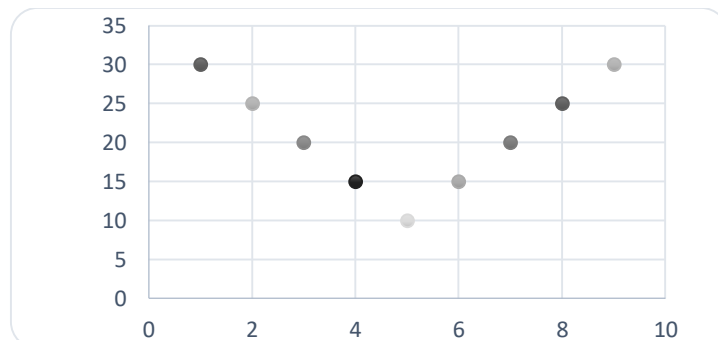
- א. ככל שמבליים יותר הציון נוטה לרדת.
- ב. אין קשר בין שעות הבילוי לציון.
- ג. ככל שמבליים פחות הציון נוטה לרדת.
- ד. ככל שהציון נוטה לרדת הסטודנט מבלה פחות.

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.

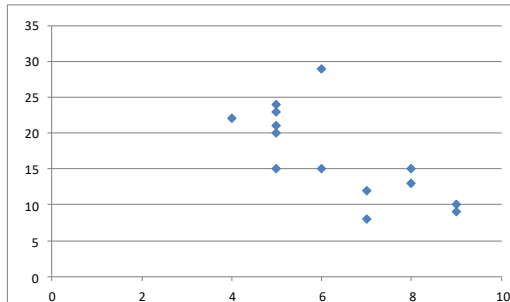
10) התרשים הבא מתאר דיאגרמת פיזור.



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)

Y - (משתנה תלוי).

במדגם התקבל $r^2 = 0.52$.

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

א. -0.52

ב. 0.72

ג. -0.72

ד. 0.52

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

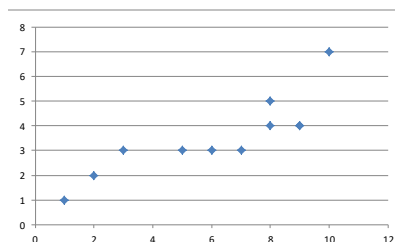
א. הערכים של המשתנים הם חיוביים.

ב. עבור כל תצפית ערך של משתנה אחד שווה לערך של המשתנה השני.

ג. הקשר הלינארי הוא בעוצמה חזקה.

ד. אף אחת מהתשובות לא בהכרח נכונה.

13) להלן דיאגרמת פיזור:



מה יהיה מקדם המתאם בין שני המשתנים?

א. 1

ב. 0.85

ג. 0.15

ד. 0

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

א. קיימת נוסחה לינארית הקושרת בין כל התצפיות.

ב. לא קיים קשר בין שני המשתנים.

ג. ככל שמשתנה אחד נוטה לרדת גם לשני יש נטייה לרדת.

ד. קיים קשר בין שני המשתנים, אך לא ניתן לדעת מאיזה סוג.

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:
ב. ספורט ולשון.

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.
ב. קטן.
ב. ראה גרף בפתרון וידאו.
ד. מקדם המתאם לא היה משתנה.
ב. לא ישפיע על מקדם המתאם.

(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 וננתח אותן (הדיאגרמות המלאות בסרטון).

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

מתאם חיובי מלא (מקדם מתאם 1):

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

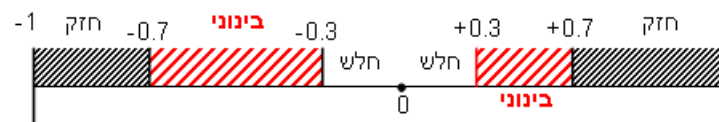
מתאם חיובי חלקי:

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

מתאם שלילי חלקי:

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

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



מקדם המתאם יסומן באות r .

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

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

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

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

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

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.

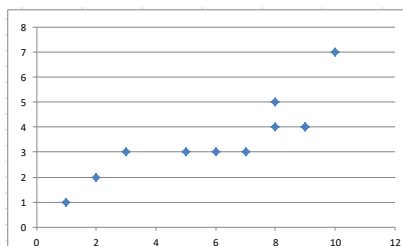
שאלות רב-ברירה:

- (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

פרק 12 - יסודות ההסתברות

תוכן העניינים

1. כללי 65

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:

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

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

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

פרק 13 - פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים

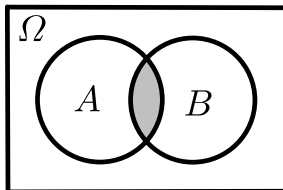
תוכן העניינים

1. כללי 69

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

רקע:

פעולת חיתוך:

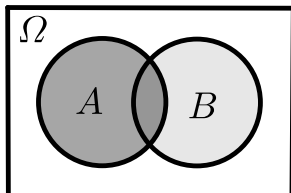


נותנת את המשותף בין המאורעות הנחתכים.
חיתוך בין המאורע A למאורע B יסומן כך: $A \cap B$.
מדובר בתוצאות שנמצאות ב- A וגם ב- B .

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:



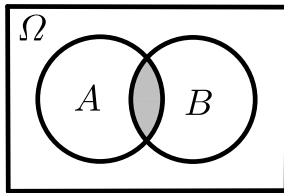
נותנת את כל האפשרויות שנמצאות לפחות באחת מהמאורעות, ומסומנת: $A \cup B$.
הפעולה נותנת את אשר נמצא ב- A או ב- B .
כלומר, לפחות אחד מהמאורעות קורה.

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:

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

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

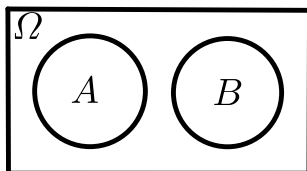
$$P(A \cap B) = 1 - P(\bar{A} \cup \bar{B})$$

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

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.

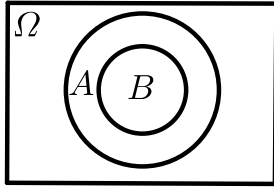
ההסתברות של חיתוך המאורעות היא אפס: $P(A \cap B) = 0$.

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 מוכללים:

נתונים שני מאורעות A ו- B , השונים מאפס.
נאמר שהמאורע B מוכלל 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
מוכללות בתוך מאורע A .

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

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

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

$$\begin{aligned} A &= \{2, 4, 6\} \\ B &= \{2, 4\} \end{aligned} \quad \text{למשל:}$$

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 וסמנו 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- Ω את מרחב המדגם וב- ϕ קבוצה ריקה.
- נתון כי A הינו מאורע בתוך מרחב המדגם.
- להלן מוגדרים מאורעות שפתרונם הוא Ω או ϕ או A .
- 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. $\overline{A} \cap B$

ד. $\overline{A} \cup \overline{B}$

ה. $\overline{\overline{A}}$

(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 "עתיד" תעבור את אחוז החסימה בהסתברות של 0.20. בהסתברות של 76% שתי המפלגות לא תעבורנה את אחוז החסימה.

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

ב. מה ההסתברות ששתי המפלגות תעבורנה את אחוז החסימה?

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

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

א. איזה אחוז מהעובדים הם גברים אקדמאיים?

ב. איזה אחוז מהעובדים הם גברים או אקדמאיים?

ג. איזה אחוז מהעובדים הם נשים לא אקדמאיות?

(9) הסיכוי של מניה A לעלות הנו 0.5 ביום מסוים והסיכוי של מניה B לעלות ביום מסוים הנו 0.4. בסיכוי של 0.7 לפחות אחת מהמניות תעלה ביום מסוים. 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: $P(A) = 0.3$, $P(B) = 0.25$, $P(A \cup B) = 0.49$

- חשבו את הסיכוי ל- $P(A \cap B)$.
- האם A ו- B מאורעות זרים?
- 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:

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

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

- האם יתכן ש- $P(A \cup B) = 0.4$?
- האם יתכן ש- $P(A \cup B) = 0.6$?
- אם A ו- B זרים מה הסיכוי $P(A \cup B)$?
- 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.

- מה אחוז האזרחים להם חשבון בבנק לאומי בלבד?
- מה ההסתברות שאזרח כלשהו יחזיק חשבון בבנק פועלים ולאומי אבל לא בבנק מזרחי?
- מה ההסתברות שלאזרח יהיה חשבון בפועלים או במזרחי אבל לא בבנק לאומי?
- מה אחוז האזרחים שיש להם חשבון בנק אחד בלבד?
- מה אחוז האזרחים שיש להם בדיוק חשבון בשני בנקים בלבד?
- מה ההסתברות שלאזרח בוגר אין חשבון בנק באף אחד מהבנקים הללו?
- 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 "ישראכרט", 20% מחזיקים כרטיס "אמריקן אקספרס", 15% מחזיקים כרטיס ויזה וגם ישראכרט, 8% מחזיקים כרטיס ישראכרט 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 ישראכרט 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. $\bar{B} = 5, 6, 7, 8, 9, B = 0, 1, 2, 3, 4, A = 0, 2, 4, 6, 8$
 $A \cup B = 0, 2, 4, 6, 8, 1, 3, A \cap B = 0, 2, 4$
- 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) $\bar{\bar{A}} = A, A \cap \phi = \phi, A \cup \phi = A, A \cap \Omega = A, A \cup \Omega = \Omega$
 $A \cup \bar{A} = \Omega, \bar{\phi} = \Omega, A \cap \bar{A} = \phi$
- (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 לכל היותר 1.7. ד. $\bar{A} \cup \bar{B}$: לכל 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

פרק 14 - קומבינטוריקה - כלל המכפלה

תוכן העניינים

1. כללי 78

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 :
 - א. הטלת קובייה פעמים.
 - ב. מספר תלת ספרתי.
 - ג. בחירת בן ובת מכתה שיש בה שבעה בנים ועשר בנות.
 - ד. חלוקת שני פרסים שונים לעשרה אנשים שונים כאשר אדם לא יכול לקבל יותר מפרס אחד.
- (2) במסעדה מציעים ארוחה עסקית.
 - א. 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.
 - ב. כמה ארוחות שונות ניתן להרכיב בעזרת התפריט הזה?
 - i. בארוחה סלט ירקות, לזניה בשרית ולימונדה.
 - ii. בארוחה סלט, לזניה ותה.
- (3) בוחרים באקראי מספר בין חמש ספרות. חשבו את ההסתברויות הבאות :
 - א. המספר הוא זוגי.
 - ב. במספר כל הספרות שונות.
 - ג. במספר כל הספרות זהות.
 - ד. במספר לפחות שתי ספרות שונות.
 - ה. במספר לפחות שתי ספרות זהות.
 - ו. המספר הוא פלינדרום (מספר הנקרא מימין ומשמאל באות הצורה).
- (4) חישה אנשים אקראיים נכנסו למעלית בבניין בן 8 קומות. חשבו את ההסתברויות הבאות :
 - א. כולם ירו בקומה החמישית.
 - ב. כולם ירדו באותה קומה.
 - ג. כולם ירדו בקומה אחרת.
 - ד. 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:
- א. חבר כנסת יכול למלא יותר מתפקיד אחד.
 - ב. חבר כנסת לא יכול למלא יותר מתפקיד אחד.
- (6) מטילים קובייה 4 פעמים.
- א. מה ההסתברות שכל התוצאות תהינה זהות?
 - ב. מה ההסתברות שכל התוצאות תהינה שונות?
 - ג. מה ההסתברות שלפחות שתי תוצאות תהינה זהות?
 - ד. מה ההסתברות שלפחות שתי תוצאות תהינה שונות?
- (7) יש ליצור מילה בת חמש אותיות, לא בהכרח עם משמעות מאותיות ה-ABC (26 אותיות).
- א. מה ההסתברות שבמילה שנוצרה אין האותיות A, D ו-L?
 - ב. מה ההסתברות שבמילה שנוצרה כל האותיות זהות?
 - ג. מה ההסתברות שבמילה שנוצרה לפחות שתי אותיות שונות זו מזו?
 - ד. 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).
- א. בקוד אין את הספרה 5.
 - ב. בקוד מופיעה הספרה 3.
 - ג. 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. 36 ב. 900 ג. 70 ד. 90
- (2) א. 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. 3375 ב. 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

פרק 15 - קומבינטוריקה- תמורה - סידור עצמים בשורה

תוכן העניינים

1. כללי 82

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?
שני ספרים מתוך ה-10 הם ספרים בסטטיסטיקה.
ב. מה ההסתברות שאם נסדר את הספרים באקראי, הספרים בסטטיסטיקה יהיו צמודים זה לזה?
ג. מה ההסתברות שהספרים בסטטיסטיקה לא יהיו צמודים זה לזה?
ד. מה ההסתברות שהספרים בסטטיסטיקה יהיו בקצות המדף (כל ספר בקצה אחר)?
- (5) אדם יצר בנגן שלו פלייליסט (רשימת השמעה) של 12 שירים שונים. 4 בשפה העברית, 5 באנגלית ו-3 בצרפתית. האדם הריץ את הפלייליסט באקראי.
א. מה ההסתברות שכל השירים באנגלית יופיעו כשירים הראשונים כמקשה אחת?
ב. מה ההסתברות שכל השירים באנגלית יופיעו ברצף (לא חובה ראשונים)?
ג. מה ההסתברות ששירים באותה השפה יופיעו ברצף (כלומר כל השירים באנגלית ברצף, כל השירים בעברית ברצף וכך גם השירים בצרפתית)?

- 6) 4 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

פרק 16 - קומבינטוריקה - דגימה סידורית ללא החזרה ועם החזרה

תוכן העניינים

1. כללי 85

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:

בחרים 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:

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

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 (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

פרק 17 - קומבינטוריקה - דגימה ללא סדר וללא החזרה

תוכן העניינים

1. כללי 87

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:

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

הערות:

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

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

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

שאלות:

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

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

א. כל המזכירות בוועד יהיו ממסלול "מדעי ההתנהגות".

ב. כל המזכירות בוועד יהיו מאותו המסלול.

ג. מכל מסלול תבחר לפחות מזכירה אחת.

(7) 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) א. $\binom{2n}{n}^2$ ב. $\sum_{i=1}^n \binom{2n}{i}^2$

סטטיסטיקה

פרק 18 - קומבינטוריקה - שאלות מסכמות

תוכן העניינים

1. כללי 90

קומבינטוריקה – שאלות מסכמות:

שאלות:

- (1) 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.
 - ג. אין תפקידים שונים בוועד.

- (2) במשרד 30 עובדים, יש לבחור ארבעה עובדים למשלחת לחו"ל. בכמה דרכים ניתן להרכיב את המשלחת?
 - א. במשלחת ארבע משימות שונות שיש למלא וכל עובד יכול למלא יותר ממשימה אחת.
 - ב. כמו בסעיף א' רק הפעם עובד לא יכול למלא יותר ממשימה אחת.
 - ג. 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).
 - א. כמה קודים שונים ניתן להרכיב?
 - ב. כמה קודים שונים ניתן להרכיב אם הקוד מתחיל בספרה ונגמר בספרה?
 - ג. כמה קודים ניתן להרכיב אם הספרות חייבות להיות צמודות זו לזו?
 - ד. בכמה קודים הספרות לא מופיעות ברצף?

- (4) בארונות 4 מגירות. ילד התבקש על ידי אמו לסדר 6 משחקים בארונות. הילד מכניס את המשחקים באקראי למגירות השונות. כל מגירה יכולה להכיל את כל המשחקים יחד.
 - א. מה ההסתברות שהילד יכניס את כל המשחקים למגירה העליונה?
 - ב. מה ההסתברות שהילד יכניס את כל המשחקים לאותה מגירה?
 - ג. מה ההסתברות ש"דומינו" יוכנס למגירה העליונה ויתר המשחקים לשאר המגירות.
 - ד. 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 4 מקומות ישיבה נגד כיוון הנסיעה.

חלון	מעבר		

- 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

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

סטטיסטיקה

פרק 19 - הסתברות מותנית-במרחב מדגם אחד

תוכן העניינים

1. כללי 97

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: $P(A|B)$.

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

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

פרק 20 - הסתברות מותנית - מרחב לא אחיד

תוכן העניינים

1. כללי 100

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:
- א. התלמיד עבר בסטטיסטיקה, מה ההסתברות שהוא עבר בכלכלה?
 - ב. התלמיד עבר בכלכלה, מה ההסתברות שהוא עבר בסטטיסטיקה?
 - ג. התלמיד עבר בכלכלה, מה ההסתברות שהוא נכשל בסטטיסטיקה?
 - ד. התלמיד נכשל בסטטיסטיקה, מה ההסתברות שהוא נכשל בכלכלה?
 - ה. התלמיד עבר לפחות מבחן אחד, מה ההסתברות שהוא 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.
- א. איזה אחוז מהתושבים הבוגרים רשומים אצל שתי החברות?
 - ב. נבחר אדם שרשום אצל חברת "סופט", מה ההסתברות שהוא רשום גם אצל חברת "בל"?
 - ג. אם אדם לא רשום אצל חברת "בל", מה ההסתברות שהוא כן רשום בחברת "סופט"?
 - ד. 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.
- א. מה ההסתברות שיש מקום בשעה 08:00 רק בחניון הגדול של המכללה?
 - ב. ידוע שבחניון הקטן יש מקום בשעה 08:00, מה הסיכוי שבחניון הגדול יש מקום?
 - ג. אם בשעה 08:00 בחניון הגדול אין מקום, מה ההסתברות שבחניון הקטן יהיה מקום?
 - ד. 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

פרק 21 - דיאגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה

תוכן העניינים

1. כללי 104

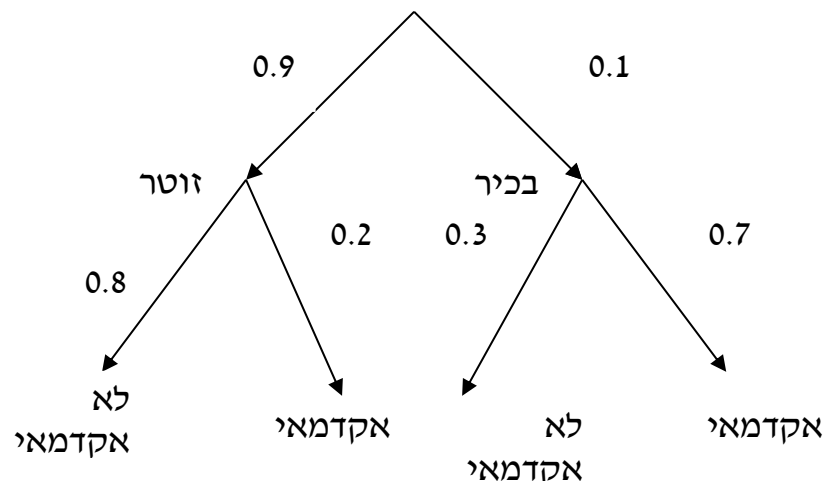
דיאגרמת עצים – נוסחת בייס ונוסחת ההסתברות השלמה:

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 20% הם אקדמאים ומבין הזוטרים 80% 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

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

נוסחת ההסתברות השלמה:

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

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

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.
- מה ההסתברות להתקבל לעבודה?
 - מועמד לא התקבל לעבודה. מה ההסתברות שהוא נכשל במבחן הראשון?
 - 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.
- מה אחוז החולים בשפעת בקרב האוכלוסייה הקשישה?
 - 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.
- מה הסיכוי שהאנייה תתגלה ע"י הרדאר?
 - אם האנייה התגלתה ע"י הרדאר, מה ההסתברות שהיא נמצאת באזור C?
 - 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.
- מה ההסתברות שאדם הגיע לקליניקה וגילו אצלו את סימפטום X?
 - אם התגלה אצל אדם סימפטום X, מה ההסתברות שהוא חולה במחלה A?
 - אם לאדם יש את סימפטום X, מה ההסתברות שהוא חולה במחלה A?
 - אם לא 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 מישו 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

פרק 22 - תלות ואי תלות בין מאורעות

תוכן העניינים

1. אי תלות בין מאורעות (מורחב) 109

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 שמההקשר 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 יחד?
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 שיווק, 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

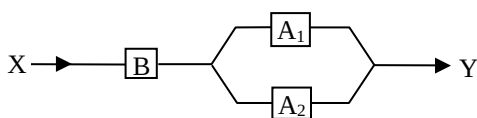
בהסתברויות שונות, אך באופן בלתי תלוי זה בזה.

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) 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

פרק 23 - שאלות מסכמות בהסתברות

תוכן העניינים

1. כללי 113

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.
 - ד. אם ההסתברות שרק המכונית החדשה שלה היא מתוצרת אירופאית?
 - ה. 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 ייצאו 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 ייצאו רק לבאולינג וביום למחרת לא ייצאו 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?

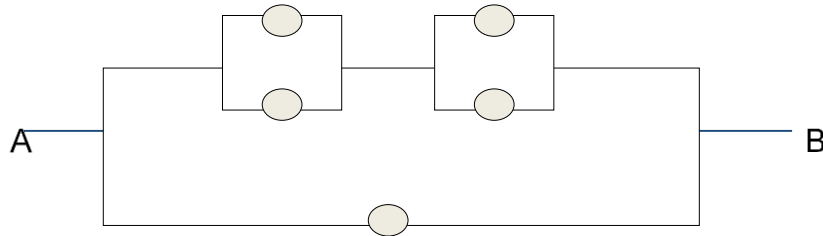
(6) בתיבה 10 מטבעות, מתוכם 7 מטבעות רגילים (ראש, זנב) ו-3 מטבעות שבשני צדדיהם טבוע ראש. אדם בוחר באקראי מטבע ומטיל אותו פעמיים. נסמן ב-A את ההטלה הראשונה ראש וב-B את ההטלה השנייה ראש.

- 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

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

(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.
- (8) ד. i. 0.478 ד. ii. 0.15 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

פרק 24 - המשתנה המקרי הבדיד - פונקציית ההסתברות

תוכן העניינים

1. כללי 118

המשתנה המקרי הבדיד – פונקציית ההסתברות:

רקע:

משתנה מקרי בדיד:

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

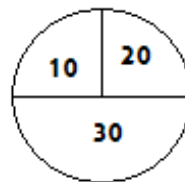
מתארים את המשתנה המקרי על ידי פונקציית ההסתברות.

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 פונקציות 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) 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

פרק 25 - המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן

תוכן העניינים

1. כללי 122

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.

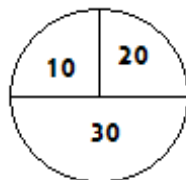
$$\text{מגדירים שונות באופן הבא: } V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = \sigma^2$$

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

פרק 26 - המשתנה המקרי הבדיד - טרנספורמציה לינארית

תוכן העניינים

1. כללי 126

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 והשונות של הרווח במשחק?

פתרון (בהקלטה):

חישבנו קודם ש: $V(X) = 68.75 = \sigma^2$, $E(X) = 22.5 = \mu$.

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 הנ"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: $k=1,2,3,4$: $P(X=k) = \frac{k}{A}$.

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

פרק 27 - התפלגויות בדידות מיוחדות - התפלגות בינומית

תוכן העניינים

1. כללי 129

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 :

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

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

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

$$E(X) = np \quad \text{תוחלת:}$$

$$V(X) = npq \quad \text{שונות:}$$

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

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

דוגמה (פתרון בהקלטה):

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

- א. 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 ששלושה יעבדו במדגם?
 - מה ההסתברות שלפחות אחד יהיה מובטל?
 - מה תוחלת ומהי השונות של מספר המובטלים במדגם?
- (2) על פי נתוני משרד התקשורת ל-70% מהאוכלוסייה יש 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 מס' 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. 0.5314 ב. 0.2335 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

פרק 28 - התפלגויות בדידות מיוחדות - התפלגות גיאומטרית

תוכן העניינים

1. כללי 133

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 בניסוי 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.

- א. מהי ההתפלגות של מספר הכדורים שהוצאו?
- ב. מה ההסתברות שהוצאו בדיוק 5 כדורים?
- ג. מה ההסתברות שהוצאו יותר מ-5 כדורים?
- ד. אם הוצאו יותר מ-3 כדורים, מה הסיכוי שהוצאו בדיוק 5 כדורים?
- ה. 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 עוגת 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

פרק 29 - המשתנה המקרי הבדיד - שאלות מסכמות

תוכן העניינים

1. כללי 136

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

תשובות סופיות:

- (1) א. תוחלת: 2, סטיית תקן: 1. ב. תוחלת: 0, סטיית תקן: 2.
ג. תוחלת: 4.5, סטיית תקן: לא ניתן.
- (2) א. 0.03. ב. תוחלת: 0.15, שונות: 0.1875.
ג. 0.1328.
- (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. ג. 0.282. ד. 0.282.
- (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.

סטטיסטיקה

פרק 30 - משתנה דו-מימדי בדיד - פונקציית הסתברות משותפת

תוכן העניינים

1. כללי 140

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:

$$p(x=0, y=0) = 0.05$$

$$p(x=0, y=1) = 0$$

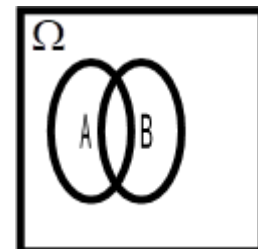
$$p(x=1, y=0) = 0.15$$

$$p(x=1, y=1) = 0.05$$

$$p(x=2, y=0) = 0$$

$$p(x=2, y=1) = 0.75$$

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 פונקציית 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 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 ביישוב 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

פרק 31 - משתנה דו מימדי בדיד - מתאם בין משתנים

תוכן העניינים

1. כללי 146

משתנה דו מימדי בדיד – מתאם בין משתנים:

רקע:

נרצה לבדוק את מידת ההתאמה הלינארית בין שני המשתנים על ידי מקדם המתאם הלינארי שמסומן ב- ρ .
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: $y = ax + b$.

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 שכלל משתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט ואילו מתאם שלילי חלקי אומר שכלל משתנה 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.
- א. בנו את פונקציית ההסתברות המשותפת ואת פונ' ההסתברות השולית.
 - ב. האם המשתנים הינם בלתי תלויים?
 - ג. ידוע שהסטודנט ניגש ליותר ממבחן אחד, מה ההסתברות שהוא נכשל בפחות משלושה מבחנים?
 - ד. האם המתאם בין X ל- Y מלא או חלקי? חיובי או שלילי? הסבירו ללא חישוב.
 - ה. חשבו את מקדם המתאם בין X לבין Y .
 - ו. 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.
- א. בנו את פונקציית ההסתברות המשותפת של X ו- Y ואת פונקציות ההסתברות השוליות.
 - ב. האם X ו- Y הם משתנים בלתי תלויים?
 - ג. מהו מקדם המתאם בין X ל- Y . האם המשתנים מתאומים?
 - ד. אם בשתי ההטלות הראשונות יצא בדיוק עץ אחד, מה ההסתברות שבשתי ההטלות האחרונות יצאו שני עצים?
 - ה. 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.
- א. בנו את פונקציית ההסתברות המשותפת.
 - ב. 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 פונקציות 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

פרק 32 - המשתנה המקרי הדו ממדי - קומבינציות ליניאריות

תוכן העניינים

1. כללי 153

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:
- $$\mu_X = 80, \sigma_X = 15, \mu_Y = 70, \sigma_Y = 20, \text{cov}(X, Y) = 200$$
- א. מצאו את התוחלת והשונות של סכום המשתנים.
- ב. מצאו את התוחלת והשונות של X ו- Y .
- ג. 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	
$P(X)$				

- א. השלימו את ההסתברויות החסרות.
 ב. האם המשתנים תלויים?
 ג. האם המשתנים בלתי מתואמים?
 ד. חשבו את השונות המשותפת.
 ה. חשבו את התוחלת והשונות של סכום המשתנים.
 ו. 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.
 א. חשבו את השונות המשותפת בין הציון הכמותי לציון המילולי.
 ב. חשבו את התוחלת והשונות של סכום הציונים בחלק הכמותי ובחלק המילולי.
 ג. חשבו את התוחלת והשונות של הפרש הציונים בין החלק הכמותי לחלק המילולי.
 ד. 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) $-\frac{n}{36}$.