

מבוא להסתברות וסטטיסטיקה



תוכן העניינים

1.	יסודות ההסתברות	1
2.	פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים	5
3.	קומבינטוריקה - כלל המכפלה	14
4.	קומבינטוריקה - תמורה - סידור עצמים בשורה	18
5.	קומבינטוריקה - סידור עצמים במעגל	21
6.	קומבינטוריקה - דגימה סידורית ללא החזרה ועם החזרה	24
7.	קומבינטוריקה - דגימה ללא סדר וללא החזרה	26
8.	כלל ההכלה וההפרדה	29
9.	קומבינטוריקה - שאלות מסכמות	34
10.	הסתברות מותנית-במרחב מדגם אחיד	41
11.	הסתברות מותנית - מרחב לא אחיד	44
12.	דיאגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה	48
13.	תלות ואי תלות בין מאורעות	53
14.	שאלות מסכמות בהסתברות	57
15.	המשתנה המקרי הבדיד - פונקציית ההסתברות	62
16.	המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן	66
17.	המשתנה המקרי הבדיד - טרנספורמציה לינארית	70
18.	תוחלת ושונות של סכום משתנים מקריים	73
19.	התפלגויות בדידות מיוחדות - התפלגות בינומית	76
20.	התפלגויות בדידות מיוחדות - התפלגות גיאומטרית	80
21.	התפלגויות בדידות מיוחדות - התפלגות אחידה	83
22.	התפלגויות בדידות מיוחדות - התפלגות פואסונית	86
23.	קירוב פואסוני להתפלגות הבינומית	89

תוכן העניינים

24.	התפלגויות בדידות מיוחדות-התפלגות היפרגאומטרית	91
25.	המשתנה המקרי הבדיד -תוחלת של פונקציה של משתנה מקרי בדיד	94
26.	המשתנה המקרי הבדיד - שאלות מסכמות	97
27.	המשתנה המקרי הרציף- התפלגויות כלליות (שימוש באינטגרלים)	104
28.	התפלגויות רציפות מיוחדות-התפלגות אחידה	113
29.	התפלגויות רציפות מיוחדות- התפלגות מעריכית	116
30.	התפלגויות רציפות מיוחדות - התפלגות נורמלית	119
31.	משתנה דו-מימדי בדיד - פונקציית הסתברות משותפת	127
32.	משתנה דו מימדי בדיד - מתאם בין משתנים	133
33.	המשתנה המקרי הדו ממדי - קומבינציות ליניאריות	140
34.	המשתנה המקרי הדו ממדי הבדיד - שאלות מסכמות	143
35.	המשתנה המקרי הדו ממדי הרציף	151
36.	מערכות חשמליות	159
37.	נוסחת התוחלת השלמה	162
38.	נוסחת השונות השלמה (שונות של משתנה מותנה)	165
39.	הסקה סטטיסטית - הקדמה	167
40.	התפלגות הדגימה ומשפט הגבול המרכזי	170
41.	אי שוויונים בהסתברות	184
42.	חישוב תוחלת ושונות על ידי פירוק לאינדקטורים	192
43.	אמידה נקודתית	195
44.	רווח סמך לתוחלת (ממוצע)	210
45.	בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן)	221
46.	הלמה של ניימן פירסון	228
47.	מבוא לבדיקת השערות על פרמטרים	235
48.	בדיקת השערות על תוחלת (ממוצע)	241

מבוא להסתברות וסטטיסטיקה

פרק 1 - יסודות ההסתברות

תוכן העניינים

1. כללי 1

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 בישוב מסוים:

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

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":

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

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. $\frac{1}{8}$

מבוא להסתברות וסטטיסטיקה

פרק 2 - פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים

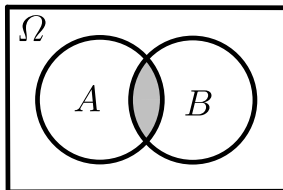
תוכן העניינים

1. כללי 5

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

רקע:

פעולת חיתוך:

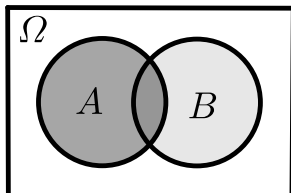


נותנת את המשותף בין המאורעות הנחתכים.
 חיתוך בין המאורע A למאורע B יסומן כך: $A \cap B$.
 מדובר בתוצאות שנמצאות ב- A וגם ב- B .

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:



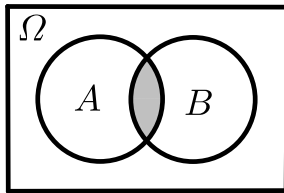
נותנת את כל האפשרויות שנמצאות לפחות באחת מהמאורעות, ומסומנת: $A \cup B$.
 הפעולה נותנת את אשר נמצא ב- A או ב- B .
 כלומר, לפחות אחד מהמאורעות קורה.

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:

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

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

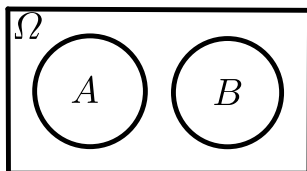
$$P(A \cap B) = 1 - P(\bar{A} \cup \bar{B})$$

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

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.

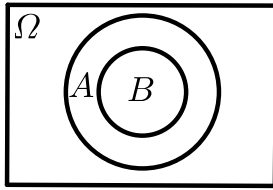
ההסתברות של חיתוך המאורעות היא אפס: $P(A \cap B) = 0$.

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: $P(A) = 0.3$, $P(B) = 0.25$, $P(A \cup B) = 0.49$

- א. חשבו את הסיכוי ל- $P(A \cap B)$.
- ב. האם A ו- B מאורעות זרים?
- ג. מה ההסתברות שרק A יקרה או שרק B יקרה?

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

מה הסיכוי למאורע A ומה ההסתברות למאורע B ?

16 קבעו אילו מהטענות הבאות נכונות:

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

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

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

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

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

- (19)** חברה מסוימת פרסמה את הנתונים הבאים לגבי האזרחים מעל גיל 21. הנתונים שהתקבלו היו: 40% מהאנשים מחזיקים כרטיס "ויזה", 52% מחזיקים כרטיס "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.

מבוא להסתברות וסטטיסטיקה

פרק 3 - קומבינטוריקה - כלל המכפלה

תוכן העניינים

1. כללי 14

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
(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			

מבוא להסתברות וסטטיסטיקה

פרק 4 - קומבינטוריקה- תמורה - סידור עצמים בשורה

תוכן העניינים

1. כללי.....18

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. $\frac{1}{35}$

מבוא להסתברות וסטטיסטיקה

פרק 5 - קומבינטוריקה- סידור עצמים במעגל

תוכן העניינים

1. סידור עצמים במעגל 21

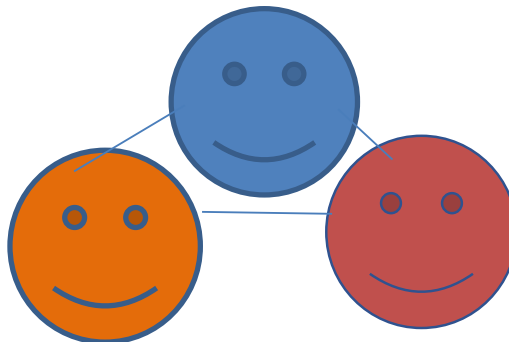
קומבינטוריקה – סידור עצמים במעגל:

רקע:

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

דוגמה (פתרון בהקלטה):

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



שאלות:

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



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



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



- (4) n בנים ו- n בנות הסתדרו במעגל באקראי.
- א. מה הסיכוי שכל הבנים יסתדרו זה לצד זה בלי להתפצל?
 - ב. מה הסיכוי שכל הבנים יסתדרו זה לצד זה בלי להתפצל וגם כל הבנות יסתדרו זו לצד זו בלי להתפצל?
 - ג. מה הסיכוי שהסידור יהיה שמימין ומשמאל לכל בן תהיה בת?

תשובות סופיות:

(1) א. 2177280 ב. 7776

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

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

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

מבוא להסתברות וסטטיסטיקה

פרק 6 - קומבינטוריקה - דגימה סידורית ללא החזרה ועם החזרה

תוכן העניינים

1. כללי 24

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

מבוא להסתברות וסטטיסטיקה

פרק 7 - קומבינטוריקה - דגימה ללא סדר וללא החזרה

תוכן העניינים

1. כללי 26

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
(3)	א. 0.1117	ב. 0.1445	ג. 0.9819
(4)	א. 0.02	ב. 0.187	ג. 0.972
(5)	א. 0	ב. 0.1923	ג. 0.009
(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. $\sum_{i=1}^n \binom{2n}{i}^2$

מבוא להסתברות וסטטיסטיקה

פרק 8 - כלל ההכלה וההפרדה

תוכן העניינים

1. כלל ההכלה וההפרדה 29

כלל ההכלה וההפרדה:

רקע:

אנו מעוניינים בנוסחה לחישוב הסתברות של איחוד מאורעות. אם קיימים n מאורעות זרים בזוגות, הסיכוי לאיחוד המאורעות הוא סכום ההסתברויות של כלל המאורעות.

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n p(A_i)$$

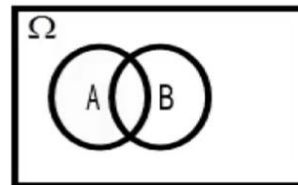


אם דנים במאורעות שאינם בהכרח זרים בזוגות, סכימת ההסתברויות של כלל המאורעות תוביל לספירה כפולה של חלק מהמאורעות. למשל: אדם מטיל קובייה. מה הסיכוי לקבל תוצאה זוגית או את התוצאה 2 לכל היותר?

בדוגמה שהוצגה לעיל מדובר בהסתברות לאיחוד שני מאורעות ומתקיים:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

אפשר להמחיש זאת באמצעות דיאגרמת ון:

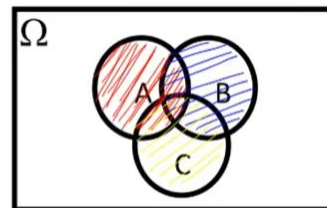


נוסחה זו היא נוסחת ההכלה וההפרדה לשני מאורעות.

במקרה של שלושה מאורעות נוסחת ההכלה וההפרדה תיראה כך:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$$

אפשר להמחיש זאת באמצעות דיאגרמת ון:



כעת נכליל את הנוסחה ל- n מאורעות:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) \dots + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right)$$

**דוגמה (פתרון בהקלטה):**

מטילים חמש קוביות. מה ההסתברות שלפחות אחת מהתוצאות הבאות לא תתקבל באף אחת מהקוביות: 1, 2, 3, 4?

שאלות:

(1) רני קיבלה ליום הולדתה חמש מתנות וסידרה אותן בשורה מימין לשמאל.



- א. מה ההסתברות שהמתנה העטופה בנייר מנוקד לא תהיה הראשונה בשורה והמתנה העטופה בסרט ירוק לא תהיה האחרונה בשורה?
 ב. מה ההסתברות שלפחות אחד מהמאורעות הבאים יתרחש?
 i. המתנה העטופה בנייר מנוקד לא תהיה השנייה בשורה.
 ii. המתנה הקטנה ביותר לא תהיה השנייה בשורה.
 iii. המתנה העטופה בנייר חום לא תהיה השלישית בשורה.

(2) אדם הטיל קובייה ארבע פעמים. נגדיר את המאורעות הבאים:



- A: התוצאה 1 התקבלה לפחות פעם אחת.
 B: התוצאה 2 התקבלה לפחות פעם אחת.
 C: התוצאה 3 התקבלה לפחות פעם אחת.

חשבו את ההסתברויות הבאות:

- א. $P(A \cap B)$
 ב. $P(A \cap B \cap C)$

(3) מפזרים באופן מקרי חמישה כדורים בעשרה תאים. הכדורים ממוספרים מ-1 עד 5. אין הגבלה על מספר הכדורים בכל תא.



חשבו את ההסתברויות הבאות:

- א. לפחות אחד משני התאים השמאליים ריק.
 ב. לפחות אחד משלושת התאים השמאליים ריק.
 ג. שני התאים השמאליים תפוסים.
 ד. ארבעת התאים השמאליים תפוסים.

(4) בוחרים מספר אקראי מהמספרים: $\Omega = \{1, 2, \dots, 1000\}$.

מה ההסתברות שהמספר שנבחר יתחלק לפחות באחד מהמספרים: 2, 3, 5?



- (5) עשרה ילדי גן נתבקשו לבחור גיבור-על מרשימה של 14 גיבורי-על. אם כל ילד בוחר באקראי מתוך הרשימה ללא תלות בילדים אחרים, מה ההסתברות שבדיוק שישה גיבורי-על ייבחרו?

- (6) ניסוי מקרי הוא בעל מרחב המדגם: $\Omega = \{1, 2, 3, \dots\}$.
 $P(n)$ הוא ההסתברות לקבל את התוצאה n ממרחב המדגם.
 נתון ש: $P(n) = A \cdot 0.5^n$. כמו כן נתון ש- A הוא קבוע חיובי.
 א. מצאו את ערכו של הפרמטר A .
 ב. חשבו את: $P(n > 6)$.
 ג. חשבו את הסיכוי שהתוצאה שתתקבל בניסוי תהיה אי-זוגית.

- (7) יוצרים מספר בן 8 ספרות מהספרות: $1, 2, \dots, 8$. במספר שיוצרים כל ספרה מופיעה בדיוק פעם אחת.
 א. מה ההסתברות שבמספר לא מופיעים הרצפים: $12, 34, 56$?
 ב. מה ההסתברות שבמספר מופיע לפחות אחד מהרצפים: $123, 234, 567$?



- (8) מסדרים בשורה שמונה נעליים שהן ארבעה זוגות.
 מה ההסתברות שלפחות שתי נעליים שהן זוג יהיו זו לצד זו בשורה?

- (9) מפזרים באופן מקרי m כדורים ל- n תאים שונים ($m \geq n$). תא יכול להכיל גם יותר מכדור אחד. מצאו ביטוי להסתברות שבכל תא יהיה לפחות כדור אחד. אין צורך לפשט את הביטוי שקיבלתם.



- (10) בכיתה יש n תלמידים, ולכל תלמיד יומן אישי. המורה אסף את היומנים של כל התלמידים. יום למחרת חילק לכל תלמיד יומן מהיומנים שאסף, אך החלוקה הייתה אקראית (תלמיד לא בהכרח קיבל את היומן האישי שלו).
 א. מה ההסתברות שאף תלמיד לא קיבל את היומן האישי שלו?
 ב. למה שואפת ההסתברות מהסעיף הקודם אם: $n \rightarrow \infty$?



- 11** על השולחן עשרה מסמכים. מכניסים כל מסמך לאחת משמונה תיקיות ריקות באופן אקראי. לכל תיקייה צבע אחר. אין הגבלה על מספר המסמכים שיכולים להימצא בכל אחת מהתיקיות.
- א. מה ההסתברות שבתיקייה האדומה יהיו בדיוק שני מסמכים?
- ב. מה ההסתברות שהתיקייה הצהובה תישאר ריקה וגם בתיקייה הירוקה יהיה לפחות מסמך אחד?
- ג. מה ההסתברות שיהיו לפחות שש תיקיות ריקות?
- ד. מה ההסתברות שיהיו בדיוק שתי תיקיות ריקות?

תשובות סופיות:

- (1) א. 0.65 ב. $\frac{59}{60}$
- (2) א. 0.233 ב. $\frac{1}{12}$
- (3) א. 0.8533 ב. 0.9565 ג. 0.1467 ד. 0.0096
- (4) 0.734
- (5) 0.1706
- (6) א. 1 ב. 0.015625 ג. $\frac{2}{3}$
- (7) א. 0.6756 ב. 0.0496
- (8) 0.6571
- (9) $1 - \frac{\sum_{i=1}^n \binom{n}{i} (n-i)^m \cdot (-1)^{i-1}}{n^m}$
- (10) א. $\sum_{i=2}^n \frac{(-1)^i}{i!}$ ב. 0.3679
- (11) א. 0.2416 ב. 0.2068 ג. 0.00003 ד. 0.4286

מבוא להסתברות וסטטיסטיקה

פרק 9 - קומבינטוריקה - שאלות מסכמות

תוכן העניינים

1. כללי 34

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

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

מבוא להסתברות וסטטיסטיקה

פרק 10 - הסתברות מותנית-במרחב מדגם אחיד

תוכן העניינים

1. כללי 41

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) $\frac{1}{4}$

מבוא להסתברות וסטטיסטיקה

פרק 11 - הסתברות מותנית - מרחב לא אחיד

תוכן העניינים

1. כללי 44

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

מבוא להסתברות וסטטיסטיקה

פרק 12 - דיאגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה

תוכן העניינים

1. כללי.....48

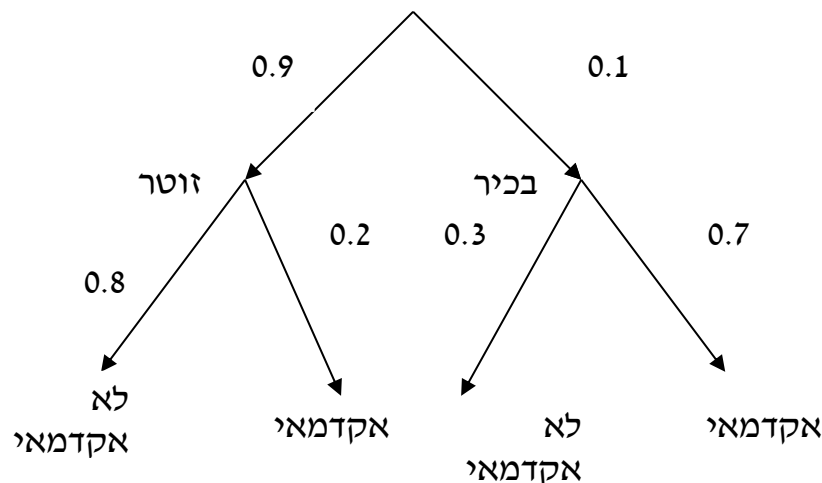
דיאגרמת עצים – נוסחת בייס ונוסחת ההסתברות השלמה:

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:

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

שאלות:

- (1) בשקית סוכריות 4 סוכריות תות ו-3 לימון. מוציאים באקראי סוכריה. אם היא בטעם תות אוכלים אותה ומוציאים סוכריה נוספת, ואם היא בטעם לימון מחזירים אותה לשקית ומוציאים סוכריה נוספת.
- א. מה ההסתברות שהסוכריה הראשונה שהוצאה בטעם תות והשנייה בטעם לימון?
- ב. מה ההסתברות שהסוכריה השנייה בטעם לימון?
- (2) באוכלוסייה מסוימת 30% הם ילדים, 50% בוגרים והיתר קשישים. לפי נתוני משרד הבריאות הסיכוי שילד יחלה בשפעת במשך החורף הוא 80%, הסיכוי שמבוגר יחלה בשפעת במשך החורף הוא 40% והסיכוי שקשיש יחלה בשפעת במשך החורף הוא 70%.
- א. איזה אחוז מהאוכלוסייה הינו קשישים שלא יחלו בשפעת במשך החורף?
- ב. מה אחוז האנשים שיחלו בשפעת במשך החורף?
- ג. נבחר אדם שחלה במשך החורף בשפעת, מה ההסתברות שהוא קשיש?
- ד. נבחר ילד, מה ההסתברות שהוא לא יחלה בשפעת במשך החורף?
- (3) בכד א' 5 כדורים כחולים ו-5 כדורים אדומים. בכד ב' 6 כדורים כחולים ו-4 כדורים אדומים. בוחרים באקראי כד, מוציאים ממנו כדור ומבלי להחזירו מוציאים כדור נוסף.
- א. מה ההסתברות ששני הכדורים שיוצאו יהיו בצבעים שונים?
- ב. אם הכדורים שהוצאו הם בצבעים שונים, מה ההסתברות שהכדור השני שהוצא יהיה בצבע אדום?
- (4) חברת סלולר מסווגת את לקוחותיה לפי 3 קבוצות גיל: נוער, בוגרים ופנסיונרים. נתון כי: 10% מהלקוחות בני נוער, 70% מהלקוחות בוגרים והיתר פנסיונרים. מתוך בני הנוער 90% מחזיקים בסמארט-פון, מתוך האוכלוסייה הבוגרת ל-70% יש סמארט-פון ומתוך אוכלוסיית הפנסיונרים 30% מחזיקים בסמארט-פון.
- א. איזה אחוז מלקוחות החברה הם בני נוער עם סמארט-פון?
- ב. נבחר לקוח אקראי ונתון שיש לו סמארט-פון. מה ההסתברות שהוא פנסיונר?
- ג. אם ללקוח אין סמארט-פון, מה ההסתברות שהוא לא בן נוער?

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

- (9) סטודנט ניגש למבחן אמריקאי. הסיכוי שהוא יודע תשובה לשאלה מסוימת הוא P , ואם הוא לא יודע את התשובה הוא מנחש. בכל מקרה הוא עונה על השאלה. נתון שלשאלה יש k תשובות אפשריות. אם הסטודנט ענה נכון על השאלה, מה הסיכוי שהוא ידע אותה?
- (10) אדם משחק נגד שני מתמודדים, רונית ודולב. האדם צריך לשחק שלושה משחקים ויש לו לבחור איזה סדר משחקים עדיף לו:
- דולב, רונית, דולב.
 - רונית, דולב, רונית.
- בכל משחק 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'.

מבוא להסתברות וסטטיסטיקה

פרק 13 - תלות ואי תלות בין מאורעות

תוכן העניינים

1. אי תלות בין מאורעות (מורחב) 53

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 למחלה 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

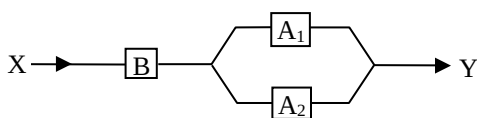
בהסתברויות שונות, אך באופן בלתי תלוי זה בזה.

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. $\frac{1}{2}$

מבוא להסתברות וסטטיסטיקה

פרק 14 - שאלות מסכמות בהסתברות

תוכן העניינים

1. כללי 57

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?
- (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?

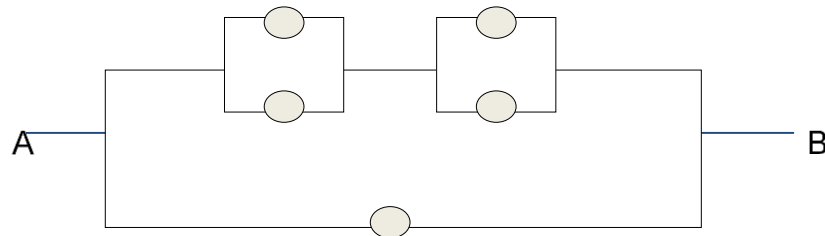
(6) בתיבה 10 מטבעות, מתוכם 7 מטבעות רגילים (ראש, זנב) ו-3 מטבעות שבשני צדדיהם טבוע ראש. אדם בוחר באקראי מטבע ומטיל אותו פעמיים. נסמן ב-A את ההטלה הראשונה ראש וב-B את ההטלה השנייה ראש.

- 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.

מבוא להסתברות וסטטיסטיקה

פרק 15 - המשתנה המקרי הבדיד - פונקציית ההסתברות

תוכן העניינים

1. כללי 62

המשתנה המקרי הבדיד – פונקציית ההסתברות:

רקע:

משתנה מקרי בדיד:

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

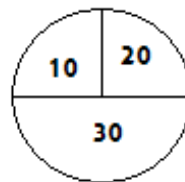
מתארים את המשתנה המקרי על ידי פונקציית ההסתברות.

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	$P(X)$

מבוא להסתברות וסטטיסטיקה

פרק 16 - המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן

תוכן העניינים

1. כללי..... 66

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.

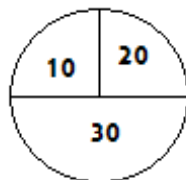
$$\text{מגדירים שונות באופן הבא: } V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = \sigma^2$$

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.

מבוא להסתברות וסטטיסטיקה

פרק 17 - המשתנה המקרי הבדיד - טרנספורמציה לינארית

תוכן העניינים

1. כללי 70

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 הני"ל?

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

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

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

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

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

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

תשובות סופיות:

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

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

- ג. תוחלת: 1650, שונות: $85,727.5^2$.
- (6) ב. $V(X) = 1.8275$.
- (7) א. $A = 10$. ב. $E(X) = 3$, $V(X) = 1$. ג. $E(X^3) = 35.4$, $V(X^3) = 616.84$. ד. $E(Y) = -2.5$, $V(Y) = 0.25$.

מבוא להסתברות וסטטיסטיקה

פרק 18 - תוחלת ושונות של סכום משתנים מקריים

תוכן העניינים

1. כללי 73

תוחלת ושונות של סכום משתנים מקריים:

רקע:

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

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

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

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

דוגמה:

אדם משחק בשני משחקי מזל בלתי תלויים. תוחלת סכום הזכייה של המשחק הראשון היא 7 עם סטיית תקן 3. תוחלת סכום הזכייה של המשחק השני היא 2- עם סטיית תקן 4. מה התוחלת ומהי השונות של סכום הזכייה הכולל של שני המשחקים יחד?

שאלות:

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

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

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

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

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

מהי התוחלת ומהי סטיית התקן של סכום הזכייה בשני המשחקים?

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

(5) נתון משתנה מקרי בעל פונקציית ההסתברות הבאה:

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

$$0 \text{ אחר } ת$$

מצאו את ערכו של A .

א. חשבו את התוחלת והשונות של X .

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

תשובות סופיות:

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

מבוא להסתברות וסטטיסטיקה

פרק 19 - התפלגויות בדידות מיוחדות - התפלגות בינומית

תוכן העניינים

1. כללי 76

התפלגויות בדידות מיוחדות – התפלגות בינומית:

רקע:

נגדיר את המושג ניסוי ברנולי :

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

פונקציית ההסתברות של X :

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

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

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

$$E(X) = np \quad \text{תוחלת:}$$

$$V(X) = npq \quad \text{שונות:}$$

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

מבוא להסתברות וסטטיסטיקה

פרק 20 - התפלגויות בדידות מיוחדות - התפלגות גיאומטרית

תוכן העניינים

1. כללי 80

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

מבוא להסתברות וסטטיסטיקה

פרק 21 - התפלגויות בדידות מיוחדות - התפלגות אחידה

תוכן העניינים

1. כללי 83

התפלגויות בדידות מיוחדות – התפלגות אחידה:

רקע:

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

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

תוחלת: $E(X) = \frac{a + b}{2}$

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

דוגמה (פתרון בהקלטה) :

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

שאלות:

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

תשובות סופיות:

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

מבוא להסתברות וסטטיסטיקה

פרק 22 - התפלגויות בדידות מיוחדות- התפלגות פואסונית

תוכן העניינים

1. כללי.....86

התפלגויות בדידות מיוחדות – התפלגות פואסונית:

רקע:

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

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

פונקציית ההסתברות של ההתפלגות הפואסונית נתונה:

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

התוחלת והשונות של ההתפלגות:

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

תכונות מיוחדות של ההתפלגות:

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

דוגמה (פתרון בהקלטה):

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

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

שאלות:

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

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

תשובות סופיות:

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

מבוא להסתברות וסטטיסטיקה

פרק 23 - קירוב פואסוני להתפלגות הבינומית

תוכן העניינים

1. כללי 89

קירוב פואסוני להתפלגות הבינומית:

רקע:

אם: $X \sim B(n, p)$ עבור n גדול ו- P קטן ניתן לקרב את ההתפלגות להיות פואסונית כאשר הפרמטר: $\lambda = np$.

כאשר פונקציית ההסתברות של ההתפלגות הפואסונית כזכור היא: $p(X = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}$.
הערה: יש הטוענים, כי n 'גדול' ו- P 'קטן' משמעו: $np \geq 10$ ו- $p \leq 0.1$.

דוגמה (פתרון בהקלטה):

בקו ייצור המוני 10% מהמוצרים כחולים. בוחרים באקראי 20 מוצרים מקו הייצור. חשבו את ההסתברות שמתוך המוצרים שיבחרו בדיוק 1 יהיה כחול.
פעם לפי ההתפלגות הבינומית ופעם לפי הקירוב הפואסוני.

שאלות:

- (1) במדינת שומקום 10% מהאוכלוסייה מובטלת. נדגמו 10 תושבים אקראיים מאותה מדינה. חשבו את הסיכוי שבמדגם יהיה לכל היותר מובטל אחד. השוו את התוצאה לקירוב הפואסוני.
- (2) מקו ייצור המוני נדגמו 1000 מוצרים. ידוע ש-5% מהמוצרים בקו הייצור פגומים. מה הסיכוי שבמדגם יתקבלו 45 מוצרים פגומים?
- (3) 1% מהתושבים באוכלוסייה גדולה חולים במחלה מסוימת. בסניף קופת חולים נרשמו 2000 תושבים אקראיים. חשבו לפי הקרוב הפואסוני שבדיוק 18 מהם יהיו חולים.
- (4) בעיר ניו יורק ישנם כתשעה מיליון תושבים, שמתוכם 900 אלף אסיאתיים. מה בקירוב ההסתברות שמתוך 100 תושבים אקראיים לפחות שני אסיאתיים?

תשובות סופיות:

- (1) ללא קירוב: 0.7361, עם קירוב: 0.7358.
- (2) 0.0458
- (3) 0.0844
- (4) 0.9995

מבוא להסתברות וסטטיסטיקה

פרק 24 - התפלגויות בדידות מיוחדות-התפלגות היפרגאומטרית

תוכן העניינים

1. כללי 91

התפלגויות בדידות מיוחדות – התפלגות היפרגאומטרית:

רקע:

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

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

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

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

דוגמה (הפתרון בהקלטה):

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

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

שאלות:

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

תשובות סופיות:

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

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

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

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

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

(4) א. 0.0256. ב. 0.8.

מבוא להסתברות וסטטיסטיקה

פרק 25 - המשתנה המקרי הבדיד - תוחלת של פונקציה של משתנה מקרי
בדיד

תוכן העניינים

1. ראשי..... 94

תוחלת של פונקציה של משתנה מקרי בדיד:

רקע:

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

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

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

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

דוגמה (פתרון בהקלטה):

נתון:

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

מצאו התפלגות ותוחלת של: $Y = X^2$.

שאלות:

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

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

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

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

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

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

חשבו את התוחלת של:

א. X^2 .

ב. 2^X .

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

א. מצאו את ערכו של A .

ב. חשבו את: $E([X - E(X)]^2)$.

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

א. חשבו את $E(W)$.

ב. חשבו את $E(W^3)$.

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

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

- (7) למיקה יש 20 חולצות ו-3 מגירות. כאשר מיקה מסדרת את 20 החולצות במגירות היא בוחרת עבור כל חולצה, באופן מקרי ובלתי תלוי בחולצות האחרות, את המגירה אליה תכניס את החולצה (כל אחת מהמגירות יכולה להכיל את כל החולצות).
 נסמן ב- X את מספר המגירות המכילות בדיוק 10 חולצות.
 מצאו את התפלגות X ואת: $E(\sqrt{X+2})$.

- (8) מטבע מוטל 10 פעמים. X = מספר הפעמים שהתקבלה התוצאה ראש.
 א. בנו את פונקציית ההסתברות של X .
 ב. הרווח במשחק הוא 4^X . מצאו את התוחלת של הרווח במשחק.

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

תשובות סופיות:

- (1) א. $E\left(\frac{1}{X}\right) = 0.4083$, $E(X) = 2.9$. ב. לא.
 (2) א. 3.2. ב. 3.45.
 (3) א. 10. ב. 1.
 (4) א. 0.95. ב. 2.21.
 (5) הוכחה.
 (6) א. $X \sim U(1,3)$. ב. 12.
 (7) $E(\sqrt{X+2}) = 1.4659$.
 (8) א. $X \sim B(10, 0.5)$. ב. 2.5^{10} .

מבוא להסתברות וסטטיסטיקה

פרק 26 - המשתנה המקרי הבדיד - שאלות מסכמות

תוכן העניינים

1. כללי 97

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?

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

חשבו: $E(X - 5)^2$.

(10) הסיכוי לעבור מבחן הנהיגה הינו P . בוחרים באקראי ארבעה נבחנים.

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

הארבעה יעברו את המבחן.

א. חשבו את ערכו של P .

ב. תלמיד ניגש לבחינה עד אשר הוא עובר אותה.

מה ההסתברות שיעבור את מבחן הנהיגה רק במבחן הרביעי?

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

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

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

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

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

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

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

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

ראש בפעם הראשונה מפסידים שקל ומפסיקים את המשחק. אחרת,

ממשיכים לזרוק וזוכים במספר שקלים לפי מספר הפעמים שהטלנו את

המטבע מההתחלה ועד שהתקבל ראש.

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

ב. בטאו את תוחלת הרווח באמצעות P .

ג. לאלו ערכי P המשחק כדאי?

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

- 19** בשדרות רוטשילד החליטו לשתול n ברושים ו-2 אורנים אחד אחרי השני בשורה. סידור העצים בשורה נעשה באקראי. נגדיר את X להיות מספר הברושים, בין הברוש הגבוה ביותר לברוש הנמוך ביותר שנשתלו.
- א. מצאו את ההתפלגות של X .
- ב. הוכיחו שהתוחלת של X היא $\frac{n-2}{3}$.

תשובות סופיות:

- (1) א. תוחלת: 2, סטיית תקן: 1. ב. תוחלת: 0, סטיית תקן: 2.
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.
- (9) 10.
- (10) א. 0.6. ב. 0.0384. ג. 0.0256.
ד. תוחלת: 0.67, שונות: 1.11. ה. 0.24.

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

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

ב. $\frac{1-2p^2}{p}$

ג. $0 < p < \sqrt{\frac{1}{2}}$

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

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

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

(16) ראו טבלה:

3	2	1	X
0.7099	0.2893	0.0008	$P(X)$

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

$$2.675 \quad (18)$$

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

מבוא להסתברות וסטטיסטיקה

פרק 27 - המשתנה המקרי הרציף- התפלגויות כלליות (שימוש באינטגרלים)

תוכן העניינים

1. כללי 104

המשתנה המקרי הרציף – התפלגויות כלליות (שימוש באינטגרלים)

רקע:

בפרק זה נעסוק בהתפלגות של משתנים מקריים רציפים (גובה אדם אקראי, זמן תגובה וכו'). משתנים רציפים הם משתנים שבתחום מסוים מקבלים רצף אינסופי של ערכים אפשריים בניגוד למשתנים בדידים. נתאר את המשתנה המקרי הרציף על ידי פונקציה הנקראת פונקציית צפיפות.

באופן כללי נסמן פונקציית צפיפות של משתנה רציף כלשהו ב- $f(x)$. השטח שמתחת לפונקציית הצפיפות נותן את ההסתברות. פונקציית צפיפות חייבת להיות לא-שלילית והשטח הכולל שמתחת לפונקציה יהיה תמיד 1.

הגדרות יסודיות:

יהא משתנה רציף X בעל פונקציית צפיפות $f(x)$.

פונקציית התפלגות מצטברת:

פונקציית ההתפלגות המצטברת מוגדרת באופן הבא: $F(t) = p(X \leq t) = \int_{-\infty}^t f(x) dx$.
כמו כן מתקיים: $p(X > t) = 1 - F(t)$ ו- $p(a < X < b) = F(b) - F(a)$.

תוחלת ושונות של משתנה רציף:

תוחלת של משתנה רציף תחושב באופן הבא: $E(X) = \int_{-\infty}^{\infty} X \cdot f(x) dx = \mu$.

שונות של משתנה רציף תחושב באופן הבא: $V(X) = \int_{-\infty}^{\infty} X^2 \cdot f(x) dx - \mu^2 = \sigma^2$.

תוחלת של פונקציה של X :

תוחלת של פונקציית משתנה רציף X , המסומנת : $g(x)$, תחושב באופן

$$\text{הבא : } E(g(x)) = \int_{-\infty}^{\infty} g(x) f(x) dx$$

אחוזונים :

האחוזון ה- p הוא ערך (נסמן אותו : x_p), שהסיכוי ליפול מתחתיו הוא p .

$$\text{כלומר : } p(X \leq x_p) = p$$

ריענון מתמטי :

נוסחאות לחישוב שטחים

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

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

משוואת קו ישר :

משוואת ישר מפורשת תסומן : $y = mx + n$, כאשר m הוא שיפוע הישר ו- n היא נקודת החיתוך של הישר עם ציר ה- y .

$$\text{שיפוע ישר העובר דרך שתי נקודות : } (x_1, y_1), (x_2, y_2) \text{ הוא : } m = \frac{y_2 - y_1}{x_2 - x_1}$$

משוואת ישר שעובר דרך נקודה ספציפית (x_1, y_1) ושיפועו הוא m , תחושב באופן

$$\text{הבא : } y - y_1 = m(x - x_1)$$

אינטגרלים מידיים:

$$\int a dx = ax + c$$

$$\int x^n dx = \frac{x^{n+1}}{n+1} + c \quad n \neq -1$$

$$\int \frac{1}{x} dx = \ln |x| + c$$

$$\int e^x dx = e^x + c$$

$$\int k^x dx = \frac{k^x}{\ln k} + c$$

$$\int \cos x dx = \sin x + c$$

$$\int \sin x dx = -\cos x + c$$

$$\int \tan x dx = -\ln |\cos x| + c$$

$$\int \cot x dx = \ln |\sin x| + c$$

$$\int \frac{1}{\cos^2 x} dx = \tan x + c$$

$$\int \frac{1}{\sin^2 x} dx = -\cot x + c$$

$$\int \frac{1}{\cos x} dx = \ln \left| \frac{1}{\cos x} + \tan x \right| + c$$

$$\int \frac{1}{x^2 + a^2} dx = \frac{1}{a} \arctan \left(\frac{x}{a} \right) + c$$

$$\int \frac{1}{\sqrt{a^2 - x^2}} dx = \arcsin \left(\frac{x}{a} \right) + c$$

$$\int \frac{f'}{f} dx = \ln |f| + c$$

$$\int e^f \cdot f' dx = e^f + c$$

$$\int \sin f \cdot f' dx = -\cos(f) + c$$

$$\int \sqrt{f} \cdot f' dx = \frac{2}{3} f^{\frac{3}{2}} + c$$

$$\int (ax+b)^n dx = \frac{1}{a} \frac{(ax+b)^{n+1}}{n+1} + c \quad n \neq -1$$

$$\int \frac{1}{ax+b} dx = \frac{1}{a} \ln |ax+b| + c$$

$$\int e^{ax+b} dx = \frac{1}{a} e^{ax+b} + c$$

$$\int k^{ax+b} dx = \frac{1}{a} \frac{k^{ax+b}}{\ln k} + c$$

$$\int \cos(ax+b) dx = \frac{1}{a} \sin(ax+b) + c$$

$$\int \sin(ax+b) dx = -\frac{1}{a} \cos(ax+b) + c$$

$$\int \tan(ax+b) dx = -\frac{1}{a} \ln |\cos(ax+b)| + c$$

$$\int \cot(ax+b) dx = \frac{1}{a} \ln |\sin(ax+b)| + c$$

$$\int \frac{1}{\cos^2(ax+b)} dx = \frac{1}{a} \tan(ax+b) + c$$

$$\int \frac{1}{\sin^2(ax+b)} dx = -\frac{1}{a} \cot(ax+b) + c$$

$$\int \frac{1}{\sin x} dx = \ln \left| \frac{1}{\sin x} - \cot x \right| + c$$

$$\int \frac{1}{x^2 - a^2} dx = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| + c$$

$$\int \frac{1}{\sqrt{x^2 \pm a^2}} dx = \ln |x + \sqrt{x^2 \pm a^2}| + c$$

$$\int f \cdot f' dx = \frac{1}{2} f^2 + c$$

$$\int \cos f \cdot f' dx = \sin(f) + c$$

$$\int \frac{f'}{\sqrt{f}} dx = 2\sqrt{f} + c$$

$$\int u \cdot v' dx = u \cdot v - \int u' \cdot v dx$$

שאלות:

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

א. מצאו את ערכו של c .

ב. בנו את פונקציית ההתפלגות המצטברת.

ג. חשבו את ההסתברויות הבאות:

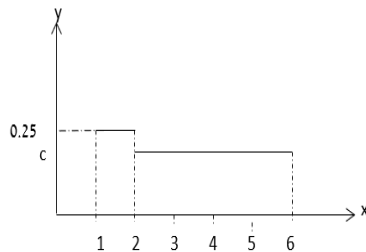
i. $P(x < 4)$

ii. $P(x > 1.5)$

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

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

v. מצאו את החציון של המשתנה.



(2) נתון משתנה מקרי רציף A שפונקציית הצפיפות שלו היא:

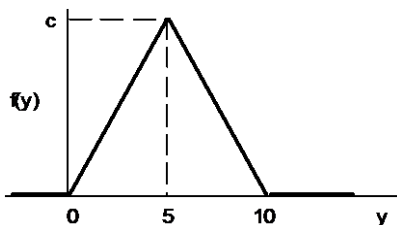
$$f(x) = \begin{cases} cx & 0 \leq x \leq b \\ 0 & \text{else} \end{cases}$$

וידוע ש- $P(0 < X < 1) = \frac{1}{4}$

א. מצאו במפורש את פונקציית הצפיפות של X .

ב. מצאו את החציון של X .

ג. מה הסיכוי ש- X קטן מ-0.5?



(3) נתונה פונקציית צפיפות של משתנה מקרי Y :

א. מצאו את c .

ב. מצאו את פונקציית ההתפלגות המצטברת של Y .

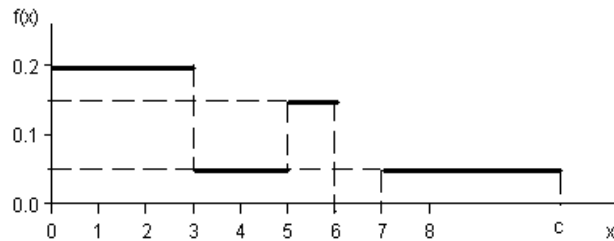
ג. חשבו את ההסתברויות הבאות:

$P(Y > 4)$, $P(7.5 \leq Y \leq 15.5)$, $P(Y \leq 3.0)$, $P(Y = 7.0)$

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

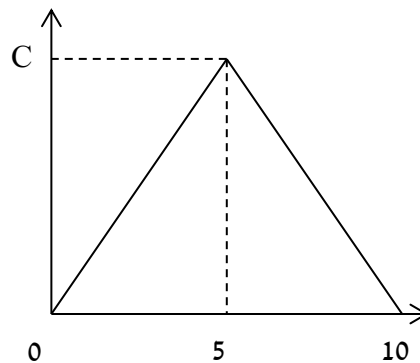
הסיקו מהו העשירון עליון: $y_{0.9}$.

(4) נתונה פונקציית צפיפות של משתנה מקרי X :



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

(5) נתונה פונקציית הצפיפות הבאה :



- א. מה ערכו של c ?
- ב. מצאו אינטרוול (תחום) סימטרי סביב הערך 5, שהסיכוי ליפול בו הינו 0.5.

(6) נתונה פונקציית צפיפות : $f(X) = \frac{2}{x}$, המוגדרת מ-1 עד K .

- א. מצאו את ערכו של K .
- ב. בנו את פונקציית ההתפלגות המצטברת.
- ג. חשבו את הסיכוי ש- X לפחות 1.5.
- ד. מצאו את העשירון התחתון של ההתפלגות.
- ה. מה התוחלת של X ?

(7) נתונה פונקציית צפיפות הבאה: $f(X) = AX^2(10 - X)$, $0 < X < 10$.

A הינו קבוע חיובי.

א. מצאו את A.

ב. חשב את: $P(x > 5 | x > 2)$.

ג. מה התוחלת ומהי השונות של X?

(8) פונקציית הצפיפות של משתנה מקרי רציף X:

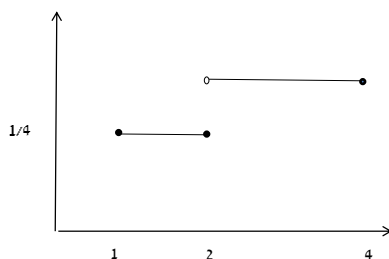
$$f(x) = 0.5 \cdot e^{2x}, \quad -\infty \leq x \leq \ln(c)$$

א. מצאו את ערכו של c.

ב. מצאו את פונקציית ההתפלגות המצטברת של ההתפלגות.

ג. חשב: $P(X > 0)$.

ד. מהו הרבעון העליון של ההתפלגות?



(9) נתונה פונקציית הצפיפות הבאה של משתנה מקרי X:

א. רשמו את נוסחת פונקציית הצפיפות.

ב. בנו את פונקציית ההתפלגות המצטברת.

ג. מצאו את החציון של ההתפלגות.

ד. חשבו את התוחלת והשונות של המשתנה.

ה. חשבו את: $E(X^3)$.

(10) במפעל מייצרים מוצר A. זמן תהליך הייצור של המוצר בשעות הוא בעל

$$f(x) = 6x(1-x), \quad 0 \leq x \leq 1$$

א. מה ההסתברות שזמן הייצור של מוצר A אקראי יהיה קטן מ-20 דקות?

ב. מה ההסתברות שזמן הייצור של מוצר A אקראי יהיה בדיוק חצי שעה?

ג. נבחרו חמישה מוצרים אקראיים מסוג A. מה תוחלת מספר המוצרים

שזמן הייצור שלהם יהיה גדול מ-20 דקות?

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

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

א. שרטטו את פונקציית ההתפלגות המצטברת.

ב. מה הסיכוי שזמן ההמתנה יהיה לפחות רבע שעה?

ג. אם חיכיתי בתור כבר 10 דקות מה ההסתברות שאאלץ לחכות בסך הכול

פחות מרבע שעה?

ד. מהו הזמן ש-90% מהלקוחות מחכים מתחתיו?

(12) פונקציית הצפיפות של משתנה מקרי נתונה על ידי הנוסחה הבאה :

$$f(x) = \begin{cases} 0 & x < 4 \\ bx - 4b & 4 \leq x \leq 5 \\ b & 5 < x \leq 6 \\ 0 & x > 6 \end{cases}$$

- א. מצאו את b .
- ב. חשבו את התוחלת של X .
- ג. y הוא משתנה אינדיקטור המקבל את הערך 1 אם X קטן מ-5. מהי השונות של y ?

(13) נתונה פונקציית הצפיפות הבאה :

$$f(x) = \begin{cases} \frac{x^2}{4} & 1 \leq x \leq 2 \\ kx & 2 < x \leq 3 \end{cases}$$

- א. מצאו את ערכו של k .
- ב. מצאו את פונקציית ההתפלגות המצטברת.
- ג. חשבו $P(x > 2.5)$.

(14) להלן משתנה מקרי בעל פונקציית צפיפות הבאה : $f(x) = \frac{1}{b-a}$, $a \leq x \leq b$

- א. מצאו את פונקציית ההתפלגות המצטברת.
- ב. חשב את התוחלת והשונות של ההתפלגות.
- ג. מצאו את התוחלת של $\frac{1}{X}$.

תשובות סופיות:

$$F(t) = \begin{cases} 0 & t < 1 \\ (t-1)0.25 & 1 \leq t \leq 2 \\ 0.25 + (t-2) \cdot \frac{3}{16} & 2 < t \leq 6 \\ 1 & t > 6 \end{cases} \quad \text{ב.} \quad \text{א. } \frac{3}{16} \quad \text{ג. } \frac{5}{8} \quad \text{ד. } \frac{1}{3}$$

$$\text{א. } b=2, c=0.5 \quad \text{ב. } 1.41 \quad \text{ג. } 0.0625 \quad \text{ד. } \frac{1}{3}$$

$$F(t) = \begin{cases} 0 & t < 0 \\ 0.02t^2 & 0 \leq t \leq 5 \\ 1 - 0.02(t-10)^2 & 5 < t \leq 10 \\ 1 & t > 10 \end{cases} \quad \text{ב.} \quad \text{א. } 0.2 \quad \text{ג. } 0, 0.18, 0.125, 0.32$$

ד. עשירון תחתון: 2.24, רבעון תחתון: 3.54, החציון: 5, עשירון עליון: 7.76.

$$F(t) = \begin{cases} 0 & t < 0 \\ 0.2t & 0 < t \leq 3 \\ 0.6 + (t-3) \cdot 0.05 & 3 < t \leq 5 \\ 0.7 + (t-5) \cdot 0.15 & 5 < t \leq 6 \\ 0.85 & 6 < t \leq 7 \\ 0.85 + (t-7) \cdot 0.05 & 7 < t \leq 10 \\ 1 & t > 10 \end{cases} \quad \text{ב.} \quad \text{א. } 0.10 \quad \text{ג. } 0.5$$

$$\text{א. } c=0.2 \quad \text{ב. } 5 \pm 1.46$$

$$F(t) = \begin{cases} 0 & t < 1 \\ 2 \cdot \ln t & 1 \leq t \leq e^{\frac{1}{2}} \\ 1 & t > e^{\frac{1}{2}} \end{cases} \quad \text{ב.} \quad \text{א. } e^{\frac{1}{2}} \quad \text{ג. } 0.189$$

$$\text{ד. } 1.051 \quad \text{ה. } 1.297$$

$$\text{א. } 0.0012 \quad \text{ב. } 0.7067 \quad \text{ג. תוחלת: 6, שונות: 4}$$

$$(8) \quad \text{א. 2.} \quad \text{ב.} \quad F(t) = \begin{cases} \frac{1}{4}e^{2t} & t \leq \ln(2) \\ 1 & t > \ln(2) \end{cases} \quad \text{ג. 0.75} \quad \text{ד. 0.549}$$

$$(9) \quad \text{א. 2.} \quad F(t) = \begin{cases} 0 & t < 1 \\ (t-1)0.25 & 1 \leq t \leq 2 \\ 0.25 + (t-2) \cdot \frac{3}{8} & 2 < t \leq 4 \\ 1 & t > 4 \end{cases} \quad \text{ב.} \quad F(t) = \begin{cases} \frac{1}{4} & 1 \leq x \leq 2 \\ \frac{3}{8} & 2 < x \leq 4 \\ 0 & \text{אחרת} \end{cases}$$

$$\text{ג. } 2\frac{2}{3} \quad \text{ד. תוחלת: 2.625, שונות: 0.6927} \quad \text{ה. 23.4375}$$

$$(10) \quad \text{א. } \frac{7}{27} \quad \text{ב. 0} \quad \text{ג. 3.704}$$

$$(11) \quad \text{א. עיין סרטוט בוידאו} \quad \text{ב. 0.0498} \quad \text{ג. 0.6321} \quad \text{ד. 11.51}$$

$$(12) \quad \text{א. } \frac{2}{3} \quad \text{ב. 5.22} \quad \text{ג. } \frac{2}{9}$$

$$(13) \quad \text{א. } \frac{1}{6} \quad \text{ב.} \quad F(t) = \begin{cases} 0 & t < 1 \\ \frac{t^3-1}{12} & 1 \leq t \leq 2 \\ \frac{7}{12} + \frac{t^2-4}{12} & 2 < t \leq 3 \\ 1 & t > 3 \end{cases} \quad \text{ג. 0.229}$$

$$(14) \quad \text{א.} \quad F(t) = \begin{cases} 0 & t < a \\ \frac{(t-b)}{b-a} & a \leq t \leq b \\ 1 & t > b \end{cases} \quad \text{ב. תוחלת: } E(X) = \frac{a+b}{2}, \text{ שונות: } V(X) = \frac{(b-a)^2}{12}$$

$$\text{ג. } \frac{\ln\left(\frac{b}{a}\right)}{b-a}$$

מבוא להסתברות וסטטיסטיקה

פרק 28 - התפלגויות רציפות מיוחדות-התפלגות אחידה

תוכן העניינים

1. כללי 113

התפלגויות רציפות מיוחדות – התפלגות אחידה:

רקע:

זו התפלגות שפונקציית הצפיפות שלה קבועה בין a לבין b .

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

פונקציית הצפיפות:

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

$$a \leq x \leq b$$

פונקציית ההתפלגות המצטברת:

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

התוחלת :

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

השונות:

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

דוגמה (פתרון בהקלטה):

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

מה הסיכוי ש- X קטן מ-25?

מה התוחלת והשונות של X ?

$$a = 20, b = 40$$

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

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

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

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

שאלות:

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

תשובות סופיות:

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

מבוא להסתברות וסטטיסטיקה

פרק 29 - התפלגויות רציפות מיוחדות- התפלגות מעריכית

תוכן העניינים

1. כללי 116

התפלגויות רציפות מיוחדות – התפלגות מעריכית:

רקע:

התפלגות זו היא התפלגות רציפה המאפיינת את הזמן עד להתרחשות מאורע מסוים. λ - הוא ממוצע מספר האירועים המתרחשים ביחידת זמן (אותו פרמטר מההתפלגות הפואסונית): $X \sim \exp(\lambda)$ כאשר $\lambda > 0$.
התפלגות זו צריכה להיות נתונה בתרגיל או שיאמר שמספר האירועים ביחידת זמן מתפלג פואסונית ואז הזמן עד להתרחשות המאורע הבא מתפלג מעריכית.

פונקציית הצפיפות של ההתפלגות:

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

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

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

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

$$P(X > a+b \mid X > a) = P(X > b) \quad \text{להתפלגות זו יש תכונת חוסר הזיכרון:}$$

דוגמה (פתרון בהקלטה):

אורך חיי סוללה מתפלג מעריכית עם תוחלת של 8 שעות.

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

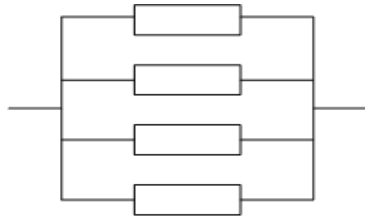
ב. מה סטיית התקן של אורך חיי הסוללה?

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

שאלות:

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

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



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

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

ב. מעוניינים להוסיף במקביל עוד רכיב למערכת. עלות הוספת רכיב היא K ₪.

כמו כן אם המערכת עבדה פחות מ-100 שעות נגרם הפסד של A ₪.

מה התנאי שבו יהיה כדאי להוסיף את הרכיב למערכת?

תשובות סופיות:

(1)	א. 0.368	ב. 0.865	ג. 0.347	
(2)	א. 24 שעות.	ב. 0.632	ג. 0.135	
(3)	א. 0.393	ב. 0.239	ג. 0.513	ד. 69.08
(4)	א. 0.264	ב. 0.368	ג. 0.233	
(5)	א. 0.8403	ב. $K < 0.0588A$		

מבוא להסתברות וסטטיסטיקה

פרק 30 - התפלגויות רציפות מיוחדות - התפלגות נורמלית

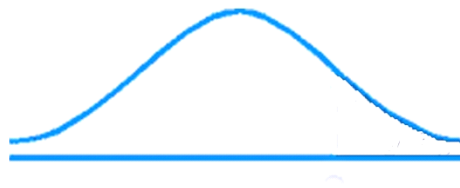
תוכן העניינים

1. כללי 119

התפלגויות רציפות מיוחדות – התפלגות נורמלית:

רקע:

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



לעקומה זו קוראים גם עקומת גאוס ועקומה אחת נבדלת מהשנייה באמצעות הממוצע וסטיית התקן שלה.

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

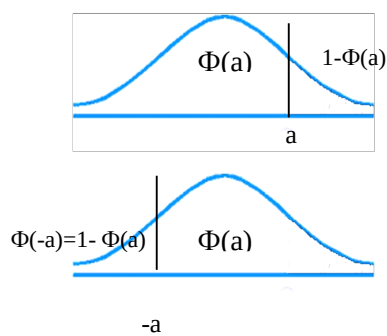
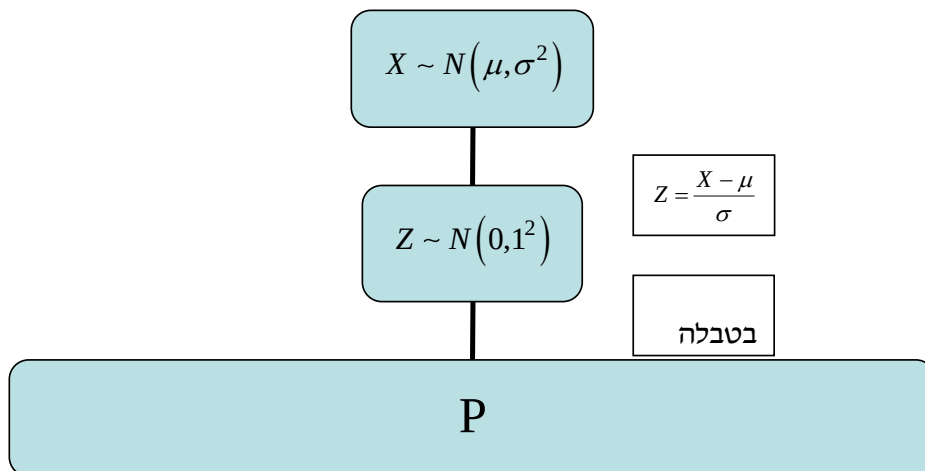
$$\text{נוסחת פונקציית הצפיפות: } f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

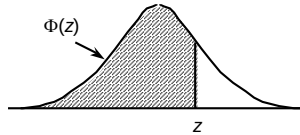
כדי לחשב הסתברויות בהתפלגות נורמלית יש לחשב את השטחים הרלוונטיים שמתחת לעקומה. כדי לחשב שטחים אלה נמיר כל התפלגות נורמלית להתפלגות נורמלית סטנדרטית על ידי תהליך הנקרא תקנון. התפלגות נורמלית סטנדרטית היא התפלגות נורמלית שהממוצע שלה הוא אפס וסטיית התקן היא אחת, והיא תסומן באות Z : $Z \sim N(0, 1^2)$.

$$\text{תהליך התקנון מבוצע על ידי הנוסחה הבאה: } Z = \frac{X - \mu}{\sigma}$$

אחרי תקנון מקבלים ערך הנקרא ציון תקן. ציון התקן משמעו בכמה סטיות תקן הערך סוטה מהממוצע.

לאחר חישוב ציון התקן של ערך מסוים נעזרים בטבלה של ההתפלגות הנורמלית הסטנדרטית לחישוב השטח הרצוי, ובאופן כללי נתאר את הסכמה הבאה:



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


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

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

דוגמה (הפתרון בהקלטה):

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

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

שאלות:

(1) הגובה של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 170 ס"מ וסטית תקן של 10 ס"מ.

- א. מה אחוז האנשים שגובהם מתחת ל-182.4 ס"מ?
- ב. מה אחוז האנשים שגובהם מעל 190 ס"מ?
- ג. מה אחוז האנשים שגובהם בדיוק 173.6 ס"מ?
- ד. מה אחוז האנשים שגובהם מתחת ל-170 ס"מ?
- ה. מה אחוז האנשים שגובהם לכל היותר 170 ס"מ?

(2) נתון שהזמן שלוקח לתרופה מסוימת להשפיע מתפלג נורמלית עם ממוצע של 30 דקות ושונות של 9 דקות רבועות.

- א. מהי פרופורציית המקרים בהן התרופה תעזור אחרי יותר משעה?
- ב. מה אחוז מהמקרים שבהן התרופה תעזור בין 35 ל-37 דקות?
- ג. מה הסיכוי שהתרופה תעזור בדיוק תוך 36 דקות?
- ד. מה שיעור המקרים שבהן ההשפעה של התרופה תסטה מ-30 דקות בפחות מ-3 דקות?

(3) המשקל של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 60 ק"ג וסטיית תקן של 8 ק"ג.

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

(4) משקל תינוקות ביום היוולדם מתפלג נורמלית עם ממוצע של 3300 גרם וסטית תקן 400 גרם.

- א. מצאו את העשירון העליון.
- ב. מצאו את האחוזון ה-95.
- ג. מצאו את העשירון התחתון.

- (5) ציוני מבחן אינטליגנציה מתפלגים נורמלית עם ממוצע 100 ושונות 225.
 א. מה העשירון העליון של הציונים במבחן האינטליגנציה?
 ב. מה העשירון התחתון של ההתפלגות?
 ג. מהו הציון ש-20% מהנבחנים מקבלים מעליו?
 ד. מהו האחוזון ה-20?
 ה. מהו הציון ש-5% מהנבחנים מקבלים מתחתיו?
- (6) נפח משקה בבקבוק מתפלג נורמלית עם סטיית תקן של 20 מ"ל, ונתון ש-33% מהבקבוקים בעלי נפח שעולה על 508.8 מ"ל.
 א. מה ממוצע נפח משקה בבקבוק?
 ב. 5% מהבקבוקים המיוצרים עם הנפח הגבוה ביותר נשלחים לבדיקה, החל מאיזה נפח שולחים בקבוק לבדיקה?
 ג. 1% מהבקבוקים עם הנפח הקטן ביותר נתרמים לצדקה, מהו הנפח המקסימלי לצדקה?
- (7) אורך חיים של מכשיר מתפלג נורמלית. ידוע שמחצית מהמכשירים חיים פחות מ-500 שעות, כמו כן ידוע ש-67% מהמכשירים חיים פחות מ-544 שעות.
 א. מהו ממוצע אורך חיי מכשיר?
 ב. מהי סטית בתקן של אורך חיי מכשיר?
 ג. מה הסיכוי שמכשיר אקראי יחיה פחות מ-460 שעות?
 ד. מהו המאיון העליון של אורך חיי מכשיר?
 ה. 1% מהמכשירים בעלי אורך החיים הקצר ביותר נשלח למעבדה לבדיקה מעמיקה. מהו אורך החיים המקסימלי לשליחת מכשיר למעבדה?
- (8) להלן שלוש התפלגויות נורמליות של שלוש קבוצות שונות ששורטטו באותה מערכת צירים. ההתפלגויות מוספרו כדי להבדיל ביניהן.
 א. לאיזו התפלגות הממוצע הגבוה ביותר?
 ב. במה מבין המדדים הבאים התפלגות 1 ו-2 זהות?
 i. בעשירון העליון.
 ii. בממוצע.
 iii. בשונות.
 ג. לאיזו התפלגות סטיית התקן הקטנה ביותר?
 i. 1.
 ii. 2.
 iii. 3.
 iv. אין לדעת.



- (9)** הזמן שלוקח לאדם להגיע לעבודתו מתפלג נורמלית עם ממוצע של 40 דקות וסטית תקן של 5 דקות.
- א. מה ההסתברות שמשך הנסיעה של האדם לעבודתו יהיה לפחות שלושת רבעי השעה?
- ב. אדם יצא לעבודתו בשעה 08:10 מביתו. הוא צריך להגיע לעבודתו בשעה 09:00. מה הסיכוי שיאחר לעבודתו?
- ג. אם ידוע שזמן נסיעתו לעבודה היה יותר משלושת רבעי השעה. מה ההסתברות שזמן הנסיעה הכולל יהיה פחות מ-50 דקות?
- ד. מה הסיכוי שבשבוע (חמישה ימי עבודה) בדיוק פעם אחת יהיה זמן הנסיעה לפחות שלושת רבעי השעה?
- (10)** ההוצאה החודשית לבית אב בעיר "טרירה" מתפלגת נורמלית עם ממוצע של 2000 דולר וסטית תקן של 300 דולר. בחרו באקראי 5 בתי אב. ההסתברות שלפחות אחד מהם מוציא בחודש מעל ל- T דולר היא 0.98976.
- א. מה ערכו של T ?
- ב. מה הסיכוי שההוצאה החודשית של בית אב בעיר תהיה לפחות סטיית תקן אחת מעל T ?
- ג. מסתבר שנפלה טעות בנתונים, ויש להוסיף 100 דולר להוצאות החודשית של כל בתי האב בעיר. לאור זאת, מה ההסתברות שההוצאה החודשית של בית אב נמוכה מ-1800 דולר?
- (11)** אורך שיר אקראי המשודר ברדיו מתפלג נורמלית עם תוחלת של 3.5 דקות וסטית תקן של שלושים שניות.
- א. מה ההסתברות שאורך של שיר אקראי המנוגן ברדיו יהיה בין 3 ל-2.5 דקות?
- ב. מהו הטווח הבין רבעוני של אורך שיר המשודר ברדיו?
- ג. ביום מסוים מנוגנים 200 שירים ברדיו. כמה שירים מתוכם תצפה שיהיו באורך הנמוך מ-3.5 דקות?
- ד. בשעה מסוימת שודרו 8 שירים. מה ההסתברות שרבע מהם בדיוק היו ארוכים מ-4 דקות והיתר לא?

תשובות סופיות:

1	א. 89.25%	ב. 2.28%	ג. 0.	ד. 50%	ה. 50%
2	א. 0%	ב. 3.76%	ג. 0%	ד. 68.26%	
3	א. 26.43%	ב. 89.44%	ג. 39.44%	ד. 0.383	
	ה. 100%				
4	א. 3812.8	ב. 3958	ג. 2787.2		
5	א. 119.2	ב. 80.8	ג. 112.6	ד. 87.4	
6	א. 500	ב. 532.9	ג. 453.48		
7	א. 500	ב. 100	ג. 0.3446	ד. 733	
	ה. 267				
8	א. 3	ב. בממוצע.	ג. 1		
9	א. 0.1587	ב. 0.0228	ג. 0.8563	ד. 0.3975	
10	א. 1925	ב. 0.2266	ג. 0.1587		
11	א. 0.1359	ב. 0.675	ג. 100	ד. 0.25	

מבוא להסתברות וסטטיסטיקה

פרק 31 - משתנה דו-מימדי בדיד - פונקציית הסתברות משותפת

תוכן העניינים

1. כללי 127

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:

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

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

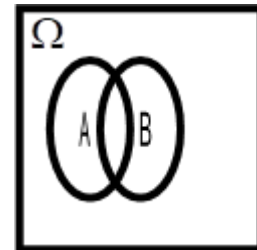
$$p(x=1, y=0) = 0.15$$

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

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

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

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

מבוא להסתברות וסטטיסטיקה

פרק 32 - משתנה דו מימדי בדיד - מתאם בין משתנים

תוכן העניינים

1. כללי 133

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:
- א. X - מספר הכדורים בתא הראשון.
 - ב. 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	$\frac{5}{36}$

(5) א. להלן טבלה: ב. X ו- Y מתואמים. ג. $\frac{2}{3}$.

$x \setminus y$	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

ד. i. $-\frac{2}{3}$ ii. -1 ה. $-\frac{2}{3}$

מבוא להסתברות וסטטיסטיקה

פרק 33 - המשתנה המקרי הדו ממדי - קומבינציות ליניאריות

תוכן העניינים

1. כללי 140

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	
$P(X)$				

- השלימו את ההסתברויות החסרות.
- האם המשתנים תלויים?
- האם המשתנים בלתי מתואמים?
- חשבו את השונות המשותפת.
- חשבו את התוחלת והשונות של סכום המשתנים.
- חשבו את התוחלת והשונות של הפרש המשתנים.

- (2) מבחן בנוי מחלק כמותי וחלק מילולי. תוחלת הציון בחלק הכמותי היא 100, עם סטיית תקן 20. תוחלת הציונים בחלק המילולי היא 90 עם סטיית תקן 15. מקדם המתאם בין הציון הכמותי לציון המילולי הוא 0.8.
- חשבו את השונות המשותפת בין הציון הכמותי לציון המילולי.
 - חשבו את התוחלת והשונות של סכום הציונים בחלק הכמותי ובחלק המילולי.
 - חשבו את התוחלת והשונות של הפרש הציונים בין החלק הכמותי לחלק המילולי.
 - עלות הבחינה 2000 שקלים. הוחלט לזכות שקל עבור כל נקודה שנצברה בחלק המילולי ושני שקלים עבור כל נקודה שנצברה בחלק הכמותי. מהי התוחלת ומהי השונות של עלות הבחינה נטו (העלות לאחר הזיכוי)?

(3) נתון: $\text{var}(X + 2Y) = 3$, $\text{var}(X - 2Y) = 2$.חשבו: $\text{cov}(X, Y)$.(4) מטילים קובייה n פעמים. נגדיר את המשתנים הבאים: X = מספר הפעמים שהתקבלה התוצאה 6. Y = מספר הפעמים שהתקבלה התוצאה 5בטאו 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) $-\frac{n}{36}$.

מבוא להסתברות וסטטיסטיקה

פרק 34 - המשתנה המקרי הדו ממדי הבדיד - שאלות מסכמות

תוכן העניינים

1. שאלות מסכמות 143

המשתנה המקרי הדו ממדי הבדיד – שאלות מסכמות:

רקע:

משתנים בלתי תלויים:

יהיו משתנים X ו- Y . הם יהיו משתנים בלתי תלויים אם עבור כל X ו- Y אפשריים מתקיים: $p(x=k, y=l) = p(x=k) \cdot p(y=l)$.

מקדם המתאם:

מגדירים את מקדם המתאם: $\rho = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y}$.

שונויות משותפת:

מגדירים את השונויות המשותפת:

$$\text{cov}(x, y) = E[(x - \mu_x)(y - \mu_y)] = E(x \cdot y) - E(x) \cdot E(y)$$

תכונות של השונויות המשותפת:

$$1. \text{cov}(X, Y) = \text{cov}(Y, X)$$

$$2. \text{cov}(X, X) = \text{var}(X)$$

$$3. \text{cov}[(aX + b), (cY + d)] = a \cdot c \cdot \text{cov}(X, Y)$$

משתנים בלתי מתואמים:

משתנים בלתי מתואמים הם משתנים שמקדם המתאם שלהם אפס וכדי שדבר כזה יקרה השונויות המשותפת צריכה להתאפס.

השפעת טרנספורמציה לינארית על מקדם המתאם:

$$\rho[(aX+b), (cY+d)] = \begin{cases} \rho(X,Y) & \text{if } a \cdot c > 0 \\ -\rho(X,Y) & \text{if } a \cdot c < 0 \end{cases}$$

תוחלת ושונות של סכום משתנים:

$$E(X+Y) = E(X) + E(Y) \quad V(X+Y) = V(X) + V(Y) + 2 \cdot \text{cov}(X, Y)$$

קומבינציות לינאריות:

נגדיר קומבינציה ליניארית כללית באופן הבא: $W = (aX + b) + (cY + d)$
אזי מתקיים:

$$E(W) = E((aX + b) + (cY + d)) = aE(X) + b + cE(Y) + d$$

$$V(W) = V((aX + b) + (cY + d)) = a^2V(X) + c^2V(Y) + 2 \cdot a \cdot c \cdot \text{cov}(X, Y)$$

שאלות:

- (1) יש ליצור סיסמה בת 3 תווים. כל תו יכול להיבחר רק מתוך כלל התווים הבאים: 1, 2, A, B, C. יהי X מספר הפעמים שהספרה 1 מופיעה בסיסמה, ויהי Y מספר הפעמים שהספרה 1 מופיעה בקצה הסיסמה (שני הקצוות).
- א. זהו את ההתפלגויות השוליות של X ו- Y כהתפלגויות מיוחדות.
 - ב. מצאו את ההתפלגות המשותפת של X ושל Y .
 - ג. מצאו את מקדם המתאם בין X ל- Y .
 - ד. מהו המתאם בין $2X$ ל- $3Y+5$?
- (2) במסיבת סוף שנה ישנו ארגז קרח ובתוכו 7 בקבוקי בירה: 4 "מכבי", 2 "גולדסטאר" ו-1 "טובורג". קרן לקחה 3 בקבוקי בירה באקראי מתוך ארגז הקרח. נסמן ב- X את מספר בקבוקי "מכבי" שנלקחו על ידי קרן, ונסמן ב- Y את מספר בקבוקי "טובורג" שנלקחו על ידי קרן.
- א. בנו את פונקציית ההסתברות המשותפת של X ושל Y .
 - ב. חשבו את התוחלת והשונות של X ושל Y .
 - ג. מצאו את השונות המשותפת של X ושל Y .
 - ד. נגדיר את W כמספר בקבוקי "גולדסטאר" שנלקחו על ידי קרן. בטאו את W באמצעות X ו- Y , וחשבו את התוחלת והשונות של W על סמך התוצאות שהתקבלו בשני הסעיפים הקודמים בלבד.
 - ה. מהו מקדם המתאם בין מספר בקבוקי "מכבי" שנלקחו על ידי קרן, למספר בקבוקים שאינם "מכבי" שנלקחו על ידי קרן?
- (3) במגירה 6 זוגות נעליים. יהודה הוציא מהמגירה 4 נעליים (לא בהכרח זוגות) באקראי. נסמן ב- W את מספר זוגות הנעליים שהוציא יהודה, ונסמן ב- R את מספר הנעליים השמאליות שהוציא יהודה.
- א. מצא את ההתפלגות המשותפת של המשתנים שהוצגו.
 - ב. האם המשתנים שהוצגו בלתי תלויים?
 - ג. מצא את התפלגות מספר הנעליים השמאליות שהוצאו אם בסך הכול הוצא זוג נעלים יחיד על ידי יהודה.
 - ד. אם ידוע שהוצאו לפחות 3 נעליים שמאליות מה הסיכוי שהוצא לכל היותר זוג אחד?

- (4) בכד 5 כדורים כחולים, 4 כדורים לבנים ו-3 כדורים ירוקים. בוחרים באקראי וללא החזרה 3 כדורים. נגדיר את המשתנים הבאים:
- X - מקבל את הערך 1 אם נבחר לפחות כדור אחד כחול, ו-0 אחרת.
 - Y - מספר הכדורים הלבנים שנבחרו.
- א. חשבו את $P(X=1)$.
- ב. בנו את פונקציית ההסתברות המשותפת של X ו- Y .
- ג. מה התוחלת של Y , אם ידוע שלא הוצאו כדורים כחולים?
- ד. מה השונות של X , אם ידוע שהוצא לכל היותר כדור לבן אחד?
- (5) ביום ההולדת הרביעי של טל הוא מחלק שלושה פרסים שונים באקראי ל-5 ילדים. בכל פעם שטל מחלק פרס הוא בוחר באקראי ילד מתוך ה-5 באופן אקראי ובלתי תלוי בבחירות הקודמות. נגדיר את המשתנים הבאים:
- X - מספר הפרסים שקיבלה יוליה.
 - Y - מספר הילדים שלא קיבלו פרס.
- א. בנו את פונקציית ההסתברות המשותפת והשוליות של X ו- Y .
- ב. האם X ו- Y הם משתנים בלתי מתואמים?
- ג. מצאו את התוחלת של $X \cdot Y^2$.
- ד. מה מקדם המתאם בין מספר הפרסים שקיבלה יוליה, למספר הילדים שקיבלו פרס?
- (6) קבעו אילו מהטענות הבאות נכונות. נמקו.
- א. אם שני משתנים הם מתואמים, אזי הם תלויים.
 - ב. אם שני משתנים הם תלויים, אזי הם מתואמים.
 - ג. אם שני משתנים הם בלתי תלויים, אזי הם בלתי מתואמים.
 - ד. אם שני משתנים הם בלתי מתואמים, אזי הם בלתי תלויים.
- (7) במקום עבודה 50 עובדים מתוכם 25 גברים ו-25 נשים. כל עובד נתבקש לבחור מתנה לחג. לכל עובד מוצגות 5 אופציות, מתוכן הוא צריך לבחור אחת. העובדים בוחרים מתנה באקראי ובאופן בלתי תלוי זה בזה.
- נסמן X_i - מספר הגברים שבחרו במתנה i .
- נסמן Y_i - מספר הנשים שבחרו במתנה i .
- א. האם X_1 ו- Y_1 הם משתנים בלתי תלויים? אין צורך לחשב רק להסביר.
- ב. האם X_1 ו- X_2 הם משתנים בלתי תלויים? אין צורך לחשב רק להסביר.
- ג. מהי ההתפלגות של $X_1 + X_2$?
- ד. האם המתאם בין X_1 ו- X_2 מלא או חלקי? חיובי או שלילי? אין צורך לחשב רק להסביר.

(8) הוכיחו את הזהות הבאה עבור שלושת המשתנים: X, Y ו- Z :

$$\text{cov}(X + Y, Z) = \text{cov}(X, Z) + \text{cov}(Y, Z).$$

(9) מספר העלים שנושרים בסתיו מהעץ בגינה מתפלג פואסונית עם תוחלת של 50

עלים בדקה. נסמן ב- Y את מספר העלים שנושרים מהעץ בין 12:00 ל-12:10,

ונסמן ב- Q את מספר העלים שנושרים בין 12:05 ל-12:30.

א. חשבו את: $\text{cov}(4Y, Q+6)$.

ב. מה המתאם בין Y ל- Q ?

(10) בסל יש 20 כדורים אדומים, 20 ירוקים ו-20 כחולים. מוציאים באקראי

מהסל 20 כדורים. מצאו את מקדם המתאם בין מספר הכדורים האדומים

שהוצאו למספר הכדורים הירוקים שהוצאו.

(11) נתון ש: $Y \sim B(1, p)$ כאשר $0 < p < 1$.

הוכיחו שאם מתקיים: $P(X = x | Y = 0) = P(X = x | Y = 1)$ לכל X ,

אז X ו- Y הם משתנים בלתי תלויים.

(12) נתון ש- $X \sim B(n, p)$ וכן: $Y \sim B(m, p)$, שאינם תלויים זה בזה.

הוכיחו שמתקיים: $X | X + Y = k \sim HG(n + m, n, k)$.

תשובות סופיות:

$$(1) \quad X \sim B\left(3, \frac{1}{5}\right), Y \sim B\left(2, \frac{1}{5}\right) \quad \text{א.}$$

ב. להלן טבלה: ג. 0.816 ד. 0.816.

X / Y	0	1	2	3	P_Y
0	$\frac{64}{125}$	$\frac{16}{125}$	0	0	$\frac{80}{125}$
1	0	$\frac{32}{125}$	$\frac{8}{125}$	0	$\frac{40}{125}$
2	0	0	$\frac{4}{125}$	$\frac{1}{125}$	$\frac{5}{125}$
P_X	$\frac{64}{125}$	$\frac{48}{125}$	$\frac{12}{125}$	$\frac{1}{125}$	1

$$(2) \quad \text{א. להלן טבלה: ב. } E(X) = \frac{12}{7}, V(X) = \frac{24}{49}, E(Y) = \frac{3}{7}, V(Y) = \frac{12}{49}.$$

X / Y	0	1	2	3	P_Y
0	0	$\frac{3}{35}$	$\frac{12}{35}$	$\frac{4}{35}$	$\frac{20}{35}$
1	$\frac{1}{35}$	$\frac{8}{35}$	$\frac{6}{35}$	0	$\frac{15}{35}$
P_X	$\frac{1}{35}$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{4}{35}$	1

$$\text{ג. } -\frac{8}{49} \quad \text{ד. } E(W) = \frac{6}{7}, V(W) = \frac{20}{49} \quad \text{ה. } -1.$$

(3) א. להלן טבלה: ב. המשתנים תלויים.

R / W	0	1	2	P_R
0	$\frac{15}{495}$	0	0	$\frac{15}{495}$
1	$\frac{60}{495}$	$\frac{60}{495}$	0	$\frac{120}{495}$
2	$\frac{90}{495}$	$\frac{120}{495}$	$\frac{15}{495}$	$\frac{225}{495}$
3	$\frac{60}{495}$	$\frac{60}{495}$	0	$\frac{120}{495}$
4	$\frac{15}{495}$	0	0	$\frac{15}{495}$
P_W	$\frac{240}{495}$	$\frac{240}{495}$	$\frac{15}{495}$	1

ג. להלן טבלה: ד. 1.

$R/w=1$	1	2	3
$P(R/w=1)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

(4) א. $\frac{185}{220}$ ב. להלן טבלה: ג. 1.714 ד. 0.071

X/Y	0	1	P_Y
0	$\frac{1}{220}$	$\frac{55}{220}$	$\frac{56}{220}$
1	$\frac{12}{220}$	$\frac{100}{220}$	$\frac{112}{220}$
2	$\frac{18}{220}$	$\frac{30}{220}$	$\frac{48}{220}$
3	$\frac{4}{220}$	0	$\frac{4}{220}$
P_X	$\frac{35}{220}$	$\frac{185}{220}$	1

(5) א. להלן טבלה: ב. X ו- Y בלתי מתואמים.

X/Y	0	1	2	3	P_Y
2	$\frac{24}{125}$	$\frac{36}{125}$	0	0	$\frac{60}{125}$
3	$\frac{36}{125}$	$\frac{12}{125}$	$\frac{12}{125}$	0	$\frac{60}{125}$
4	$\frac{4}{125}$	0	0	$\frac{1}{125}$	$\frac{5}{125}$
P_X	$\frac{64}{125}$	$\frac{48}{125}$	$\frac{12}{125}$	$\frac{1}{125}$	1

ג. 4.128 ד. 0.

(6) א. נכון. ב. לא נכון. ג. נכון. ד. לא נכון.

- (7) א. בלתי תלויים. ב. תלויים. ג. $x_1 + x_2 \sim B\left(n = 25, p = \frac{2}{5}\right)$.
- ד. חלקי ושלילי.
- (8) שאלת הוכחה.
- (9) א. 1000. ב. 0.316.
- (10) -0.5.
- (11) שאלת הוכחה.
- (12) שאלת הוכחה.

מבוא להסתברות וסטטיסטיקה

פרק 35 - המשתנה המקרי הדו ממדי הרציף

תוכן העניינים

1. משתנה דו ממדי רציף..... 151

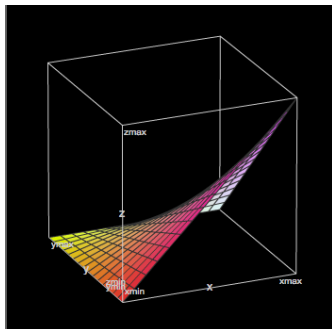
משתנה מקרי דו ממדי רציף:

רקע:

יהיו X ו- Y משתנים מקריים רציפים המוגדרים בתחום R מסוים. פונקציית הצפיפות המשותפת שלהם תסומן על ידי $f(x, y)$. פונקציית צפיפות משותפת צריכה לקיים את שני התנאים הבאים:

$$(1) \quad f(x, y) \geq 0 \quad \text{לכל } (x, y) \in R$$

$$(2) \quad \iint_R f(x, y) dx dy = 1$$



דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} : \text{נתונה הפונקציה}$$

הראו שפונקציה זו יכולה להיות פונקציית צפיפות משותפת.

פונקציית צפיפות שולית:

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy : \text{פונקציית הצפיפות השולית של } X \text{ תתקבל באופן הבא}$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx : \text{פונקציית הצפיפות השולית של } Y \text{ תתקבל באופן הבא}$$

דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} : \text{מצאו לפונקציית הצפיפות}$$

את פונקציית הצפיפות השולית של X , וחשבו את $E(X)$ דרכה.

אי-תלות בין משתנים רציפים:

X ו- Y יהיו משתנים מקרים בלתי תלויים, אם עבור כל X ו- Y בתחום ההגדרה R מתקיים ש:

$$f(x, y) = f(x) \cdot f(y)$$

דוגמה (פתרון בהקלטה):

האם X ו- Y , המתפלגים לפי פונקציית הצפיפות המשותפת:

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases}$$

הם משתנים בלתי תלויים?

חישוב הסתברויות עבור משתנה מקרי רציף דו ממדי:

הנפח הכלוא מתחת למשטח $f(x, y)$ בתחום מסוים ייתן את ההסתברות ש- X

$$\text{ו-} Y \text{ יהיו בתחום הזה: } P[(x, y) \in A] = \iint_A f(x, y) dx dy$$

דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \quad \text{משתנים מתפלגים לפי פונקציית הצפיפות:}$$

$$P(X < 0.5 \cap Y < 0.5) \quad \text{חשבו את הסיכוי:}$$

פונקציית התפלגות מצטברת משותפת:

פונקציית התפלגות מצטברת משותפת הינה פונקציה של שני משתנים רציפים המחזירה את הסיכוי שהמשתנים יהיו קטנים מערכים מסוימים:

$$F(s, t) = P(X \leq s \cap Y \leq t) = \int_{-\infty}^t \int_{-\infty}^s f(x, y) dx dy$$

דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \quad \text{משתנים מתפלגים לפי פונקציית הצפיפות:}$$

מצאו את פונקציית ההתפלגות המצטברת המשותפת

$$\text{ועל פיה חשבו את הסיכוי: } P(X < 0.5 \cap Y < 0.5)$$

פונקציית צפיפות מותנית:

אם ל- X ול- Y ישנה פונקציית צפיפות משותפת $f(x, y)$, אז מגדירים את פונקציית הצפיפות המותנית של X , בהינתן ש- $Y = y$ לכל ערכי y

$$\text{המקיימים: } f(y) > 0 \text{ על ידי: } f(x|y) = \frac{f(x, y)}{f(y)}$$

באופן דומה, פונקציית הצפיפות המותנית של Y בהינתן ש- $X = x$ לכל ערכי x

$$\text{המקיימים: } f(x) > 0 \text{ על ידי: } f(y|x) = \frac{f(x, y)}{f(x)}$$

דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \quad \text{משתנים מתפלגים לפי פונקציית הצפיפות:}$$

$$\text{מצאו את: } f(x|y)$$

תוחלת מותנית:

ל- X ול- Y ישנה פונקציית צפיפות משותפת $f(x, y)$.

$$\text{התוחלת של } X \text{ בהינתן ש- } Y = y \text{ תהיה: } E(X|Y=y) = \int_{-\infty}^{\infty} x \cdot f(x|y) dx$$

ובאופן דומה, התוחלת של Y בהינתן ש- $X = x$ תהיה:

$$E(Y|X=x) = \int_{-\infty}^{\infty} y \cdot f(y|x) dy$$

דוגמה (פתרון בהקלטה):

$$f(x, y) = \begin{cases} 4x(1-y) & 0 \leq x, y \leq 1 \\ 0 & \text{else} \end{cases} \quad \text{משתנים מתפלגים לפי פונקציית הצפיפות:}$$

$$\text{מצאו את: } E(X|Y)$$

שאלות:

- (1) נתונה פונקציית הצפיפות הבאה: $f(x, y) = x + y$, המוגדרת בתחום שבו: $0 \leq x \leq 1$ וגם: $0 \leq y \leq 1$. הוכיחו שמדובר בפונקציית צפיפות.
- (2) נתונה פונקציית הצפיפות הבאה: $f(x, y) = Ax(x - y)$, המוגדרת בתחום שבו: $0 \leq x \leq 2$ וגם: $-x \leq y \leq x$. מצאו את ערכו של הפרמטר A .
- (3) נתונה פונקציית הצפיפות הבאה: $f(x, y) = \frac{(x \cdot y)^3 + x \cdot y}{C}$, המוגדרת בתחום שבו: $0 \leq x \leq 1$ וגם: $0 \leq y \leq 1$.
 א. מצאו את ערכו של C .
 ב. מצאו את $f(y)$.
 ג. האם X ו- Y הינם משתנים בלתי תלויים?
- (4) משתנה מקרי דו ממדי מתפלג לפי פונקציית הצפיפות הבאה: $f(x, y) = \frac{1}{800}$, המוגדרת בתחום שבו: $60 \leq x \leq y$ וגם: $60 \leq y \leq 100$.
 א. הראו שפונקציה זו מקיימת את התנאים של פונקציית צפיפות.
 ב. מצאו את פונקציית הצפיפות השולית של Y .
 ג. חשבו את $E(X)$, $V(X)$.
 ד. האם X ו- Y הם משתנים בלתי תלויים?
 ה. חשבו את מקדם המתאם בין X ל- Y .
 ו. חשבו את הסיכוי: $P(Y > X + 10)$.
- (5) משתנה מקרי דו ממדי מתפלג לפי פונקציית הצפיפות הבאה:
 $f(x, y) = \lambda\mu \cdot e^{-(\lambda x + \mu y)}$, המוגדרת בתחום שבו: $x, y > 0$.
 א. מצאו את פונקציית הצפיפות של X ואת פונקציית הצפיפות של Y .
 ב. האם X ו- Y הם משתנים תלויים?
 ג. מהו מקדם המתאם בין X ל- Y ?
 ד. חשבו את הסיכוי: $P(Y > X)$.

(6) Y הינו משתנה מקרי אחיד רציף המתפלג בקטע: $[2, 4]$.

בנוסף, נתון ש- X הינו משתנה מקרי רציף המקיים: $0 \leq x \leq y$, $f(x|y) = \frac{2x}{y^2}$.
מצאו את השונות המשותפת של X ו- Y .

(7) נתונים שני משתנים מקרים רציפים X ו- Y . פונקציות הצפיפות

$$f(x, y) = \begin{cases} x & 0 < y < 1 \\ 0 & 1 - y \leq x \leq 1 + y \\ 0 & \text{else} \end{cases}$$

המשותפות שלהם היא:

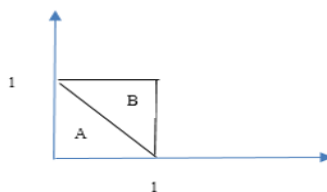
- מצאו את $f(x)$.
- מצאו את $f(y|x)$.
- מצאו את $E(Y|X)$.

(8) יהיו X ו- Y משתנים רציפים המתפלגים אחיד בתוך משולש שקדקודיו: $(-1, 2)$, $(-1, 0)$, $(0, 1)$.

- רשמו את פונקציית הצפיפות המשותפת.
- מצאו את פונקציית הצפיפות השולית של X ו- Y .
- חשבו את התוחלת של X ו- Y .
- האם X ו- Y משתנים בלתי מתואמים?
- האם X ו- Y משתנים בלתי תלויים?

(9) פונקציית צפיפות משותפת מקבלת ערך אחיד באופן הבא:

הצפיפות על פני משולש A הינה 1.5 והצפיפות על פני משולש B היא 0.5.
האם פונקציית הצפיפות המשותפת היא לגיטימית?



- מצאו את $f(x)$.
- מצאו את $f(x|y)$.

(10) נתונה פונקציית הצפיפות המשותפת: $f(x, y) = cx$. פונקציה זו מוגדרת

- בתחום שבו: $0 \leq x \leq 1$ וכן: $0 \leq y \leq x^2$.
- מצאו את הקבוע C .
- חשבו את ההסתברות ש- $6Y < 1 - X$.

11 נתונים X ו- Y שני משתנים מקריים רציפים כך ש: $Y \sim U(0,1)$

ו- $X|Y=y \sim U(0, \sqrt{y})$. חשבו את: $E(Y|X=0.5)$.

12 נתונה פונקציית הצפיפות: $f(x,y) = 2e^{-x} \cdot e^{-2y}$ בתחום שבו: $x, y \geq 0$.

חשבו את הסיכוי: $P(X < Y)$.

13 נתונה פונקציית הצפיפות המשותפת: $f(x,y) = \frac{e^{-y} \cdot e^{-\frac{x}{y}}}{y}$, המוגדרת לרביע הראשון. חשבו את: $P(X > 1|Y=2)$.

14 יוסי וערן עובדים באותו המשרד. הם מגיעים לעבודה בכל יום בין 8:00 ל-9:00. נניח שזמן ההגעה של כל אחד מתפלג אחיד ובאופן בלתי תלוי זה בזה. מה הסיכוי שיוסי יצטרך לחכות לערן יותר מ-10 דקות?

15 נתונים שני משתנים מקרים רציפים: $X \sim N(Y,1)$ ו- $Y \sim U(0,2)$.

א. מצאו את פונקציית הצפיפות המשותפת של X ו- Y .

ב. מצאו את $E(X^2|Y)$.

ג. מצאו את $E(X)$.

16 פונקציית הצפיפות המשותפת של X ו- Y היא: $f(x,y) = 1$.

פונקציה זו מוגדרת בתחומי: $0 \leq x, y \leq 1$.

הוכיחו ש: $E(|X-Y|^n) = \frac{2}{(n+1)(n+2)}$.

17 $X \sim \exp(1)$ וכן: $Y \sim \exp(1)$, הינם משתנים מקרים בלתי תלויים.

נגדיר את: $Z = \frac{X}{X+Y}$.

הוכיחו: $Z \sim U(0,1)$.

תשובות סופיות:

(1) שאלת הוכחה.

(2) $A = \frac{1}{8}$

(3) א. $\frac{5}{16}$ ב. $f(y) = 0.8y^3 + 1.6y$ ג. תלויים.

(4) א. שאלת הוכחה. ב. $f(y) = \frac{y-60}{800}$ ג. $E(X) = 73\frac{1}{3}$, $V(X) = 88\frac{8}{9}$

ד. לא. ה. 0.5 ו. 0.5625

(5) א. $f(y) = \mu e^{-\mu y}$, $f(x) = \lambda e^{-\lambda x}$ ב. לא.

ג. 0. ד. $\frac{\lambda}{\lambda + \mu}$

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

(7) א. $f(x) = \begin{cases} x^2 & 0 \leq x < 1 \\ 2x - x^2 & 1 \leq x \leq 2 \\ 0 & \text{else} \end{cases}$

$$E(y|x) = \begin{cases} 1 - \frac{x}{2} & 0 \leq x < 1 \\ \frac{x}{2} & 1 \leq x \leq 2 \\ 0 & \text{else} \end{cases} \quad \text{ג.}$$

$$f(y|x) = \begin{cases} \frac{1}{x} & 0 \leq x < 1 \quad 1-x < y < 1 \\ \frac{1}{2-x} & 1 \leq x \leq 2 \quad x-1 < y < 1 \\ 0 & \text{else} \end{cases} \quad \text{ב.}$$

(8) א. $f(x, y) = \begin{cases} 1 & 1+x < y < 1-x-1 < x < 0 \\ 0 & \text{else} \end{cases}$

$$f(y) = \begin{cases} y & 0 \leq y < 1 \\ 2-y & 1 \leq y \leq 2 \\ 0 & \text{else} \end{cases}, \quad f(x) = \begin{cases} -2x & -1 < x < 0 \\ 0 & \text{else} \end{cases} \quad \text{ב.}$$

ג. $E(X) = -\frac{2}{3}$, $E(Y) = 1$ ד. כן.

ה. לא.

$$(9) \quad \text{א. כן.} \quad \text{ב.} \quad f(x) = \begin{cases} 1.5 - x & 1 < x < 0 \\ 0 & \text{else} \end{cases}$$

$$f(x|y) = \begin{cases} \frac{1.5}{1.5-y} & 0 \leq x < 1-y \\ \frac{0.5}{1.5-y} & 1-y \leq x \leq 1 \\ 0 & \text{else} \end{cases} \quad \text{ג.}$$

$$(10) \quad \text{א. 4.} \quad \text{ב. 0.0947.}$$

$$(11) \quad \frac{7}{12}$$

$$(12) \quad \frac{1}{3}$$

$$(13) \quad e^{-\frac{1}{2}}$$

$$(14) \quad \frac{25}{72}$$

$$(15) \quad \text{א.} \quad f(x, y) = \begin{cases} \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2}} & 0 < y < 2 \\ 0 & \text{else} \end{cases}$$

$$\text{ב.} \quad y^2 + 1$$

$$\text{ג. 1.}$$

$$(16) \quad \text{שאלת הוכחה.}$$

$$(17) \quad \text{שאלת הוכחה.}$$

מבוא להסתברות וסטטיסטיקה

פרק 36 - מערכות חשמליות

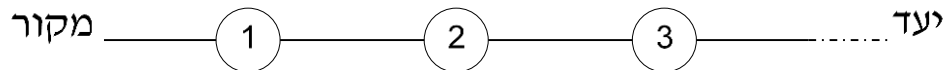
תוכן העניינים

1. כללי 159

מערכות חשמליות:

רקע:

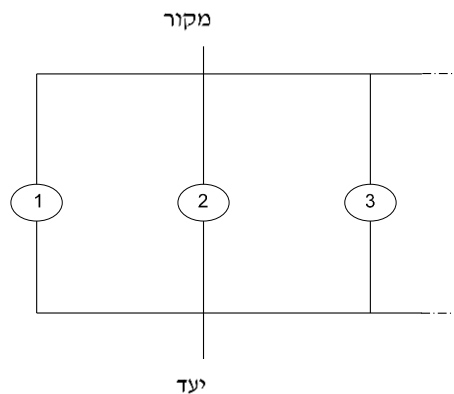
מערכת חשמלית בטור הינה מערכת חשמלית בה הרכיבים מסודרים באופן הבא:



נסמן ב- A_i את המאורע: רכיב i פועל.

כדי שהמערכת כולה תפעל צריך להתקיים ש: $\bigcap_{i=1}^n A_i$.

מערכת חשמלית במקביל הינה מערכת חשמלית בה הרכיבים מסודרים באופן הבא:



כדי שהמערכת החשמלית כולה תפעל צריך להתקיים ש: $\bigcup_{i=1}^n A_i$.

דוגמה (הפתרון בהקלטה):

במערכת חשמלית 4 רכיבים בלתי תלויים שלכל אחד מהם סיכוי P לפעול. בטאו באמצעות P את הסיכוי שהמערכת תפעל.

א. כל הרכיבים מחוברים בטור זה לזה.

ב. כל הרכיבים מחוברים במקביל זה לזה.

שאלות:

(1) נתונים שלושה רכיבים חשמליים המחוברים בטור.

אורך החיים של כל מכשיר מתפלג באופן הבא:

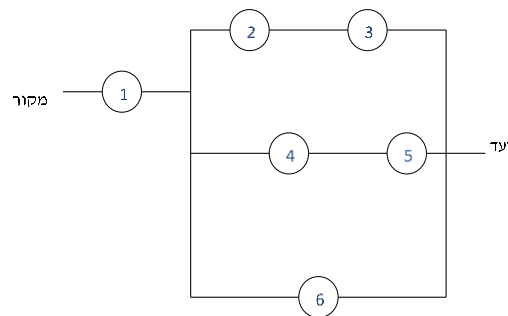
$$X_1 \sim U(2, 4)$$

$$X_2 \sim N(3, 1)$$

$$X_3 \sim \exp(1)$$

כל רכיב פועל באופן בלתי תלוי זה לזה. כל הרכיבים הופעלו כעת. מה הסיכוי שבעוד 3 שעות המערכת תפעל?

(2) המערכת החשמלית הבאה מכילה 6 רכיבים כמוראה בשרטוט:



כל רכיב פועל באופן בלתי תלוי זה לזה. רכיבים מספר 1, 2, 6 פועלים בסיכוי 0.9. רכיב מספר 3 פועל בסיכוי 0.8. רכיבים מספר 4, 5 פועלים בסיכוי P .

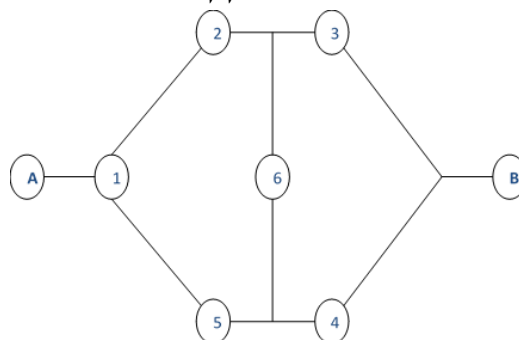
מצאו את P , אם הסיכוי שהמערכת תפעל הוא 0.887148.

(3) בין שני המחשבים A ו-B נמצאים 6 שרתים כמוראה בשרטוט. כל אחד מהשרתים תקין בסיכוי 0.9. על מנת שהודעה תצליח לעבור ממחשב A ל-B צריך להיות לפחות מסלול אחד שבו כל השרתים תקינים.

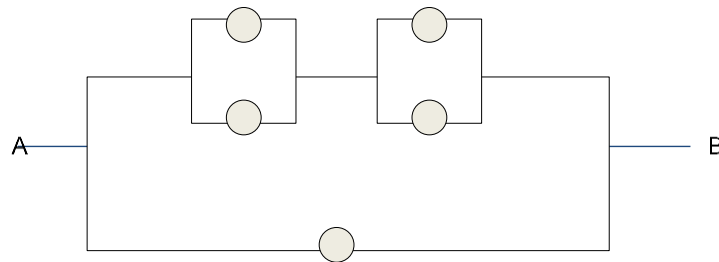
א. מה ההסתברות לכך שההודעה תעבור בהצלחה ממחשב A ל-B?

ב. ההודעה לא הצליחה לעבור ממחשב A למחשב B.

מה הסיכוי ששרת מספר 1 לא תקין?

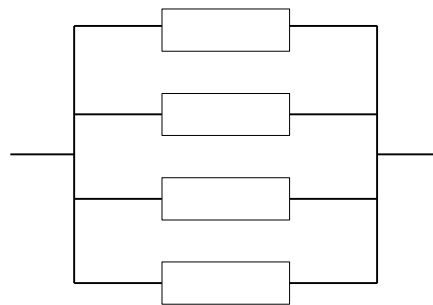


(4) נתונה המערכת החשמלית הבאה :



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

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



על מנת שהמערכת תפעל בצורה תקינה נדרש שלפחות אחד מהמרכיבים יהיה תקין. אורך החיים של כל רכיב מתפלג מעריכית עם ממוצע של 100 שעות.
 א. מה ההסתברות שהמערכת תפעל בצורה תקינה במשך 100 שעות לפחות?
 ב. נרצה להוסיף במקביל עוד רכיב למערכת. עלות הוספת רכיב היא K ₪. כמו כן אם המערכת עבדה פחות מ-100 שעות נגרם הפסד של A ₪. מה התנאי שבו יהיה כדאי להוסיף את הרכיב למערכת?

תשובות סופיות:

(1)	0.1245
(2)	0.7
(3)	א. 0.880632 ב. 0.837745
(4)	שאלת הוכחה.
(5)	א. 0.8403 ב. $0.0588A > K$

מבוא להסתברות וסטטיסטיקה

פרק 37 - נוסחת התוחלת השלמה

תוכן העניינים

1. כללי 162

נוסחת התוחלת השלמה:

רקע:

כאשר התפלגות של משתנה X תלויה במשתנה אחר Y , מתקיים: $E(X) = E[E(X/Y)]$.

עבור משתנה Y בדיד כלשהו: $E(X) = \sum_y E(X | Y = y) \cdot P(Y = y)$.

עבור משתנה Y רציף כלשהו: $E(X) = \int_{-\infty}^{\infty} E(X | Y = y) \cdot f(y) dy$.

דוגמה (הפתרון בהקלטה):

מרצה מלמד שתי כיתות. הוא מעוניין לבדוק 5 מחברות בחינה.

בכיתה מספר 1: 10 בנים ו-20 בנות.

בכיתה מספר 2: 15 בנים ו-15 בנות.

המרצה בוחר כיתה באופן מקרי וממנה בוחר 5 מחברות בחינה באקראי.

יהי X מספר מחברות הבחינה של בנים שנבחרו.

יש לחשב את $E(X)$.

שאלות:

(1) בכד 2 כדורים ירוקים, 4 כדורים כחולים ו-4 אדומים. שולפים כדור באקראי. אם הכדור ירוק מטיילים קובייה עד אשר מתקבלת התוצאה 1. אם הכדור כחול מטיילים קובייה עד אשר מתקבלת התוצאה 5. אם הכדור אדום מטיילים קובייה עד אשר מתקבלת תוצאה זוגית. חשבו את תוחלת מספר הפעמים שהקובייה הוטלה.

(2) מטיילים n מטבעות ומוציאים מהמשחק את כל המטבעות שהראו ראש. כעת מטיילים את כל המטבעות שנותרו. בטאו באמצעות n את תוחלת מספר הראשים שהתקבלו בסבב השני של ההטלות.

(3) בהגרלה מבצעים את התהליך הבא: בסיכוי 0.25 מגרילים מספר ממכונה A בסיכוי 0.75 מגרילים מספר ממכונה B. במכונה A המספר המוגרל מתפלג פואסונית עם תוחלת 6. במכונה B המספר המוגרל מתפלג פואסונית עם תוחלת 2. אם הוגרל המספר 0 זוכים ב-15₪. אחרת זוכים ב-50₪. חשבו את תוחלת סכום הזכיה.

(4) נתון ש- $Y/X \sim U(0, X)$, כאשר פונקציית הצפיפות של X

$$f(x) = \begin{cases} 0.1 & 2 \leq x \leq 6 \\ 0.2 & 6 < x \leq 9 \\ 0 & \text{אחרת} \end{cases} \quad \text{הינה:}$$

חשבו את $E(Y)$.

(5) בתכנית הריאליטי "המרוץ למיליון" מגיעים לנקודה שבה שלוש אפשרויות בפני המתמודדים: נקודה A - שבה חוזרים אחרי 1 שעות לנקודת המוצא. נקודה B - שבה חוזרים אחרי 2 שעות לנקודת המוצא ונקודה C - המובילה תוך 2 שעות לנקודת הסיום. המתמודדים בוחרים בכל פעם את הנקודה באופן מקרי. נסמן ב- X את זמן ההגעה לנקודת הסיום. חשבו את $E(Y)$.

תשובות סופיות:

(1) 4.4.

(2) $\frac{n}{4}$.

(3) 46.4.

(4) 3.05.

(5) .5

מבוא להסתברות וסטטיסטיקה

פרק 38 - נוסחת השונות השלמה (שונות של משתנה מותנה)

תוכן העניינים

1. נוסחת השונות השלמה (שונות של משתנה מותנה) 165

נוסחת השונות השלמה (המותנית):

רקע:

כאשר התפלגות של משתנה X תלויה במשתנה אחר Y , מתקיים: $E(X) = E[E(X|Y)]$.
 כמו כן, מתקיים לגבי השונות: $\text{Var}(X) = \text{Var}(E[X|Y]) + E[\text{Var}(X|Y)]$.

דוגמה (הפתרון בהקלטה):

מרצה מלמד שתי כיתות. הוא מעוניין לבדוק 5 מחברות בחינה.
 בכיתה מספר 1: 10 בנים ו-20 בנות.
 בכיתה מספר 2: 15 בנים ו-15 בנות.
 המרצה בוחר כיתה באופן מקרי וממנה בוחר 5 מחברות בחינה באקראי.
 יהי X מספר מחברות הבחינה של בנים שנבחרו.
 יש לחשב את $V(X)$.

שאלות:

- (1) בכד 2 כדורים ירוקים, 4 כדורים כחולים ו-4 אדומים. שולפים כדור באקראי. אם הכדור ירוק מטיילים קובייה עד אשר מתקבלת התוצאה 1.
אם הכדור כחול מטיילים קובייה עד אשר מתקבלת התוצאה 5.
אם הכדור אדום מטיילים קובייה עד אשר מתקבלת תוצאה זוגית.
חשבו את התוחלת והשונות של מספר הפעמים שהקובייה הוטלה.

- (2) נטיל קובייה: $Y + 4$ פעמים. נתון ש- $Y \sim P(4)$.
נגדיר את X כמספר הפעמים שהתקבלה התוצאה 2.
א. מצאו את התוחלת של X .
ב. מצאו את השונות של X .

- (3) נתון ש- $Y | X \sim U(0, X)$, כאשר פונקציית הצפיפות של X

$$f(x) = \begin{cases} 0.1 & 2 \leq x \leq 6 \\ 0.2 & 6 < x \leq 9 \\ 0 & \text{אחרת} \end{cases} \quad \text{הינה:}$$

חשבו את $V(Y)$.

תשובות סופיות:

- (1) תוחלת: 4.4, שונות: 22.64.
(2) תוחלת: $\frac{4}{3}$, שונות: $\frac{11}{9}$.
(3) 4.4.

מבוא להסתברות וסטטיסטיקה

פרק 39 - הסקה סטטיסטית - הקדמה

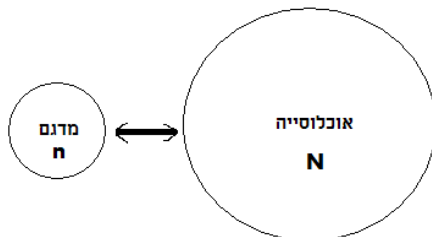
תוכן העניינים

1. כללי 167

הסקה סטטיסטית – הקדמה:

רקע:

אוכלוסייה:



קבוצה שאליה מפנים שאלה מחקרית. למשל, חברת תרופות שמעוניינת לפתח תרופה למחלת הסוכרת מתעניינת באוכלוסיית חולי הסוכרת בעולם.

מדגם:

חלק מתוך האוכלוסייה. למשל, אם נדגום באקראי 10 אנשים מתוך חולי הסוכרת אז זהו מדגם מתוך אוכלוסיית חולי הסוכרת.

במקרים רבים אין אפשרות לחקור את כל האוכלוסייה כיוון שאין גישה לכולה, היא גדולה מידי, אנו מוגבלים בזמן ובאמצעים טכניים ולכן מבצעים מדגם במטרה לבצע הסקה סטטיסטית מהמדגם לאוכלוסייה. הדגימה בקורס תהיה דגימה מקרית - הכוונה לדגימה שבה לכל תצפית באוכלוסייה יש את אותו סיכוי להיכלל במדגם.

סטטיסטי:

גודל המחושב על המדגם.

פרמטר:

גודל המתאר את האוכלוסייה.

הסימונים לפרמטר וסטטיסטי הם שונים:

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

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

דוגמה (פתרון בהקלטה):

25% מאזרחי המדינה תומכים בהצעת החוק של חבר כנסת מסוים. הוחלט לדגום 200 אזרחים ומתוכם לבדוק מהו אחוז התומכים בהצעת החוק.

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

שאלות:

- (1) מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מי האוכלוסייה?
 - מה המשתנה?
 - מהם הפרמטרים?
 - מהו גודל המדגם?
- (2) להלן התפלגות מספר מקלטי הטלויזיה למשפחה ב"העוגן". נגדיר את X להיות מספר המקלטים של משפחה אקראית. מתכננים לדגום מאוכלוסייה זו 4 משפחות ולהתבונן בממוצע מספר מקלטי הטלויזיה במדגם.
- מיהי האוכלוסייה ומהו המשתנה הנחקר?
 - מהו הסטטיסטי שיילקח מהמדגם ומה סימונו?

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

- (3) נתון כי 20% מהשכירים במדינה הם אקדמאיים. נבחרו באקראי 10 שכירים באותה אוכלוסייה ומתכננים לפרסם את מספר האקדמאיים שנדגמו.
- מיהי האוכלוסייה?
 - מה המשתנה באוכלוסייה?
 - מהם הפרמטרים?
 - מהו הסטטיסטי?

תשובות סופיות:

- (1) א. כלל הסטודנטים במכללה שסיימו סטטיסטיקה א. ב. ציון. ג. ממוצע: 78, סטיית תקן: 15. ד. 2.
- (2) א. האוכלוסייה: 1000 משפחות ב"העוגן, המשתנה הנחקר: מס' מקלטים. ב. $\bar{X}$ = ממוצע מדגם.
- (3) א. השכירים במדינה. ב. השכלה: אקדמאי, לא אקדמאי. ג. שיעור ההצלחות באוכלוסייה: 0.2. ג. מס' האקדמאים במדגם.

מבוא להסתברות וסטטיסטיקה

פרק 40 - התפלגות הדגימה ומשפט הגבול המרכזי

תוכן העניינים

1. התפלגות ממוצע המדגם ומשפט הגבול המרכזי..... 170
2. התפלגות סכום תצפיות בלתי תלויות ומשפט הגבול המרכזי..... 178
3. חוק המספרים הגדולים..... 181

התפלגות ממוצע המדגם ומשפט הגבול המרכזי:

רקע:

בפרק זה נדון בהתפלגות של ממוצע המדגם: $\bar{x} = \frac{\sum x_i}{n}$.

מכיוון שממדגם למדגם אנו יכולים לקבל ממוצע מדגם שונה, אזי ממוצע המדגם הוא משתנה מקרי ויש לו התפלגות.

גדלים המתארים התפלגות כלשהי או אוכלוסייה כלשהי נקראים פרמטרים.

להלן רשימה של פרמטרים החשובים לפרק זה:

ממוצע האוכלוסייה נסמן ב- μ (נקרא גם תוחלת).

שונות אוכלוסייה נסמן ב- σ^2 .

סטיית תקן של אוכלוסייה: σ .

תכונות התפלגות:

ממוצע כל ממוצעי המדגם האפשריים שווה לממוצע האוכלוסייה: $E(\bar{x}) = \mu_{\bar{x}} = \mu$.

שונות כל ממוצעי המדגם האפשריים שווה לשונות האוכלוסייה מחולק ב- n .

תכונה זו נכונה רק במדגם מקרי: $V(\bar{x}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$.

יש יחס הפוך בין גודל המדגם לבין שונות ממוצעי המדגם.

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

טעות תקן: $\sigma(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$.

דוגמה (פתרון בהקלטה):

השכר הממוצע במשק הינו 9000 ₪ עם סטיית תקן של 4000. דגמו באקראי 25 עובדים.

א. מיהי אוכלוסיית המחקר? מהו המשתנה הנחקר?

ב. מהם הפרמטרים של האוכלוסייה?

ג. מה התוחלת ומהי סטיית התקן של ממוצע המדגם?

דגימה מהתפלגות נורמאלית:

אם נדגום מתוך אוכלוסייה שהמשתנה בה מתפלג נורמאלית עם ממוצע μ ושונות σ^2 .

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad Z_{\bar{x}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

ממוצע המדגם גם יתפלג נורמאלית:

דוגמה (פתרון בהקלטה):

משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.
מה ההסתברות שבמדגם של 4 תינוקות אקראיים בעת הולדתם המשקל הממוצע של התינוקות יהיה מתחת ל-3.5 ק"ג?

משפט הגבול המרכזי:

אם אוכלוסייה מתפלגת כלשהו עם ממוצע μ ושונות σ^2 אזי עבור מדגם מספיק

$$\bar{x} \rightsquigarrow N\left(\mu, \frac{\sigma^2}{n}\right) \quad (n \geq 30)$$

גדול (ממוצע המדגם מתפלג בקירוב נורמאלי):

דוגמה (פתרון בהקלטה):

משקל חפיסת שוקולד בקו ייצור מתפלג עם ממוצע 100 גרם וסטיית תקן של 4 גרם.
דגמו מקו הייצור 36 חפיסות שוקולד אקראיות.
מה ההסתברות שהמשקל הממוצע של חפיסות השוקולד שנדגמו יהיה מתחת ל-102 גרם?

שאלות:

- (1) מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מיהי האוכלוסייה?
 - מה המשתנה?
 - מהם הפרמטרים?
 - מהו גודל המדגם?
 - מהו תוחלת ממוצע המדגם?
 - מהי טעות התקן?

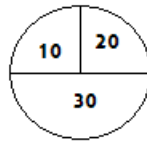
- (2) להלן התפלגות מספר מקלטי הטלויזיה למשפחה בישוב מסוים :

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

- בנו את פונקציית ההסתברות של X .
 - חשבו את התוחלת, השונות וסטיית התקן של X .
 - אם נדגום 4 משפחות מהישוב עם החזרה מה תהיה התוחלת, מהי השונות ומהי סטיית התקן של ממוצע המדגם?
- (3) אם נטיל קובייה פעמיים ונתבונן בממוצע התוצאות שיתקבלו, מה תהיה התוחלת ומה תהיה סטיית התקן של ממוצע זה?
- (4) משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.
- מה ההסתברות שתינוק אקראי בעת הלידה ישקול פחות מ-3800 גרם?
- נתון כי ביום מסוים נולדו 4 תינוקות.
- מה ההסתברות שהמשקל הממוצע שלהם יעלה על 4 ק"ג?
 - מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-2.5 ק"ג?
 - מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת בלא יותר מ-50 גרם?
 - הסבירו ללא חישוב כיצד התשובה לסעיף הקודם הייתה משתנה אם היה מדובר על יותר מ-4 תינוקות?

- (5) הגובה של המתגייסים לצה"ל מתפלג נורמאלית עם תוחלת של 175 ס"מ וסטיית תקן של 10 ס"מ. ביום מסוים התגייסו 16 חיילים.
- א. מה ההסתברות שהגובה הממוצע שלהם יהיה לפחות 190 ס"מ?
 - ב. מה ההסתברות שהגובה הממוצע שלהם יהיה בדיוק 180 ס"מ?
 - ג. מה ההסתברות שהגובה הממוצע שלהם יסטה מתוחלת הגבהים בפחות מ-5 ס"מ?
 - ד. מהו הגובה שבהסתברות של 90% הגובה הממוצע של המדגם יהיה נמוך ממנו?
- (6) הזמן הממוצע שלוקח לאדם להגיע לעבודתו 30 דקות עם שונות של 16 דקות רבועות. האדם נוסע לעבודה במשך שבוע 5 פעמים.
- לצורך הפתרון הניחו שזמן הנסיעה לעבודה מתפלג נורמאלית.
- א. מה ההסתברות שבמשך שבוע משך הנסיעה הממוצע יהיה מעל 33 דקות?
 - ב. מהו הזמן שבהסתברות של 90% ממוצע משך הנסיעה השבועי יהיה גבוה ממנו?
 - ג. מה ההסתברות שממוצע משך הנסיעה השבועי יהיה מרוחק מ-30 דקות בלפחות 2 דקות?
 - ד. כיצד התשובה לסעיף הקודם הייתה משתנה אם האדם היה נוסע לעבודה 6 פעמים בשבוע?
- (7) נפח היין בבקבוק מתפלג נורמאלית עם תוחלת של 750 סמ"ק וסטיית תקן של 10 סמ"ק.
- א. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה בדיוק 755 סמ"ק?
 - ב. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה יותר מ-755 סמ"ק?
 - ג. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה לפחות 755 סמ"ק?
 - ד. בקבוקי היין שבארגז נמזגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?
- (8) משתנה מתפלג נורמאלית עם תוחלת 80 וסטיית תקן 4.
- א. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?
 - ב. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה שגודל המדגם הוא 16?
 - ג. הסבר את ההבדל בתשובות של שני הסעיפים.

9) בקזינו ישנה רולטה. על הרולטה רשומים המס' הבאים כמוראה בשרטוט:



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.

- א. בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.
- ב. מה התוחלת ומה השונות של סכום הזכייה?
- ג. אם האדם ישחק את המשחק 5 פעמים מה התוחלת ומה השונות של ממוצע סכום הזכייה בחמשת המשחקים?
- ד. אם האדם משחק את המשחק 50 פעם מה ההסתברות שבסה"כ יזכה ב-1050 ₪ ומעלה?

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

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

12) אורך צינור שמפעל מייצר הינו עם ממוצע של 70 ס"מ וסטיית תקן של 10 ס"מ. א. נלקחו באקראי 100 מוטות, מה ההסתברות שממוצע אורך המוטות יהיה בין 68 ל 78 ס"מ? ב. יש לחבר 2 בניינים באמצעות מוטות. המרחק בין שני הבניינים הינו 7200 ס"מ. מה ההסתברות ש-100 המוטות יספיקו למלאכה? ג. מה צריך להיות גודל המדגם המינימאלי, כדי שבהסתברות של 5% ממוצע המדגם יהיה קטן מ-69 ס"מ. היעזרו במשפט הגבול המרכזי.

13) נתון משתנה מקרי בדיד בעל פונקציית ההסתברות הבאה:

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

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

14 נתון ש- $X \sim N(\mu, \sigma^2)$. דגמו 5 תצפיות מאותה התפלגות והתבוננו במוצע

המדגם $\bar{X}$. לכן: $P(\bar{X} > \mu)$ יהיה (בחרו בתשובה הנכונה):

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

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

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

16 נתון ש- $X \sim N(\mu, \sigma^2)$. אם נדגום n תצפיות מתוך ההתפלגות ונגדיר: $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$,

אזי (בחרו בתשובה הנכונה):

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

17 משקל חפיסת שוקולד בקו ייצור מתפלג עם ממוצע 100 גרם. החפיסות נארזות בקרטון המכיל 36 חפיסות שוקולד אקראיות. ההסתברות שהמשקל הממוצע של חפיסות השוקולד בקרטון יהיה מעל 99 גרם הוא 0.9932.

- א. מהי סטיית התקן של משקל חפיסת שוקולד בודדת?
- ב. מה הסיכוי שמתוך 4 קרטונים בדיוק קרטון אחד יהיה עם משקל ממוצע לחפיסה הנמוך מ-100 גרם?

(18) משתנה מקרי כלשהו מתפלג עם סטיית תקן של 20. מה הסיכוי שאם נדגום 100 תצפיות בלתי תלויות מאותה התפלגות אזי ממוצע המדגם יסטה מתוחלתו בפחות מ-2?

(19) מספר המכוניות הנכנסות לחניון "בציר" במשך היום מתפלג פואסונית עם קצב של מכונית אחת לדקה. שומר מסר נתונים על מספר המכוניות שנכנסות בכל שעה לגבי 40 שעות שאסף נתונים. מה ההסתברות שממוצע מספר המכוניות שנכנסו לחניון לשעה בשעות אלה יהיה לפחות 63?

(20) הוכיחו שאם משתנה מתפלג כלשהו עם תוחלת μ ושונות σ^2 , ומבצעים מדגם בגודל n של תצפיות בלתי תלויות מהמשתנה, אזי מתקיימות התכונות הבאות

$$\text{לגבי ממוצע המדגם: } E(\bar{x}) = \mu \text{ ו- } V(\bar{x}) = \frac{\sigma^2}{n}.$$

תשובות סופיות:

(1) א. כלל הסטודנטים במכללה שסיימו סטטיסטיקה א. ב. ציון.

ג. ממוצע: 78, סטיית תקן: 15.

ה. 78.

(2) א. להלן טבלה:

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

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

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

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

(4) א. 0.8413 ב. 0.0013

(5) א. 0 ב. 0

(6) א. 0.0465 ב. 27.71

(7) א. 0 ב. 0.1587

(8) א. 0.5468 ב. 0.6826

(9) א. להלן טבלה:

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

ב. התוחלת: 22.5, השונות: 68.75

ג. התוחלת: 22.5, השונות: 13.75

(10) 0.0475

(11) 0.1814

(12) א. 0.9772 ב. 0.0228 ג. 271

(13) 0.5

(14) ב'

(15) ד'

(16) ג'

(17) א. 2.429 ב. 0.25

(18) 0.6826

(19) 0.0071

(20) שאלת הוכחה.

התפלגות סכום תצפיות בלתי תלויות ומשפט הגבול המרכזי:

רקע:

כעת נדון בסטטיסטי המבטא את סכום התצפיות במדגם: $T = \sum_{i=1}^n X_i$.
 כאשר כל התצפיות נדגמו באקראי מאותה אוכלוסייה, כלומר, היו: $X_1, \dots, X_n$ -
 משתנים מקריים בלתי תלויים בעלי התפלגות זהה שתוחלתה μ ושונותה σ^2 אזי:
 התוחלת והשונות של סכום התצפיות: $E(T) = n\mu$, $V(T) = n\sigma^2$.

דגימה מתוך התפלגות נורמלית:

אם: $X \sim N(\mu, \sigma^2)$, אזי: $Z = \frac{T - n\mu}{\sqrt{n\sigma^2}}$, $T \sim N(n\mu, n\sigma^2)$.

משפט הגבול המרכזי:

אם X מתפלג כלשהו וידוע כי: $E(X) = \mu$, $V(X) = \sigma^2$, אזי עבור מדגם מספיק גדול (לפחות 30): $T \sim N(n\mu, n\sigma^2)$.

דוגמה (פתרון בהקלטה):

- בעיר מסוימת המשכורת הממוצעת של עובד הינה 8000 ₪. עם סטיית תקן של 2000 ₪. נדגמו 100 עובדים מהעיר שמפקידים את משכורותיהם לסניף בנק.
- מה התוחלת וסטיית התקן של סך המשכורות שיופקדו לסניף הבנק על ידי העובדים הללו?
 - מה ההסתברות שלסניף יופקד פחות מ-780 אלף ₪ ע"י אותם עובדים?

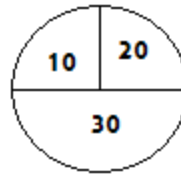
שאלות:

- (1) המשקל באוכלוסייה מסוימת מתפלג נורמאלית עם תוחלת של 60 ק"ג וסטיית תקן של 10 ק"ג.
- מה הסיכוי שאדם אקראי מהאוכלוסייה ישקול מתחת ל-65 ק"ג?
 - מה הסיכוי שהמשקל הממוצע של 4 אנשים אקראיים יהיה מתחת ל-65 ק"ג?
 - מה הסיכוי שהמשקל הכולל של 4 אנשים אקראיים יהיה מתחת ל-240 ק"ג?
- (2) נפח יין בבקבוק מתפלג נורמאלית עם תוחלת של 750 מ"ל וסטיית תקן של 20 מ"ל. אדם קנה מארז של 4 בקבוקי יין.
- מהי התוחלת ומהי סטיית התקן של נפח היין במארז?
 - את היין שבמארז האדם מזג לכלי שקיבולתו 3.1 ליטר. מה ההסתברות שהיין יגלוש מהכלי?
 - אם לא היה נתון שנפח היין מתפלג נורמאלית. האם התשובה לסעיף א' הייתה משתנה? האם התשובה לסעיף ב' הייתה משתנה?
- (3) בספר כלשהו 500 עמודים. קצב הקריאה הממוצע הוא עמוד אחד ב-4 דקות עם סטיית תקן של 1 דקות.
- מה ההסתברות לסיים את הפרק הראשון (40 עמודים) תוך שעתיים וחצי?
 - מהו האחוזון ה-95 לזמן סיום קריאת הספר?
- (4) במגדל נבנו 40 יחידות דיור. כמו כן נבנו 135 מקומות חנייה לבניין. להלן פונקציית ההסתברות של מספר המכוניות ליחידת דיור:

x	1	2	3	4	5
$P(X = x)$	0.1	0.2	0.3	0.25	0.15

- נניח שמספר המכוניות ליחידת דיור בלתי תליות זו בזו ועם אותה פונקציית ההסתברות לכל יחידת דיור (אין צורך בתיקון רציפות).
- מהי ההסתברות שיהיה מקום בחניון המגדל לכל מכוניות הבניין?
 - בהינתן ויש מקום במגדל לכל המכוניות, מה הסיכוי שבפועל מספר המכוניות נמוך מ-130?

(5) בקזינו ישנה רולטה עליה מסומנים המספרים הבאים:



אדם מסובב את הרולטה וזוכה בסכום הרשום.

א. אם האדם משחק 50 פעמים, מה ההסתברות שבסך הכול יזכה בסכום של 1050 ₪ ומעלה?

ב. האדם מגיע בכל יום לקזינו ומשחק 50 פעם, עד אשר מגיע היום בו הוא יזכה ב-1050 ₪ ומעלה. מה התוחלת ומהי השונות של מספר הימים שיבלה בקזינו?

(6) נתון ש- $X_i \sim \exp(\lambda=1)$, כאשר: $i=1, 2, \dots, 100$.

$$\text{חשבו את הסיכוי: } P\left(\sum_i X_i \geq 115\right)$$

(7) אורך חיי סוללה (בשעות) הוא בעל פונקציית הצפיפות הבאה: $0 < x < 1$, $f(x) = 2x$. ברגע שסוללה מתרוקנת מחליפים אותה במיידית בסוללה אחרת. כמה סוללות יש להחזיק במלאי אם רוצים שבסיכוי של 90% לפחות המלאי יספיק עבור 35 שעות לפחות?

תשובות סופיות:

(1) א. 0.6915 ב. 0.8413 ג. 0.5

(2) א. תוחלת: 3000 מ"ל, סטיית תקן: 40 מ"ל. ב. 0.0062

ג. סעיף א' - לא משתנה, סעיף ב' - לא פתיר, התבסס על התפלגות נורמלית.

(3) א. 0.0571 ב. 2036.8

(4) א. 0.883 ב. 0.7949

(5) א. 0.8997 ב. תוחלת 1.111, שונות 0.1239

(6) 0.0668

(7) 56

חוק המספרים הגדולים:

רקע:

חוק המספרים הגדולים מתייחס להשפעת הגדלת גודל המדגם על הסיכוי של פרופורציית המדגם (או ממוצע המדגם) להיות קרובה מהפרופורציה האמיתית (או מהממוצע האימתי).

החוק לגבי פרופורציה:

נניח שמבצעים מדגם מקרי מתוך אוכלוסייה אינסופית בה פרופורציית ההצלחות היא p , ככל שהמדגם גדול יותר: כך הסיכוי שפרופורציית המדגם ($\hat{p}$) תהיה בקרבת הפרופורציה באוכלוסייה (P) גבוה יותר.

ולכן הסיכוי לקבל ערך חריג הרחוק מהפרופורציה של האוכלוסייה קטן יותר. בכתיבה מתמטית רושמים את חוק המספרים הגדולים לגבי הפרופורציה באופן

$$\text{הבא: } \lim_{n \rightarrow \infty} \hat{P}_n = P$$

בספרות המקצועית קוראים לחוק הזה החוק החזק של המספרים הגדולים. את החוק החלש של המספרים הגדולים רושמים באופן הבא: $P(|\hat{P} - P| < \varepsilon) \xrightarrow{n \rightarrow \infty} 1$.

הערה: ככל שהמדגם גדל הסיכוי שפרופורציית המדגם תהיה בדיוק הפרופורציה האמיתית הולך וקטן.

החוק לגבי ממוצע: נניח שמבצעים מדגם מקרי מתוך התפלגות שבה התוחלת μ והשונות סופית. ככל שהמדגם גדול יותר, כך הסיכוי שממוצע המדגם ($\bar{X}$) יהיה בקרבת הממוצע באוכלוסייה (μ) גבוה יותר. לכן, הסיכוי לקבל ערך חריג הרחוק מהממוצע של האוכלוסייה קטן יותר. בכתיבה מתמטית רושמים את חוק המספרים הגדולים לגבי הממוצע באופן הבא: $\lim_{n \rightarrow \infty} \bar{X}_n = \mu$.

בספרות המקצועית קוראים לחוק הזה החוק החזק של המספרים הגדולים. את החוק החלש של המספרים הגדולים רושמים באופן הבא: $P(|\bar{X} - \mu| < \varepsilon) \xrightarrow{n \rightarrow \infty} 1$.

דוגמה (פתרון בהקלטה):

באוכלוסייה מסוימת 20% מהאוכלוסייה מובטלת. איזה סיכוי יותר גבוה? במדגם בגודל 100 פרופורציית המובטלים תסטה מהפרופורציה האמיתית בלא יותר מ-4%. במדגם בגודל 200 פרופורציית המובטלים תסטה מהפרופורציה האמיתית בלא יותר מ-4%. הסבירו.

שאלות:

- (1) באוכלוסייה מסוימת 20% מהאוכלוסייה מובטלת. איזה סיכוי יותר גבוה?
א. אחד מתוך מדגם של חמישה יהיה מובטל.
ב. שניים מתוך עשרה יהיו מובטלים. הסבירו וחשבו.
- (2) באוכלוסייה מסוימת 20% מהאוכלוסייה מובטלת. איזה סיכוי יותר גבוה?
א. לפחות 3 מתוך 10 יהיו מובטלים
ב. לפחות 30 מתוך 100 יהיו מובטלים. הסבירו.
- (3) גובה של אוכלוסייה מסוימת מתפלג נורמלית עם ממוצע 170 ס"מ וסטיית תקן של 10 ס"מ. דוגמים 4 אנשים באקראי.
א. מה ההסתברות שהגובה הממוצע שלהם יהיה מעל 176 ס"מ.
ב. כיצד התשובה לסעיף הקודם הייתה משתנה אם הינו מגדילים את גודל המדגם? נמקו.
- (4) ידוע שבהצעת חוק מסוימת תומכים 60% מהציבור. נדגמו באקראי 10 אנשים.
א. מה ההסתברות שבדיוק 60% מהמדגם תומכים בהצעת החוק?
ב. כיצד התשובה הייתה משתנה אם היו דוגמים 20 אנשים?
- (5) שני חוקרים ביצעו מדגם מאותה אוכלוסייה. חוקר א דגם 20 תצפיות והשני דגם 40 תצפיות וכל אחד מהם חישב את ממוצע המדגם: $\bar{X}_{20}$ ו- $\bar{X}_{40}$. ידוע שהתפלגות היא נורמלית ושהתוחלת באוכלוסייה הינה 500. בסעיפים הבאים נמקו אילו הסתברויות מהזוגות המוצגים גדולה יותר או שהן שוות ונמקו.
א. $P(\bar{X}_{20} > 500)$ או $P(\bar{X}_{40} > 500)$
ב. $P(480 < \bar{X}_{20} < 520)$ או $P(480 < \bar{X}_{40} < 520)$
ג. $P(\bar{X}_{20} < 490)$ או $P(\bar{X}_{40} < 490)$
- (6) נתון ש: $X \sim G(P=0.1)$. מבצעים מדגם אקראי בגודל n מההתפלגות הזו ומחשבים את ממוצע המדגם: $\bar{X}_n$. הוכיחו: $\lim_{n \rightarrow \infty} \bar{X}_n = 10$.

תשובות סופיות:

- (1) אחד מתוך מדגם של חמישה יהיה מובטל.
- (2) לפחות 3 מתוך 10 יהיו מובטלים.
- (3) א. 0.1151. ב. קטנה.
- (4) א. 0.2508. ב. קטן.
- (5) א. $P(\bar{X}_{40} > 500) = P(\bar{X}_{20} > 500)$.
- ב. $P(480 < \bar{X}_{40} < 520) > P(480 < \bar{X}_{20} < 520)$.
- ג. $P(\bar{X}_{40} < 490) < P(\bar{X}_{20} < 490)$.
- (6) שאלת הוכחה.

מבוא להסתברות וסטטיסטיקה

פרק 41 - אי שוויונים בהסתברות

תוכן העניינים

184	1. אי שוויון מרקוב
188	2. אי שוויון צ'בישב

אי-שוויון מרקוב:

רקע:

אי-שוויון מרקוב רלבנטי לשימוש עבור כל משתנה מקרי אי-שלילי.

הפרמטר a הינו ממשי וחיובי ואז חייב להתקיים ש: $P(X \geq a) \leq \frac{E[X]}{a}$,

לכן מתקיים גם ש: $P(X < a) \geq 1 - \frac{E[X]}{a}$.

דוגמה:

אורך חיים של מכשיר מתפלג עם תוחלת של 500 שעות. חשבו לפי אי-שוויון מרקוב את ההסתברות שאורך חיים של מכשיר יהיה לפחות 1500 שעות.

X = אורך חיים של מכשיר (רציף).

$$X \geq 0, E[X] = 500$$

$$P(X \geq 1500) \leq \frac{E[X]}{a} = \frac{500}{1500} = \frac{1}{3}$$

לכן, $0 \leq P(X \geq 1500) \leq \frac{1}{3}$.

שאלות:

- (1) ידוע מניסיון העבר כי ציון במבחן הגמר של סטודנט הוא משתנה מקרי שתוחלתו 65. מצאו חסם עליון להסתברות שציון מבחן הגמר של סטודנט יהיה לפחות 75.
- (2) התפלגות מספר הילדים למשפחה במדינה מסוימת היא עם תוחלת של 2 ילדים. נלקחו 5 משפחות אקראיות. העריכו את הסיכוי שבסה"כ בחמשת המשפחות יש יותר מ-15 ילדים.
- (3) X משתנה מקרי רציף אי-שלילי, שתוחלתו 15. האם ייתכן ש: $P(X > 65) = 0.3$?
- (4) X משתנה מקרי בדיד, המקיים: $X > -2$, ותוחלתו 6. מצאו חסם תחתון ל- $P(X < 10)$.
- (5) X משתנה מקרי המקיים: $P(X \geq 0) = 1$ ו- $s > 0$ קבוע ממשי. הוכיחו כי: $P(X < sE(X)) \geq \frac{s-1}{s}$.
- (6) נתון ש- $X_i \sim P(\lambda)$, $i = 1, 2, \dots, n$ הם משתנים בלתי תלויים. מצאו חסמים להסתברות ש- $\sum_{i=1}^n X_i \geq 3n\lambda$.
- (7) הוכיחו את אי-שוויון מרקוב. רמז: היעזרו במשתנה אינדיקטור המקבל את הערך 1 כאשר $X \geq a$.

(8) בשכונה חדשה בונים n בתים חדשים הנבנים בחלקת אדמה עגולה. כל בית נצבע בלבן בסיכוי p ללא תלות בבתיים האחרים. בניין שלא נצבע בלבן נצבע באפור. בית לבן הוא בית בודד אם הוא נמצא בין שני בתים אפורים. נגדיר את X להיות מספר הבתים הלבנים הבודדים.

א. מצאו את התוחלת של X .

ב. כעת נניח ש- $p = \frac{1}{4}$. הראו שהסיכוי שמספר הבתים הלבנים הבודדים

$$\text{יהיה קטן מ-} \frac{n}{4} \text{ הוא לפחות } \frac{7}{16}.$$

(9) הוכיחו את אי-שוויון צ'בישב: אם X הוא משתנה מקרי שתוחלתו ושונויותו הן סופיות, אז לכל ערך k חיובי מתקיים: $P\{|X - E(X)| \geq k\} \leq \frac{\text{Var}(X)}{k^2}$.

רמז: היעזרו באי-שוויון מרקוב.

תשובות סופיות:

$$P(X \geq 75) \leq \frac{65}{75} = \frac{13}{15} \quad (1)$$

$$0 \leq P\left(\sum_{i=1}^n X_i \geq 3n\lambda\right) \leq \frac{1}{3} \quad (2)$$

(3) לא יתכן.

$$\frac{1}{3} \quad (4)$$

(5) שאלת הוכחה.

$$0 \leq P\left(\sum_{i=1}^n X_i \geq 3n\lambda\right) \leq \frac{1}{3} \quad (6)$$

(7) שאלת הוכחה.

(8) א. $n \cdot p(1-p)^2$. ב. שאלת הוכחה.

(9) שאלת הוכחה.

אי-שוויון צ'בישב:

רקע:

אם X הוא משתנה מקרי שתוחלתו ושונותו הן סופיות, אז לכל ערך a חיובי

$$P\{|X - E(X)| \geq a\} \leq \frac{\text{Var}(X)}{a^2} \quad \text{מתקיים:}$$

$$P\{|X - E(X)| < a\} \geq 1 - \frac{\text{Var}(X)}{a^2} \quad \text{מכאן גם נובע שמתקיים:}$$

אי-שוויון צ'בישב נותן חסמים להסתברות סימטרית סביב התוחלת ללא צורך בידעת ההתפלגות של המשתנה המקרי X .

דוגמה:

נתון משתנה מקרי עם סטיית תקן של 3. האם יתכן שההסתברות שהסטייה של המשתנה המקרי מתוחלתו תהיה קטנה מ-5 היא 0.6?

$$\sigma(X) = 3$$

$$P\{|X - E(X)| < a\} \geq 1 - \frac{\text{Var}(X)}{a^2} \quad \text{נציב:}$$

$$P\{|X - E(X)| < 5\} \geq 1 - \frac{3^2}{5^2} = 1 - \frac{9}{25} = \frac{16}{25} = 0.64$$

$$P\{|X - E(X)| < 5\} \neq 0.6, \quad \text{לכן לא יתכן.}$$

שאלות:

- (1) מצאו חסמים להסתברויות הבאות עבור משתנה מקרי רציף בעל תוחלת 8 וסטית תקן 3:
- א. $p(2 < x < 14)$.
- ב. $p(|x - 8| \geq 9)$.
- (2) האם קיים משתנה מקרי X בעל תוחלת μ וסטית תקן σ שעבורו מתקיים $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.7$? הסבירו.
- (3) מספר המטוסים המגיעים לנמל תעופה ב-20 דקות מתפלג התפלגות פואסונית עם תוחלת של 100. היעזרו באי-שוויון צ'בישב כדי למצוא גבול תחתון להסתברות שמספר המטוסים המגיעים בתקופה בת 20 דקות נתונה תהיה בין 80 ל-120.
- (4) מטילים מטבע 120 פעמים. מה ניתן להגיד על הסיכוי שהתוצאה עץ תתקבל בין 50 ל-70 פעמים לפי אי-שוויון צ'בישב?
- (5) מתוך קו יצור של רכיבים שאורכם הממוצע הנו 10 ס"מ ושונותם 3 סמ"ר יש לקחת מדגם. מהו גודל המדגם שיבטיח שבהסתברות של 0.9 לפחות ימצא ממוצע המדגם בין 9 ל-11 ס"מ?
- (6) אחוז התומכים במפלגה מסוימת הנו 40%. נלקח מדגם מקרי בגודל 200. תנו חסם תחתון לכך שאחוז התומכים במדגם יהיה בין 35% ל-45%.
- (7) בוחרים קוד n ספרתי באופן מקרי.
- א. עבור $n = 10$, העריכו את ההסתברות שממוצע הספרות במספר יסטה מתוחלתו בלפחות 1.
- ב. מה אורך הקוד המינימלי (n) שיבטיח שבהסתברות של לפחות 95%, ממוצע הספרות יסטה מתוחלתו בפחות מ-0.75?

(8) בעיר מסוימת ל-5% מהמשפחות אין מכונית, ל-20% יש מכונית אחת, ל-35% יש שתי מכוניות, ל-30% שלוש מכוניות וליתר ארבע מכוניות. נניח שמספר המשפחות בעיר גדול מאוד. העריכו את ההסתברות שמספר המכוניות הכולל בעשר משפחות אקראיות יהיה לפחות 17 ולכל היותר ל-27.

(9) $X_1, X_2, \dots, X_n$ הם משתנים מקיים המתפלגים גיאומטרית עם פרמטר p באופן

$$P\left(\frac{\sum_{i=1}^n X_i}{n} \geq \frac{2}{p}\right) \leq \frac{1-p}{n} : 0 < p < 1. \text{ הוכיחו שמתקיים:}$$

(10) $X_1, X_2, \dots, X_n$ הם משתנים מקיים המתפלגים פואסונית עם פרמטר $\lambda_i = i$, באופן

$$T = \sum_{i=1}^n X_i : \text{נסמן את:}$$

$$P(|2T - n(n+1)| < 2n) \geq \frac{n-1}{2n} : \text{הוכיחו שמתקיים:}$$

תשובות סופיות:

- (1) א. בין $\frac{3}{4}$ ל-1. ב. בין 0 ל- $\frac{1}{9}$.
- (2) לא יתכן.
- (3) 0.75.
- (4) לפחות 0.7.
- (5) לפחות 30.
- (6) 0.52.
- (7) א. לכל היותר 0.825. ב. 294.
- (8) 0.7056.
- (9) שאלת הוכחה.
- (10) שאלת הוכחה.

מבוא להסתברות וסטטיסטיקה

פרק 42 - חישוב תוחלת ושונות על ידי פירוק לאינדקטורים

תוכן העניינים

1. כללי 192

חישוב תוחלת ושונות על ידי פירוק לאינדיקטורים:

רקע:

נלמד שיטה לחישוב תוחלת ושונות של משתנה מקרי, על ידי פירוק לסכום של משתני אינדיקטור. אינדיקטור הינו משתנה שפונקציית ההסתברות של נראית כך:

X	1	0
$P(X)$	P	$1-P$

נגיד ש- X_i הינו משתנה אינדיקטור כאשר: $i = 1, 2, \dots, n$ ו- $X = \sum_{i=1}^n X_i$.

ניעזר בנוסחאות תוחלת ושונות סכום משתנים מקרים כדי לחשב את התוחלת והשונות של X .

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$$

$$V(X) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \cdot \sum_{i < j} \text{COV}(X_i, X_j)$$

כאשר עבור משתנים אינדיקטורים מתקיים ש:

$$E(X_i) = P(X_i = 1)$$

$$V(X_i) = P(X_i = 1) \cdot P(X_i = 0)$$

$$\text{COV}(X_i, X_j) = P(X_i = 1, X_j = 1) - P(X_i = 1)P(X_j = 1)$$

דוגמא (פתרון בהקלטה):

יוסי החליט להזמין 8 חברים למסיבת יום הולדתו. הוא הכין 8 הזמנות שעליהן רשם את השם של כל אחד מהחברים. ההזמנות הוכנסו למעטפות וחולקו באקראי ל-8 החברים. נסמן ב- X את מספר ההזמנות שהגיעו לחבר הנכון. חשבו את $E(X)$ ואת $V(X)$.

שאלות:

- (1) יהיו X ו- Y משתני אינדיקטורים. הוכיחו ש:
- $E(X) = P(X=1)$.
 - $V(X) = P(X=1) \cdot [1 - P(X=1)]$.
 - $COV(X, Y) = P(X=1, Y=1) - P(X=1)P(Y=1)$.
- (2) 400 אנשים נבחרו מכלל האוכלוסייה.
- א. חשבו את הסיכוי שביום מסוים בשנה יהיה בדיוק אדם אחד מתוך ה-400 שיש לו יום הולדת.
 - ב. נגדיר את X_i משתנה אינדיקטור המקבל את הערך 1 אם ביום i בדיוק אדם אחד מתוך ה-400 עם יום הולדת באותו היום. חשבו את התוחלת והשונות של X_i .
 - ג. חשבו את התוחלת והשונות של מספר הימים בשנה שבהם יש יום הולדת בדיוק לאחד מ-400 האנשים הללו.
- (3) 3 משחקים הוכנסו באקראי ל-5 מגרות. מגירה יכולה להכיל יותר ממשחק אחד. נסמן ב- W את מספר המגרות בהן בדיוק משחק אחד. חשבו את התוחלת והשונות של W על ידי פירוק לאינדיקטורים.
- (4) A, B ו- C הם שלושה מאורעות כך ש: $P(A) = 0.3, P(B) = 0.2, P(C) = 0.1$.
- נגדיר את Y להיות מספר המאורעות מתוך השלושה שמתקיימים. חשבו את התוחלת והשונות של Y כאשר:
 - א. המאורעות בלתי תלויים זה בזה.
 - ב. $C \subset B \subset A$.
 - ג. A, B ו- C זרים זה לזה.
- (5) נטיל קובייה 10 פעמים. נסמן ב- W את מספר התוצאות השונות שהתקבלו.
- א. מצאו את $E(W)$.
 - ב. מצאו את $V(W)$.

- (6) נסדר בשורה 6 כוסות קולה ו-4 כוסות מים. רצף של שתי כוסות נקרא "ג'נק", אם שתי הכוסות הן ברצף של קולה. נסמן ב- X את מספר הרצפים מסוג "ג'נק" שיש לשתי כוסות. למשל, הסידור הבא:
- קולה, מים, קולה, מים, קולה, מים, קולה, מים, קולה, קולה, $X = 2$.
- חשבו את התוחלת והשונות של X .
- (7) נסדר בשורה n זוגות גרביים באקראי (בסך הכול $2n$ גרביים). חשבו את התוחלת והשונות של מספר הזוגות מתוך n הזוגות שבהם זוג הגרביים אינם עומדים זה לצד זה.
- (8) בקייטנה 100 ילדים. מחלקים לכל ילד 2 ארטיקים מתוך 200 הארטיקים שנרכשו לקייטנה. מתוך 200 הארטיקים שנרכשו 100 בטעם תות ו-100 הם בטעם לימון. נסמן ב- X את מספר הילדים שקיבלו 2 ארטיקים בטעמים שונים. נסמן ב- Y את מספר הילדים שקיבלו שני ארטיקים בטעם לימון.
- א. חשבו את התוחלת והשונות של X .
- ב. בטאו את Y כפונקציה של X וחשבו את התוחלת והשונות של Y .
- ג. מהי השונות המשותפת של X ו- Y ?

תשובות סופיות:

- (1) שאלת הוכחה.
- (2) א. 0.3667. ב. תוחלת: 0.3667, שונות: 0.2322.
- ג. תוחלת: 133.85, שונות: 88.89.
- (3) תוחלת: 1.92, שונות: 1.1136.
- (4) א. תוחלת: 0.6, שונות: 0.46. ב. תוחלת: 0.6, שונות: 1.04.
- ג. תוחלת: 0.6, שונות: 0.24.
- (5) א. 5.03. ב. 0.568.
- (6) תוחלת: 3, שונות: $\frac{2}{3}$.
- (7) תוחלת: $n-1$, שונות: $1 - \frac{1}{n} + \frac{n-1}{2(2n-1)}$.
- (8) א. תוחלת: 50.251, שונות: 25.126.
- ב. $Y = -0.5X + 50$. ג. -12.563.

מבוא להסתברות וסטטיסטיקה

פרק 43 - אמידה נקודתית

תוכן העניינים

1. אומדן חסר הטייה 195
2. אומדן ניראות מקסימלית 202

אומד חסר הטייה:

רקע:

$\hat{\theta}$ יהיה אומד חסר הטיה ל- θ , אם התוחלת של $\hat{\theta}$ תהיה שווה ל- θ : $E(\hat{\theta}) = \theta$.

דוגמה (פתרון בהקלטה):

המשתנה X הוא בעל פונקציית ההסתברות הבאה:

3	2	1	X
4θ	$1 - 60\theta$	2θ	הסתברות

מעוניינים לאמוד את θ על סמך שתי תצפיות מההתפלגות: X_2 ו- X_1 .

א. הראו שהאומד: $T_1 = \frac{2X_1 + X_2}{2}$, הוא אומד מוטה ל- θ .

הטיה של אומד היא: $E(\hat{\theta}) - \theta$. כמובן שלאומד חסר הטיה אין הטיה.

ב. מהי ההטיה של האומד T_1 ?

ג. תקנו את T_1 , כך שיהיה אומד חסר הטיה.

אם יש שני אומדים חסרי הטיה עדיף זה עם השונות היותר קטנה.

ד. מוצא האומד הבא: $T_3 = 1.5X_1 - X_2 - 1$.

האם הוא עדיף על האומד שהצעת בסעיף ג'?

אם $\hat{\theta}$ אומד חסר הטיה ל- θ , אז $g(\hat{\theta})$ יהיה אומד חסר הטיה עבור $g(\theta)$, רק אם g תהיה לינארית.

ה. מצאו אומד חסר הטיה ל: $P(X = 3)$.

אומד חסר הטיה לשונות האוכלוסייה σ^2 : $S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$.

ו. מצאו אומד חסר הטיה לשונות של X .

תזכורות חשובות:

אם: $Y = aX + b$, אזי: $E(Y) = aE(X) + b$, $V(Y) = a^2 \cdot V(X)$, $\sigma_Y = |a| \sigma_X$.

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

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

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

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

שאלות:

- (1) הציון במבחן מסוים של תלמידי כתה ח' הנו משתנה מקרי בעל תוחלת μ וסטיית תקן 10. כדי לאמוד את התוחלת μ , נלקח מדגם של 5 ציונים: $X_1, \dots, X_5$. שלושה חוקרים הציעו אומדים לתוחלת על סמך מדגם זה:

$$T_1 = \frac{X_1 + \dots + X_5}{5} \quad \text{חוקר א' הציע:}$$

$$T_2 = \frac{2X_1 - X_3 + X_4}{2} \quad \text{חוקר ב' הציע:}$$

$$T_3 = \frac{2X_1 + X_3}{2} \quad \text{חוקר ג' הציע:}$$

- איזה מן האומדים הוא חסר הטיה?
- הציעו תיקון לאומד המוטה כך שיהיה חסר הטיה.
- במדגם התקבלו הציונים הבאים: 100, 82, 58, 78, 65. חשבו את האומדנים המתקבלים עבור האומדים חסרי ההטיה.
- איזה מבין שני האומדים חסרי ההטיה עדיף? נמקו.

- (2) כדי לאמוד את המשקל הממוצע של הנשים בארה"ב, נבחר מדגם של $2n$ נשים. נסמן את שונות הגובה ב- σ^2 . הוצעו שני אומדים לממוצע המשקל על סמך מדגם

$$T_1 = \frac{1}{n} \sum_{i=1}^n X_i, \quad T_2 = \frac{1}{2n} \sum_{i=1}^{2n} X_i \quad \text{זה:}$$

- בדקו לגבי כל אומד אם הוא בלתי מוטה.
- איזה אומד עדיף? נמקו.

- (3) $X \sim B(n, p)$. כלומר, X הינו משתנה מקרי המתפלג בינומית עם פרמטר P (סיכוי להצלחה בניסיון בודד) במדגם בגודל n .

- פתחו אומד חסר הטיה ל- P .
- מהו אומד חסר הטיה לסיכוי לכישלון בניסיון בודד?
- מהו אומד חסר הטיה ל- $E(X)$?
- מצאו אומד חסר הטיה ל- $E(X^2)$.

(4) בתיק מניות שתי מניות. מספר המניות שיעלו ביום מסוים הוא משתנה מקרי התלוי בפרמטר לא ידוע: θ , $0 \leq \theta \leq 2$.

פונקציית ההסתברות של X - מספר המניות שיעלו ביום מסוים:

$$P(X=0) = 1 - \frac{\theta}{2}, \quad P(X=1) = \frac{\theta}{3}, \quad P(X=2) = \frac{\theta}{6}$$

א. מצאו אומד בלתי מוטה ל- θ , שמתבסס על מספר המניות שיעלו ביום מסוים.

ב. מצאו אומד בלתי מוטה ל- θ , שמתבסס על מספר המניות שעלו ביום,

במשך שלושה ימים - X_1, X_2, X_3 (לכל אחד מהם אותה התפלגות כנ"ל

והם בלתי תלויים).

(5) בקרב המטפלות בת"א, מספר התינוקות שבטיפולן הוא משתנה מקרי בעל התפלגות התלויה בפרמטר θ באופן הבא:

הסיכוי שמטפלת תטפל בתינוק אחד בלבד הוא 3θ ,

הסיכוי שמטפלת תטפל ב-2 תינוקות הוא $1 - 4\theta$,

הסיכוי שמטפלת תטפל ב-3 תינוקות הוא θ .

במדגם מיקרי של 4 מטפלות מת"א, נמצא כי שתיים מהם מטפלות בתינוק אחד בלבד, אחת מהן בשנים ואחת השלושה תינוקות.

א. מצאו אומד חסר הטיה לפרמטר θ על סמך תצפית בודדת.

ב. מצאו אומד חסר הטיה לפרמטר θ על סמך 4 תצפיות.

ג. מהו האומדן לפרמטר θ על סמך תוצאות המדגם.

ד. מצאו אומד חסר הטיה לסיכוי שלמטפלת בת"א תטפל בתינוק בודד אחד.

ה. מצאו אומדים חסרי הטיה לתוחלת ולשונות של מספר התינוקות בטיפול אצל מטפלת מת"א. חשבו אומדנים.

(6) קבעו אילו מהטענות הבאות נכונות:

א. אם T הוא אומד בלתי מוטה עבור פרמטר θ , אז $5T$ אומד בלתי מוטה

עבור הפרמטר 5θ .

ב. אם T הוא אומד בלתי מוטה עבור פרמטר θ , אז T^2 אומד בלתי מוטה

עבור הפרמטר θ^2 .

- (7) במפעל שתי מכונות המייצרות מוצרים. במכונה הראשונה ההסתברות שמכשיר תקין היא p , ובמכונה השנייה ההסתברות שמכשיר תקין היא $2p$. דוגמים 20 מכשירים מהייצור של כל מכונה. נסמן ב- X את מספר המכשירים התקינים שיוצרו על ידי המכונה הראשונה, וב- Y את מספר המכשירים התקינים שיוצרו על ידי המכונה השנייה. איזה מבין האומדים הבאים אינו אומד חסר הטיה ל- p ?

א. $\frac{X}{20}$.

ב. $\frac{Y}{20}$.

ג. $\frac{X+Y}{60}$.

ד. $\frac{2X+Y}{80}$.

- (8) יהיו T_1 ו- T_2 אומדים חסרי הטיה ובלתי תלויים לפרמטר θ .
 א. מצאו אומד חסר הטיה ל- θ^2 , המתבסס על T_1 ו- T_2 .
 ב. מצאו אומד חסר הטיה ל- $\theta(1-\theta)$, המתבסס על T_1 ו- T_2 .

- (9) נתון ש- X הינו משתנה מקרי עם תוחלת μ ושונות σ^2 . נדגמו n תצפיות בלתי תלויים מאותה אוכלוסיה.

א. הראו ש- $\sum_{i=1}^n p_i x_i$ אומד חסר הטיה ל- μ , כאשר: $\sum_{i=1}^n p_i = 1$.

ב. נתבונן במכפלת שתי התצפיות הראשונות: $X_1 \cdot X_2$.

הראו שהוא אומד חסרי הטיה ל- μ^2 .

- (10) $X_i \sim N(\mu, 1)$, כאשר: $i = 1, 2, \dots, n$. נתון שהתצפיות הינן בלתי תלויות זו בזו. מצאו אומד חסר הטיה ל- μ^2 .

(11) נתונות n תצפיות בלתי תלויות מתוך התפלגות בעלת הצפיפות הבאה :

$$f(x) = \begin{cases} \frac{1+\beta x}{2} & -1 < x < 1 \\ 0 & \text{else} \end{cases}$$

- א. הראו כי האומד $3\bar{X}$ הנו אומד בלתי מוטה ל- β .
 ב. מצאו את השונות של האומד מהסעיף הקודם.

(12) $X_1, X_2, \dots, X_n$ הינם משתנים מקריים רציפים בלתי תלויים בעלי פונקציית

$$f(x) = \begin{cases} X \cdot A & 0 \leq x \leq \theta \\ 0 & \text{אחרת} \end{cases} \quad \text{הצפיפות הבאה :}$$

- א. בטאו את ערכו של A באמצעות θ , כדי שפונקציית הצפיפות תהיה לגיטימית.
 ב. מצאו אומד חסר הטיה ל- θ , על סמך n התצפיות.

תשובות סופיות:

$$(1) \quad \text{א. } T_1 \text{ ו- } T_2. \quad \text{ב. } \frac{2}{3}T_3. \quad \text{ג. } T_1 = 76.6, T_2 = 110. \quad \text{ד. } T_1.$$

$$(2) \quad \text{א. ראו בוידאו.} \quad \text{ב. } T_2.$$

$$(3) \quad \text{א. } \frac{x}{n}. \quad \text{ב. } 1 - \frac{x}{n}. \quad \text{ג. } X. \quad \text{ד. } \theta.$$

$$(4) \quad \text{א. } \frac{3x}{2}. \quad \text{ב. } \frac{3\bar{x}}{2}.$$

$$(5) \quad \text{א. } 1 - \frac{x}{2}. \quad \text{ב. } 1 - \frac{1}{2}\bar{x}. \quad \text{ג. } 0.125. \quad \text{ד. } 3\left(1 - \frac{1}{2}\bar{x}\right).$$

$$\text{ה. לשונות } 0.917.$$

$$(6) \quad \text{א. נכון.} \quad \text{ב. לא נכון.}$$

$$(7) \quad \text{ב'}$$

$$(8) \quad \text{א. } T_1 \cdot T_2. \quad \text{ב. } T_1 - T_1 \cdot T_2.$$

$$(9) \quad \text{א. שאלת הוכחה.} \quad \text{ב. שאלת הוכחה.}$$

$$(10) \quad \bar{X}^2 - \frac{1}{n}$$

$$(11) \quad \text{א. שאלת הוכחה.} \quad \text{ב. } V(3\bar{X}) = \frac{3 - \beta^2}{n}$$

$$(12) \quad \text{א. } A = \frac{2}{\theta^2}. \quad \text{ב. } \theta = \frac{3 - \bar{X}}{2}$$

אומד נראות מקסימלית:

רקע:

להלן נלמד את שיטת הנראות המקסימלית למציאת אומדים.
נניח ש- X משתנה מקרי בדיד עם פונקציית הסתברות $P(x, \theta)$,
כאשר θ הפרמטר הבלתי ידוע.

יהיו: $X_1, X_2, \dots, X_n$ תוצאות מדגם מקרי בגודל n הנלקח מאוכלוסייה זו.
נבנה את פונקציית ההסתברות המשותפת (פונקציית הדגימה).
אם אנו יודעים את תוצאות המדגם, ולא את הפרמטר, קוראים לפונקציית הנראות
שהיא פונקציה של הפרמטר.

נגדיר את פונקציית הנראות:

$$L(\theta) = P(x_1, \theta) \cdot P(x_2, \theta) \cdot \dots \cdot P(x_n, \theta) = \prod_{i=1}^n P(x_i, \theta)$$

פונקציית הנראות היא ההסתברות לקבל את התצפית הראשונה (כפונקציה של θ),
כפול ההסתברות לקבל את התצפית השנייה, וכולי. כלומר, המשמעות של פונקציית
הנראות היא ההסתברות לקבל את המדגם שהתקבל, כפונקציה של הפרמטר המבוקש θ .

אם מדובר במשתנה רציף, נכפיל את פונקציות הצפיפות ולא את פונקציות ההסתברות:

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta) = \prod_{i=1}^n f(x_i, \theta)$$

דוגמה (פתרון בהקלטה):

הסיכוי של שחקן כדורסל לקלוע לסל הוא p (לא ידוע). השחקן זורק כדורים לסל
עד שהוא קולע בפעם הראשונה. נניח כי הזריקות בלתי תלויות זו בזו.
הכדור נכנס לסל לראשונה בניסיון השלישי.
השחקן חוזר על התהליך שוב, והפעם הכדור נכנס לסל בניסיון החמישי.
מצאו את פונקציית הנראות של p .

אומד נראות מקסימלית עבור θ הוא האומד $\hat{\theta}$, שממקסם את פונקציית הנראות $L(\theta)$. כלומר, אנו מחפשים את האומד שיגרום לכך שהמדגם המקרי שקיבלנו יהיה כמה שיותר סביר.

שלבים למציאת אומד נראות מקסימלית:

- לוקחים את פונקציית ההסתברות המשותפת של המדגם (או צפיפות משותפת אם המשתנה רציף).
- מציבים את תוצאות המדגם ומקבלים את פונקציית הנראות (פונקציה של הפרמטר הנחקר).
- מוצאים מקסימום לפונקציית הנראות (לעיתים כדאי להוסיף $\ln$ כדי להקל על המלאכה).

המשך דוגמה:

חשבו את אומדן הנראות המקסימלית עבור p .

משפט: אם $\hat{\theta}$ הוא אומד נראות מקסימלית עבור θ , אזי $g(\hat{\theta})$ הוא אומד נראות מקסימלית עבור $g(\hat{\theta})$, בהנחה והפונקציה היא חד-חד ערכית (אינווריאנטיות).

המשך דוגמה:

מצאו אומדן נראות מקסימלית לסיכוי של שחקן הכדורסל לקלוע לסל פעמיים ברצף.

שאלות:

- (1) הסיכוי של שחקן לנצח במשחק הוא p (לא ידוע).
 השחקן משחק במשחק עד אשר הוא מנצח בפעם הראשונה.
 נתון שהשחקן ניצח לראשונה רק במשחק השני.
 א. חשבו את פונקציית הנראות של p , וציירו גרף שלה.
 ב. מצאו אומדן נראות מקסימלית עבור p .
 ג. מצאו אומדן נראות מקסימלית ל- p , אם ביום אחד הוא נאלץ לשחק 4 פעמים וביום אחר הוא נאלץ לשחק 5 פעמים, עד אשר ניצח.
- (2) מספר הלקוחות שנכנסים לחנות מסוימת, מתפלג פואסונית עם תוחלת של λ לקוחות ביום.
 א. מצאו אומדן נראות מקסימלית ל- λ , על סמך מספר הלקוחות שנכנסים ביום מסוים.
 ב. מצאו אומדן נראות מקסימלית ל- λ , על סמך מספר הלקוחות שנכנסים ב- n ימים מסוימים.
- (3) הזמן שלוקח לאדם לחכות בתור מתפלג מעריכית עם פרמטר λ .
 דגמו 4 אנשים מקריים שחיכו בתור ומדדו את זמני ההמתנה שלהם.
 התוצאות שהתקבלו בדקות הן: 3, 5, 7 ו-3.
 א. פתחו אומדן נראות מקסימלית לפרמטר זה על סמך n תצפיות כלשהן.
 ב. מהו האומדן לפרמטר?
- (4) משך זמן הכנת שיעורי הבית (בשעות) של בני נוער, ביום אחד, מתפלג אחיד: $U(0, q)$.
 כדי לאמוד את θ , נשאלו ביום מסוים מספר בני נוער כמה שעות הם הכינו שיעורי-בית באותו יום.
 א. אלעד הכין ביום מסוים שיעורי בית במשך שעה שלמה. חשבו את פונקציית הנראות של θ המתבססת על תצפית זו, וציירו את הגרף שלה.
 ב. מצאו אומדן נראות מקסימלית ל- θ על סמך התצפית.
 ג. משכי הכנת שיעורי בית (שעות) של 3 בני נוער היו 1.5, 3, 1.
 ד. מצאו אומדן נראות מקסימלית ל- θ על סמך המדגם הזה.
 ד. מצאו באופן כללי אומדן נראות מקסימלית ל- θ , על סמך מדגם של n בני נוער – $X_1, \dots, X_n$.

(5) הגובה של אוכלוסייה מסוימת מתפלג נורמלית עם תוחלת ידועה של 170 ס"מ ושונות σ^2 לא ידועה.

א. מצאו אומד נראות מקסימלית עבור השונות על סמך מדגם $X_1, \dots, X_n$ מתצפיות מהאוכלוסייה.

ב. נדגמו 5 אנשים בלתי תלויים בעלי הגבהים: 170, 182, 165, 174, 174. מהו האומדן לשונות הגבהים באוכלוסייה?

(6) פתחו אומד נראות מקסימלית לפרמטר p בהתפלגות הבינומית, על סמך מדגם בגודל n , בו X הוא מספר ההצלחות במדגם.

(7) X הוא משתנה מקרי בעל פונקציית הצפיפות: $f(x) = \begin{cases} 2\theta x e^{-\theta x^2} & x > 0 \\ 0 & x \leq 0 \end{cases}$

א. מצאו אומד נראות מקסימלית ל- θ על סמך n תצפיות בלתי תלויות: $X_1, \dots, X_n$.

ב. מצאו אומד נראות מקסימלית ל- θ^2 .

(8) בכד א' 10 כדורים שחורים ו-10 לבנים ובכד ב' 5 כדורים שחורים ו-15 לבנים. דוגמים באקראי כדור, בלי לדעת מאיזה כד.

א. מצא אומד נראות מקסימלית לכד שממנו הוצא הכדור על סמך הצבע של הכדור.

ב. מהו האומדן אם הצבע הוא שחור?

(9) הזמן שלוקח ליוסי לפתור תשבץ מתפלג מעריכית עם תוחלת לא ידועה. נתנו ליוסי לפתור חמישה תשבצים ובממוצע לקח לו 32 דקות לפתור אותם.

א. מה אומדן הנראות המקסימלית לתוחלת זמן הפתרון של תשבץ על ידי יוסי (אין חובה לפתח).

ב. מה אומדן הנראות המקסימלית לסיכוי שייקח לו לפחות חצי שעה לפתור את התשבץ הבא?

- 10) מספר הלקוחות הממתינים בתור במוקד טלפוני הוא משתנה מקרי X , בעל התפלגות התלויה בפרמטר θ , באופן הבא:

X	0	1	2
$P(X)$	$4\theta^2$	$4\theta - 8\theta^2$	$1 - 4\theta + 4\theta^2$

בחמישה זמנים שונים שנבחרו באקראי נמצאו: 0, 0, 1, 0, 0 לקוחות ממתינים בתור.

- א. מצאו אומדן בשיטת הנראות המקסימלית עבור הפרמטר θ , על-סמך המדגם הנתון.
 ב. מצאו אומדן בשיטת הנראות המקסימלית לסיכוי שלא יהיו לקוחות בתור.

- 11) אדם מחזיק בידו שני מטבעות: מטבע הוגן ומטבע שאינו הוגן – שהסיכוי לקבל בו תוצאה של עץ הוא 0.2. האדם מטיל את אחד המטבעות פעמיים ומודיע לך כמה פעמים הוא קיבל עץ. אתה צריך לנחש איזה מטבע הוא הטיל: את ההוגן או את זה שאינו הוגן.

- א. מצא אומדן בשיטת הנראות המקסימלית לסוג המטבע שהוטל.
 ב. מהו האומדן אם האדם קיבל פעמיים עץ?

- 12) מעוניינים לאמוד את אחוז המובטלים באוכלוסייה. דוגמים 50 אנשים אקראיים ומתקבל ש-4 מהם מובטלים.
 א. מצא אומדן נראות מקסימלית לשיעור המובטלים באוכלוסייה.
 ב. מצא אומדן לשיעור העובדים באוכלוסייה.
 ג. מצא אומדן ליחס בין שיעור העובדים לשיעור המובטלים באוכלוסייה.

- 13) במשחק מחשב שלוש רמות משחק:
 ברמה 1 הסיכוי של יוסי לסיים את המשחק הוא 0.9.
 ברמה 2 הסיכוי של יוסי לסיים את המשחק הוא 0.7.
 ברמה 3 הסיכוי של יוסי לסיים את המשחק הוא 0.4.
 יוסי בחר ברמה מסוימת, אך אינו יודע באיזו רמה הוא בחר.
 הוא משחק במשחק ברמה שבחר פעמיים.
 א. הציעו א.נ.מ. לרמה של המשחק שיוסי שיחק, על סמך מספר הפעמים שסיים את המשחק.
 ב. אם יוסי סיים את שני המשחקים, מה יהיה האומדן לרמה?
 ג. מהו א.נ.מ. לסיכוי, שמתוך שני משחקים הוא יצליח בדיוק משחק אחד?

(14) $X_1, X_2, \dots, X_n$ מתפלגים אחיד בקטע: $[-\theta, \theta]$.

מצא אומד נראות מקסימלית לפרמטר θ .

(15) $X_1, X_2, \dots, X_n$ מתפלגים בדיד לפי פונקציית ההסתברות הבאה:

$$P(X = k) = \frac{\binom{2}{k} \cdot P^k \cdot (1-P)^{2-k}}{1 - (1-P)^2} \quad K = 1, 2$$

הוכח שא.נ.מ ל- P , הינו: $2 - \frac{2}{X}$.

(16) במכשיר חשמלי יש 2 סוללות שפועלות באופן ב"ת זו בזו, והוא מפסיק לפעול ברגע שאחת הסוללות מפסיקה לעבוד. הסיכוי של סוללה לתפקד לפחות חודש הוא P . כאשר המכשיר מפסיק לפעול מחליפים את שתי הסוללות שלו. בתחילת הניסוי נלקחו 80 מכשירים כאלה עם סוללות חדשות ולאחר חודש נמצא ש-30 מהם עדיין פועלים.

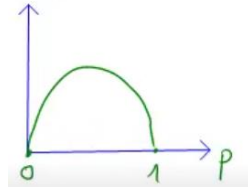
- מצא אומדן נראות מקסימלית עבור P .
- רשמו את האומד שבו השתמשתם בחלק א' באופן כללי, עבור מדגם של n מכשירים שמתוכם נמצאו Y מכשירים שעדיין פועלים לאחר חודש אחד.
- בהנחה שאורך החיים (בחודשים) של סוללה בודדת הוא מעריכי, עם פי צפיפות: $f(t) = \theta e^{-\theta t}$ עבור $t > 0$. מצא א.נ.מ. עבור θ , המבוסס על Y . מהו האומדן המתאים מן המדגם הנתון?

(17) חיוג אוטומטי של מכשיר טלפון משדר אות אחת לשתי דקות. אם לאחר 20 דקות (10 אותות חיוג) המספר שאליו מטלפנים עדיין תפוס, החיוג האוטומטי נפסק.

- רשמו את פונקציית ההסתברות של המשתנה X – מספר הפעמים שהחייגן האוטומטי מחייג למספר הטלפון המבוקש, אם ההסתברות לקבלת צליל "פנוי" בשידור אחד של אות חיוג הוא P .
- מתוך 12 ניסיונות חיוג אוטומטי למשרד הרישוי בזמנים שונים במשך 5 ימים, התקבלו התוצאות הבאות: בשני ניסיונות הופסק החיוג האוטומטי ובשאר הניסיונות שבהם הצליח המטלפן להשיג את המספר המבוקש, מספר החיוגים האוטומטיים עד לקבל צליל "פנוי" היו: 1, 6, 2, 7, 3, 8, 2, 1, 5. מצאו אומדן נראות מקסימלית עבור P , על סמך התוצאות שהתקבלו.

תשובות סופיות:

להלן גרף:



$$(1) \quad L(p) = (1-p) \cdot p \quad \text{א.}$$

$$\text{ב. } 0.5 \quad \text{ג. } \frac{2}{9}$$

$$(2) \quad X \quad \text{א.} \quad \bar{X} \quad \text{ב.}$$

$$(3) \quad \frac{1}{\bar{X}} \quad \text{א.} \quad \frac{2}{9} \quad \text{ב.}$$

$$(4) \quad \text{א. } 1 \quad \text{ב. } 1 \quad \text{ג. } 3 \quad \text{ד. } \hat{\theta} = \max\{X_1, \dots, X_n\}$$

$$(5) \quad \text{א. } \frac{\sum_{i=1}^n (x_i - 170)^2}{n} \quad \text{ב. } 40.2$$

$$(6) \quad \frac{x}{n}$$

$$(7) \quad \text{א. } \sum X_i^2 \quad \text{ב. } \left(\frac{n}{\sum X_i^2} \right)^2$$

$$(8) \quad \text{א. ראה סרטון.} \quad \text{ב. כד א'.$$

$$(9) \quad \text{א. } 32 \quad \text{ב. } 0.3916$$

$$(10) \quad \text{א. } 0.45 \quad \text{ב. } 0.81$$

$$(11) \quad \text{א. ראה סרטון.} \quad \text{ב. הוגן.$$

$$(12) \quad \text{א. } 0.08 \quad \text{ב. } 0.92 \quad \text{ג. } 11.5$$

$$(13) \quad \hat{\theta} = \begin{cases} 3 & X = 0,1 \\ 1 & X = 2 \end{cases} \quad \text{א.} \quad \hat{p} = \begin{cases} 2 \cdot 0.4 \cdot 0.6 & X = 0,1 \\ 2 \cdot 0.9 \cdot 0.1 & X = 2 \end{cases} \quad \text{ג.}$$

$$(14) \quad \max |X_i|$$

$$(15) \quad \text{שאלת הוכחה.}$$

$$(16) \quad \text{א. } 0.6124 \quad \text{ב. } \hat{p} = \sqrt{\frac{y}{n}} \quad \text{ג. } 0.49$$

$$(17) \quad \text{א.} \quad P(x) = \begin{cases} (1-p)^{x-1} p & 1 \leq x \leq 9 \\ (1-p)^9 & x = 10 \end{cases} \quad \text{ב. } 0.1818$$

נספח:
התפלגויות רציפות

ההתפלגות	פונקציית הצפיפות	פונקציית ההתפלגות המצטברת	תוחלת	שונות	הערות	אנ"מ
$X \sim U(a, b)$	$f(x) = \frac{1}{b-a}$ $a \leq x \leq b$	$\frac{t-a}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$		$b = \max(X_i)$ $a = \min(X_i)$
$X \sim \exp(\lambda)$	$f(x) = \lambda e^{-\lambda x}$	$1 - e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	הזמן עד להתרחשות מאורע מסוים. λ - הוא ממוצע האירועים ביחידת זמן.	$\hat{\lambda} = \frac{1}{\bar{X}}$
$X \sim N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\Phi(t)$	μ	σ^2		$\bar{\mu} = \bar{X}$ $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ $Z = \frac{X - \mu}{\sigma}$

התפלגויות בדידות

ההתפלגות	פונקציית ההסתברות $P(X = k)$	תוחלת	שונות	הערות	אנ"מ
בינומית $B(n, p)$ $0 \leq p \leq 1$	$\binom{n}{k} p^k (1-p)^{n-k}$ $k = 0, 1, \dots, n$	np	$np(1-p)$	(1)	$\hat{p} = \frac{Y}{n}$
גיאומטרית $G(p)$ $0 < p \leq 1$	$(1-p)^{k-1} p$ $k = 1, 2, \dots, \infty$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	(2)	$\hat{p} = \frac{1}{\bar{X}}$
אחידה $U(a, b)$	$\frac{1}{b-a+1}$ $K = a, \dots, b$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2 - 1}{12}$	(3)	$b = \max(X_i)$ $a = \min(X_i)$
פואסונית $P(\lambda)$ $\lambda > 0$	$\frac{e^{-\lambda} \lambda^k}{k!}$ $k = 0, 1, \dots, \infty$	λ	λ	(4)	$\hat{\lambda} = \bar{X}$

(1) מספר ההצלחות ב- n ניסויי ברנולי ב"ת. p - ההסתברות להצלחה.

(2) מספר הניסויים עד להצלחה הראשונה בסדרת ניסויי ברנולי ב"ת, p - ההסתברות להצלחה.

(3) בחירה אקראית של מספר בין a ו- b .

(4) מספר אירועים ביחידת זמן, λ - קצב האירועים.

מבוא להסתברות וסטטיסטיקה

פרק 44 - רווח סמך לתוחלת (ממוצע)

תוכן העניינים

1. רווח סמך כששונות האוכלוסיה ידועה 210
2. קביעת גודל מדגם 216
3. רווח סמך כששונות האוכלוסיה לא ידועה 218

רווח סמך כששונות האוכלוסייה ידועה:

רקע:

ממוצע המדגם הוא אומדן לממוצע האוכלוסייה, אך לא באמת ניתן להבין ממנו על גודלו של ממוצע האוכלוסייה. ההסתברות שממוצע המדגם יהיה בדיוק כמו הממוצע האמיתי הוא אפסי.

מה שנהוג לעשות כדי לאמוד את ממוצע האוכלוסייה, זה לבנות רווח סמך.

נבנה מרווח בטחון שהסיכוי שהפרמטר μ ייכלל בתוכו הוא: $1-\alpha$.

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

A - גבול התחתון של רווח הסמך.

B - הגבול העליון של רווח הסמך.

$L = B - A$ - אורך רווח הסמך.

דוגמה (פתרון בהקלטה):

חוקר דגם 25 חיילים שנבחנו במבחן הפסיכומטרי. הוא בנה רווח סמך לממוצע הציונים במבחן הפסיכומטרי בקרב אוכלוסיית החיילים וקיבל בין 510 ל-590. רווח הסמך נבנה ברמת סמך של 95%.

1. מהי אוכלוסיית המחקר?
2. מה המשתנה באוכלוסייה?
3. מה הפרמטר שהחוקר רצה לאמוד?
4. מהו רווח הסמך?
5. מה אורך רווח הסמך?
6. מהי רמת הביטחון של רווח הסמך?

בפרק זה נרצה לבנות רווח סמך לתוחלת (μ) במקרה ש- σ^2 (שונות האוכלוסייה) ידועה.

פרמטר אותו נרצה לאמוד: μ .

אומד נקודתי: $\bar{x}$.

תנאים לבניית רווח הסמך: $X \sim N$ או $n \geq 30$.

σ^2 (שונות האוכלוסייה) ידועה.

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

דוגמה (פתרון בהקלטה):

על פי נתוני היצרן אורך חיי סוללה מתפלג נורמאלית עם סטיית תקן של 1 שעה. מעוניינים לאמוד את תוחלת חיי סוללה. נדגמו באקראי 4 סוללות, אורך החיים הממוצע שהתקבל הוא 13.5 שעות. בנו רווח סמך ברמת סמך של 95% לתוחלת אורך חיי סוללה.

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

ε - נותן את שגיאת האמידה המקסימלית, דבר שנקרא גם טעות סטטיסטית, טעות דגימה.

דוגמה (פתרון בהקלטה):

בהמשך לשאלה עם הסוללות. מה ניתן להגיד בביטחון של 95% על שגיאת האמידה?

קשרים מתמטיים ברווח הסמך:

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

שאלות:

- (1) חוקר התעניין לאמוד את השכר הממוצע במשק. על סמך מדגם הוא קבע שבביטחון של 95% כי השכר הממוצע במשק נע בין 9200 ל-9800 ₪.
- מי האוכלוסייה במחקר?
 - מה המשתנה הנחקר?
 - מה הפרמטר שאותו רוצים לאמוד?
 - מה רווח הסמך לפרמטר?
 - מהי רמת הסמך לפרמטר?
 - מה אורך רווח הסמך?
 - מה הסיכוי שטעות הדגימה תעלה על 300 ₪?
- (2) מעוניינים לאמוד את התפוקה היומית הממוצעת של מפעל מסוים ברמת סמך של 95%. במדגם אקראי של 100 ימים התקבלה תפוקה ממוצעת 4950 מוצרים ביום. לצורך פתרון הנח שסטיית התקן האמתית ידועה ושווה 150 מוצרים ביום. בנו את רווח הסמך.
- (3) מעוניינים לאמוד את ממוצע אורך החיים של מכשיר. מנתוני היצרן ידוע שאורך החיים מתפלג נורמאלי עם סטיית תקן של 20 שעות. נדגמו 25 מכשירים ונמצא כי ממוצע אורך החיים שלהם היה 230 שעות.
- בנו רווח סמך ברמת סמך של 90% לאורך החיים הממוצע של מכשיר.
 - בנו רווח סמך ברמת סמך של 95% לאורך החיים הממוצע של מכשיר.
 - הסבירו כיצד ומדוע השתנה רווח הסמך.
- (4) דגמו 200 עובדים מהמשק הישראלי. השכר הממוצע שלהם היה 9700 ₪. נניח שסטיית התקן של השכר במשק היא 3000 ₪.
- בנו רווח סמך ברמת סמך של 95% לתוחלת השכר במשק.
 - מה ניתן לומר בביטחון של 95% על הסטייה המרבית בין ממוצע המדגם לתוחלת השכר?
 - מה היה צריך להיות גודל המדגם אם הינו רוצים להקטין את רווח הסמך ב-50%?
 - אם היינו מגדילים את גודל המדגם ובונים רווח סמך באותה רמת סמך האם היה ניתן לטעון בביטחון רב יותר שרווח הסמך מכיל את הפרמטר?

- (5) בנו רווח סמך לממוצע הציונים של מבחן אינטליגנציה. ידוע שסטיית התקן היא 15 והמדגם מתבסס על 100 תצפיות. רווח הסמך שהתקבל הוא (105,99). שחזרו את:
- א. ממוצע המדגם.
 - ב. שגיאת האמידה המקסימאלית.
 - ג. רמת הסמך.
- (6) זמן החלמה מאנגינה מתפלג עם סטיית תקן של יומיים. חברת תרופות מעוניינת לחקור אנטיביוטיקה חדשה שהיא פיתחה. במחקר השתתפו 60 אנשים שחלו באנגינה וקיבלו את האנטיביוטיקה החדשה. בממוצע הם החלימו לאחר 4 ימים.
- א. בנו רווח סמך לתוחלת זמן ההחלמה תחת האנטיביוטיקה החדשה ברמת סמך של 90%.
 - ב. מה היה קורה לאורך רווח הסמך אם היה תקציב להגדלת גודל המדגם פי 4? הסבירו.
 - ג. מה היה קורה לאורך רווח הסמך אם היינו בונים את רווח הסמך ברמת סמך גדולה יותר? הסבירו.
- (7) חוקר בנה רווח סמך לממוצע וקיבל את רווח הסמך הבא: $82 < \mu < 92$. נתון שסטיית התקן בהתפלגות שווה ל-10 ושהמדגם מתבסס על 16 תצפיות. התפלגות המשתנה היא נורמאלית.
- א. מהו ממוצע המדגם?
 - ב. מהי רמת הסמך של רווח הסמך שנבנה?
 - ג. מה הסיכוי ששגיאת האמידה באמידת ממוצע האוכלוסייה תעלה על 5?
- (8) חוקר בנה רווח סמך לתוחלת כאשר השונות בהתפלגות ידועה ברמת סמך של 95%. אם החוקר כעת יבנה על סמך אותם נתונים רווח סמך ברמת סמך קטנה מ-95%, איזה מהמשפטים הבאים לא יהיה נכון.
- א. אורך רווח הסמך החדש יהיה קטן יותר.
 - ב. גודל המדגם יהיה כעת קטן יותר.
 - ג. המרחק בין ממוצע המדגם לקצות רווח הסמך יהיו קטנים יותר ברווח הסמך החדש.
 - ד. רמת הביטחון לבנות רווח הסמך החדש תהיה קטנה יותר.

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

א. $\mu = 51$.

ב. $\bar{X} = 6$.

ג. $\bar{X} = 51$.

ד. אורך רווח הסמך הינו 3.

(10) איזה מהגורמים הבאים אינו משפיע על גודלו של רווח בר סמך, כאשר שונות האוכלוסייה ידועה (בחרו בתשובה הנכונה):

א. רמת הביטחון.

ב. סטיית התקן באוכלוסייה.

ג. מספר המשתתפים.

ד. סטיית התקן במדגם.

(11) חוקר בנה רווח סמך לממוצע וקיבל את רווח הסמך הבא: $63 < \mu < 83$. נתון שסטיית התקן בהתפלגות הייתה ידועה לו ושהמדגם התבסס על 40 תצפיות.

א. אם החוקר היה רוצה לבנות רווח סמך באורך 10.

כמה תצפיות עליו היה לדגום?

ב. רווח הסמך שנבנה על ידי החוקר היה ברמת סמך של 95%.

בנו את רווח הסמך שהיה מתקבל ברמת סמך של 98%.

(12) נתון משתנה מקרי רציף מתפלג אחיד: $X_i \sim U(\mu - 0.5, \mu + 0.5)$. נרצה לאמוד את μ . מצאו רווח סמך ל- μ ברמת-ביטחון של 0.95 אם במדגם של 45 תצפיות התקבל: $\bar{x} = 74$.

(תזכורת על השונות בהתפלגות אחידה רציפה: $(Var(X_i) = \frac{(b-a)^2}{12})$.

תשובות סופיות:

- (1) א. העובדים במשק. ב. שכר ב-ש. ג. μ . ד. $9200 < \mu < 9800$.
ה. 0.95. ו. 600. ז. 0.05.
- (2) $4920.6 < \mu < 4979.4$
- (3) א. $223.42 < \mu < 236.58$. ב. $222.16 < \mu < 237.84$.
ג. ראה סרטון.
- (4) א. $9284 < \mu < 10,116$. ב. הסטיה המירבית בין $\bar{x}$ ל- μ היא 416 שם בביטחון של 95%.
ג. 800. ד. לא.
- (5) א. 102. ב. 3. ג. 0.9544.
- (6) א. $4.42 < \mu < 83.5$. ב. יקטן פי 2. ג. גדל.
- (7) א. 87. ב. 5. ג. 0.9544.
- (8) ב'.
- (9) ג'.
- (10) ד'.
- (11) א. 160. ב. $61.13 < \mu < 84.87$.
- (12) 74 ± 0.084

קביעת גודל מדגם:

רקע:

אם מעוניינים לאמוד את ממוצע האוכלוסייה כאשר סטיית התקן של האוכלוסייה ידועה: σ ברמת סמך של $1-\alpha$ ושגיאת אמידה שלא תעלה על ε מסוים, נציב

$$.n \geq \left(\frac{z_{1-\frac{\alpha}{2}} \cdot \sigma}{\varepsilon} \right)^2 \quad \text{בנוסחה הבאה:}$$

כדי להציב בנוסחה צריך שהמשתנה הנחקר יתפלג נורמלית או שהמדגם ייצא בגודל של לפחות 30 תצפיות.

דוגמה (פתרון בהקלטה):

חברת תעופה מעוניינת לאמוד את תוחלת משקל המטען של נוסע. נניח שמשקל מטען של נוסע מתפלג נורמאלית עם סטיית תקן של 2 ק"ג. כמה נוסעים יש לדגום אם מעוניינים שבביטחון של 98% הסטייה המרבית בין ממוצע המדגם לממוצע האמיתי לא יעלה על 0.5 ק"ג? (תשובה: 87).

שאלות:

- (1) משתנה מקרי מתפלג נורמאלית עם סטיית תקן ידועה 12. מה צריך להיות גודל המדגם כדי לבנות רווח סמך ברמת סמך של 98% שאורכו לא יעלה על 2?
- (2) מעוניינים לאמוד את הדופק הממוצע של מתגייסים לצבא. מעוניינים שבביטחון של 95% שגיאת האמידה המרבית תהיה 0.5. נניח שהדופק מתפלג נורמאלית על סטיית תקן של 3 פעימות לדקה.
- א. כמה מתגייסים יש לדגום?
- ב. אם ניקח מדגם הגדול פי 4 מהמדגם של סעיף א ונאמוד את הממוצע באותה רמת סמך כיצד הדבר ישפיע על שגיאת האמידה?
- (3) יהי X משתנה מקרי עם ממוצע μ וסטיית תקן σ . חוקר רוצה לבנות רווח בר סמך ל- μ ברמת ביטחון של 0.95, כך שהאורך של הרווח יהיה 0.5σ . מהו גודל המדגם הנדרש?

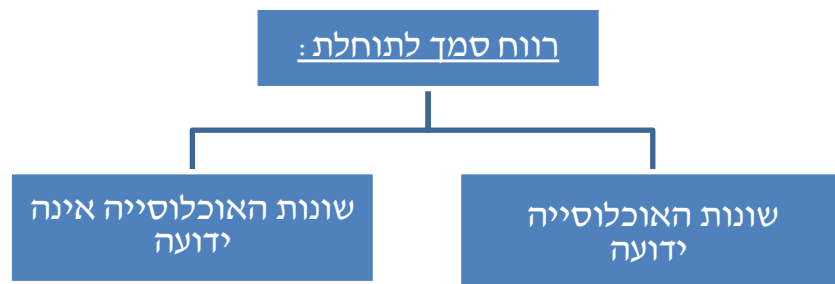
תשובות סופיות:

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

רווח סמך כששונות האוכלוסייה לא ידועה:

רקע:

בבואנו לבנות רווח סמך לתוחלת אנו צריכים להתמקד בשני המצבים הבאים:

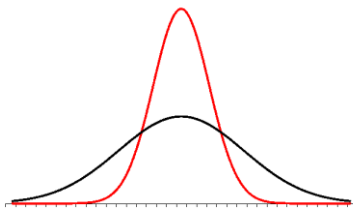


בפרק זה נעסוק במקרה ששונות האוכלוסייה (σ^2) אינה ידועה לנו.

מקרה יותר פרקטי.

התנאי: $X \sim N$ או שהמדגם גדול.

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



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

התפלגות T:

הינה התפלגות סימטרית פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה

להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים.

התפלגות T תלויה במושג שנקרא דרגות חופש. דרגות החופש הן: $df = n-1$.

ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה.

כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה (פתרון בהקלטה):

הזמן שלוקח לפתור שאלה מסוימת בחשבון מתפלג אצל תלמידי כיתות ח' נורמאלית.

במטרה לאמוד את תוחלת זמן הפתרון נדגמו 4 תלמידים בכיתה ח'. להלן התוצאות

שהתקבלו בדקות: 4.7, 5.2, 4.6, 5.3.

בנו רווח סמך ברמת סמך של 95% לממוצע זמן הפתרון לשאלה בקרב תלמידי כיתה ח'.

שאלות:

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

- (6) נדגמו 120 אנשים אקראיים מעל גיל 50. עבור כל אדם נבדק מספר שנות השכלתו. להלן התוצאות שהתקבלו: $\bar{x} = 13.8$, $S = 2$. בנו רווח סמך ברמת סמך של 96% לממוצע ההשכלה של אזרחים מעל גיל 50.
- (7) שני סטטיסטיקאים בנו רווח בר-סמך לאותו פרמטר μ . לכל אחד מהסטטיסטיקאים מדגם אחר, אך באותו גודל 10. שניהם קבעו אותה רמת סמך. סטטיסטיקאי א': הניח $\sigma = 20$. סטטיסטיקאי ב': חישב לפי המדגם וקיבל $S = 20$. למי משני הסטטיסטיקאים יהיה רווח סמך ארוך יותר?
- א. סטטיסטיקאי א.
 ב. סטטיסטיקאי ב.
 ג. אותו אורך רווח סמך לשני הסטטיסטיקאים.
 ד. תלוי בתוצאות המדגם של כל סטטיסטיקאי.
- (8) נתון ש: $X \sim N(\mu, \sigma^2)$ ביצעו מדגם בגודל 16 וקיבלו סטיית תקן מדגמית 10. אורך רווח הסמך שהתקבל הוא: 8.765. מהי רמת הביטחון של רווח הסמך?

תשובות סופיות:

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

מבוא להסתברות וסטטיסטיקה

פרק 45 - בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן)

תוכן העניינים

1. בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן) 221

בדיקת השערות כללית (סיכוי לטעויות ועוצמת מבחן):

רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטיקה. בתהליך זה ישנן שתי השערות שנבדקות:

1. השערת האפס: המסומנות ב- H_0 .

2. השערה אלטרנטיבית (השערת המחקר): המסומנת ב- H_1 .

בדרך כלל השערת האפס מסמנת את אשר היה מקובל עד עכשיו, את השגרה הנורמה ואילו ההשערה האלטרנטיבית את החדשנות בעצם ההשערה האלטרנטיבית מדברת על הסיבה שהמחקר נעשה.

דוגמה:

ישנה תרופה קיימת למחלה A אשר גורמת ל-10% מהמשתמשים בה לתופעות לוואי. חברת תרופות טוענת שפיתחה תרופה שיעילה באותה מידה, אך מקטינה את הסיכוי לתופעות הלוואי. לכן יש לבצע מחקר שעל סמך תוצאותיו ננסה להכריע איזה השערה נקבל:

H_0 : התרופה החדשה הנה קונבנציונאלית וגורמת ל-10% תופעות לוואי.

H_1 : התרופה החדשה מקטינה את אחוז הסובלים מתופעות לוואי מתחת ל-10%.

בתהליך של בדיקת השערות יוצרים כלל שנקרא כלל הכרעה. הכלל יוצר אזורים:

1. אזור דחייה: דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.

2. אזור קבלה: קבלה של השערת האפס ודחייה של האלטרנטיבה.

כלל ההכרעה מתבסס על איזשהו סטטיסטי. בתהליך יש ללכת לתוצאות המדגם ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה וכך להגיע למסקנה. המסקנה היא בעירבון מוגבל כיוון שהיא תלויה בכלל ההכרעה ובתוצאות המדגם. אם נשנה את כלל ההכרעה אז אנחנו יכולים לקבל מסקנה אחרת, אם נבצע מדגם חדש אז אנחנו עלולים לקבל תוצאה אחרת.

לכן יתכנו טעויות במסקנות שלנו :

מציאות	הכרעה		
		H_0	H_1
	H_0	אין טעות	טעות מסוג 1
	H_1	טעות מסוג 2	אין טעות

הגדרת הטעויות:

טעות מסוג ראשון: להכריע לדחות את H_0 למרות שבמציאות H_0 נכונה.

טעות מסוג שני: להכריע לקבל את H_0 למרות שבמציאות H_1 נכונה.

הגדרת הסתברויות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):

$$\alpha = P(H_0 \text{ לדחות את } H_0 \mid H_0 \text{ נכונה}) = P_{H_0}(H_0 \text{ לדחות})$$

הסיכוי לבצע טעות מסוג 2:

$$\beta = P(H_0 \text{ לקבל את } H_1 \mid H_1 \text{ נכונה}) = P_{H_1}(H_0 \text{ לקבל})$$

רמת בטחון:

$$1 - \alpha = P(H_0 \text{ לקבל את } H_0 \mid H_0 \text{ נכונה}) = P_{H_0}(H_0 \text{ לקבל})$$

עוצמה:

$$1 - \beta = \pi = P(H_1 \text{ לדחות את } H_0 \mid H_1 \text{ נכונה}) = P_{H_1}(H_0 \text{ לדחות})$$

דוגמה (פתרון בהקלטה):

בכד יש 10 כדורים. יתכן ש-5 מהם לבנים והיתר שחורים (כד א' – השערת האפס) או ש-7 מהם לבנים והיתר שחורים (כד ב' – השערה אלטרנטיבית). כדי להחליט איזה מהכדים ברשותנו, הוחלט להוציא כדור ולהשתמש בכלל ההחלטה הבא: אם הכדור שהוצא הוא לבן שזהו כד ב' H_1 .

א. חשבו את רמת המובהקות ואת רמת הביטחון של המבחן המוצע.

ב. חשבו את הסיכוי לטעות מסוג שני והעוצמה של המבחן המוצע.

שאלות:

- (1) אדם חשוד בביצוע פשע. מהן הטעויות האפשריות בהכרעת הדין?
- (2) ילד קנה שקית סוכריות אטומה שבה ציפה ל-10 סוכריות תות ו-5 לימון. ישנה שקית אחרת אותה הוא לא רצה בה 6 סוכריות תות ו-9 לימון. הוא החליט להוציא באקראי סוכרייה, אם היא תהיה לימון הוא יחזיר את השקית לחנות. מה הסיכויים לכל סוג של טעות בהכרעתו?
- (3) יהי X מספר שלם הנבחר באקראי מבין המספרים השלמים. הסיכוי ש- X יקבל ערך כלשהו נתון על ידי הנוסחה: $p(X=k) = \frac{1}{n}$ עבור: $k=1,2,\dots,n$. נתונות ההשערות הבאות לגבי התפלגות של X : $H_0: n=4$, $H_1: n=6$. כמו כן נתון כלל ההכרעה הבא: נדחה את השערת האפס אם: $X > 3$. חשבו את הסיכוי לטעות מסוג ראשון וטעות מסוג שני ואת העוצמה?
- (4) איכות של מוצר מסווגת ל-4 רמות איכות: מצוין, טוב, בינוני וירוד. להלן התפלגות טיב המוצר בשני מפעלים:
- | מפעל / איכות | מצוין | טוב | בינוני | ירוד |
|--------------|-------|-----|--------|------|
| "היוצר" | 0.6 | 0.2 | 0.2 | 0 |
| "שמשון" | 0.1 | 0.2 | 0.3 | 0.4 |
- בוחרים ממשלוח מוצר באקראי, אך לא יודעים מאיזה מפעל המשלוח הגיע. על סמך בדיקת האיכות מנסים להכריע האם מדובר במפעל "היוצר" (השערת האפס) או במפעל "שמשון" (השערה אלטרנטיבית).
- א. להלן כלל החלטה: אם מדובר במוצר שטיבו "טוב" נכריע שהמוצר בא ממפעל "שמשון", מהן ההסתברויות לסוגי הטעויות השונים?
- ב. להלן כלל החלטה: אם מדובר במוצר שטיבו "בינוני" או גרוע מכך נכריע שהמוצר בא ממפעל "שמשון", מהן ההסתברויות לסוגי הטעויות השונים?
- ג. איזה כלל החלטה עדיף? נמקו!
- (5) במטרה לבדוק האם מטבע תקין הטילו אותו 8 פעמים. הוחלט שאם מספר העצים יהיה בין 1 ל-7 כולל יוחלט שהמטבע תקין, אחרת נחליט שהמטבע מזויף.
- א. רשמו את השערות המחקר.
- ב. מה ההסתברות לטעות מסוג ראשון?
- ג. מהי עצמת המבחן אם במציאות אכן המטבע אינו תקין כי הסיכוי לעץ בו הוא 20%.

(6) להלן השערות:

$H_0: X \sim t(5)$ - התפלגות T עם חמש דרגות חופש.

$H_1: X \sim Z$ - התפלגות נורמלית.

כלל החלטה: נדחה את השערת האפס אם X גדול מ-2.015.

א. מהי רמת המובהקות של כלל ההחלטה?

ב. מהי העוצמה של כלל ההחלטה?

(7)

במפעל מסוים נפלטים לאוויר חומרים רעילים. במצב שיגרה העוצמה הממוצעת של החומר הרעיל אמורה להיות 6,000 יחידות עם סטיית תקן 900. במצב חירום העוצמה הממוצעת היא 7,000 עם סטיית תקן 900. במפעל מערכת התראה נתמכת על ידי 9 חיישנים. אם ממוצע העוצמה של החומר הרעיל לפי תשעת החיישנים עולה על 6,600 יחידות מופעלת מערכת ההתראה. נתון שעוצמת הזיהום מתפלגת נורמלית.

א. מה הסיכוי להתראת שווא? (באיזה סוג טעות מדובר)?

ב. מה הסיכוי שבמצב חירום מערכת ההתראה לא תפעל? (באיזה סוג טעות מדובר)?

ג. מה ההסתברות שאם המצב הוא מצב חירום מערכת ההתראה תפעל? (איך קוראים להסתברות זו)?

ד. בסעיפים הבאים נשנה בכל סעיף נתון מסוים. כל סעיף עומד בפני עצמו, כיצד השינוי ישנה את הסיכוי לטעות מסוג ראשון ושני?

i. המפעל יקנה עוד 4 חיישנים.

ii. מצב חירום מוגדר כעת בתוחלת של 7,500 יחידות.

iii. מערכת ההתראה תופעל אם ממוצע של תשעת החיישנים יהיה מעל 6,700.

(8)

במטרה לבדוק האם במקום עבודה מסוים פרופורציית הבנים נמוכה מפרופורציית הבנות נדגמו באקראי 10 עובדים. הוחלט שאם מספר הבנים במדגם יהיה לכל היותר 2 תתקבל הטענה שפרופורציית הבנים נמוכה מפרופורציית הבנות.

א. מה רמת המובהקות של כלל ההכרעה הנ"ל?

ב. מהי העוצמה בהנחה ובחברה 30% בנים?

(9) זמן ההשפעה של משכך הכאבים "אופטלנוס" מתפלג נורמאלי עם תוחלת של 40 דקות וסטיית תקן של 12 דקות. חברת התרופות המייצרת את התרופה מנסה לשפר את התרופה כך שתוחלת הזמן עד להשפעה תתקצר. לצורך כך, דגמו 25 מטופלים שיקבלו את התרופה "אופטלנוס פורטה", ממוצע זמן התגובה של המטופלים היה 34.5 דקות. חברת התרופות החליטה מראש שאם ממוצע הזמן עד להשפעה יהיה נמוך מ-35 דקות, היא תמשיך בתהליך שיווק "אופטלנוס פורטה".

- א. מהי רמת המובהקות של המבחן המוצע?
- ב. על סמך תוצאות המדגם, מהי המסקנה ומהי הטעות האפשרית במסקנה?
- ג. מהי עצמת המבחן המוצע אם במציאות התרופה "אופטלנוס פורטה" מפחיתה את התוחלת לכדי 32 דקות?
- ד. כיצד תשתנה התשובה לסעיף ג' אם החברה הייתה מחליטה שהיא תמשיך בתהליך שיווק התרופה החדשה כאשר ממוצע המדגם יהיה נמוך מ-36 דקות?

(10) ציוני פסיכומטרי מתפלגים נורמלית עם סטיית תקן 120. מכון טוען שלימודים אצלו מעלים את ממוצע הציונים ביותר מ-30 נקודות. נלקחו 20 שלמדו במכון ו-20 שניגשו לבחינה בלמידה עצמית. הוחלט במשרד פרסום לקבל את טענת המכון רק אם במדגם ממוצע הציונים של אלה שלמדו במכון יהיה גבוהה בלפחות 50 נקודות מאלה שלא היו.

- א. מהי רמת המובהקות של המחקר?
- ב. מה הסיכוי לעשות טעות מסוג שני II בהנחה שהמכון מעלה את ממוצע הציונים ב-60 נקודות?
- ג. כיצד התשובות לסעיף א' ו' ב' היו משתנות אם מסתבר שסטיית התקן בציוני הפסיכומטרי הינה 100. הסבירו ללא חישוב.

(11) קו ייצור נחשב תקין אם יש בו לכל היותר 4% פגומים, ונחשב שאינו תקין אחרת. מנהל האיכות דוגם בכל יום מקו הייצור 500 מוצרים. אם במדגם יהיה לפחות 30 מוצרים פגומים יפסיקו באותו היום את קו הייצור.

- א. מה ההסתברות להפסיק את קו הייצור כשהוא תקין. איך קוראים להסתברות זאת?
- ב. מה ההסתברות להמשיך ביום מסוים את קו הייצור למרות שאינו תקין כי היו 8% פגומים בקו הייצור. איך קוראים להסתברות זאת?

(12) מעוניינים לבדוק האם בפקולטה מסוימת ישנה העדפה לגברים. הוחלט לדגום 200 מתקבלים ועל סמך מספר הבנים לקבוע אם טענת המחקר מתקבלת. חוקר א' קבע רמת מובהקות של 5% וחוקר ב' החליט לקבל את טענת המחקר אם במדגם יהיו לפחות 120 בנים. למי מבין החוקרים רמת מובהקות גדולה יותר?

13 מספר המכוניות הנכנסות לחניון "עזרים" מתפלג פואסונית. בשנה שעברה המכוניות נכנסו לחניון בקצב של 2 מכוניות לדקה. בעקבות תלונות על עומס יתר בכניסה לחניון מעוניין מנהל החניון לבדוק האם קצב כניסת המכוניות לחניון גדל השנה. מנהל החניון החליט לספור את מספר המכוניות שיכנסו לחניון בדקה אקראית. אם מספר המכוניות שיספרו יהיה לפחות 4 יפתח מנהל החניון שער נוסף לחניון.

א. רשמו את השערות מנהל החניון ואת כלל ההחלטה שלו. האם כלל ההכרעה הגיוני?

ב. מהי רמת המובהקות של כלל ההכרעה?

ג. מהי העוצמה של כלל ההחלטה, אם כיום קצב כניסת המכוניות לחניון גדל ל-4 מכוניות בדקה?

14 עודד עובד במפעל שבו מתחילים לעבוד בשעה 8:00. עודד בדרך כלל מאחר לעבודה והמנהל החליט לרשום את שעת הגעתו. המנהל טוען שמשך האיחור של עודד (בדקות), X , הוא משתנה אחיד $U(0, 60)$. עודד טוען שהוא לא מגיע באיחור כה גדול, אלא שהתפלגות X היא בעלת התפלגות מעריכית עם תוחלת איחור של 20 דקות.

לבדיקת טענת המנהל (H_0) כנגד טענת עודד (H_1), המבוסס על משך האיחור של חגי ביום אחד. מוצאים שני ככלי הכרעה:

כלל 1: דחה את השערת האפס אם משך האיחור יהיה לפחות 40 דקות.

כלל 2: דחה את השערת האפס אם משך האיחור יהיה לכל היותר 20 דקות.

חשבו את הסיכוי לטעות מסוג ראשון ושני לכל אחת מכללי ההכרעה. מי עדיף?

תשובות סופיות:

- (1) ראה סרטון וידאו.
- (2) $\beta = \frac{2}{5}$, $\alpha = \frac{1}{3}$
- (3) $\beta = 0.5$, $\alpha = 0.25$
- (4) א. $\beta = 0.8$, $\alpha = 0.2$ ב. $\beta = 0.3$, $\alpha = 0.2$ ג. כלל ב'.
- (5) א. השערות: H_0 - מטבע תקין. H_1 - מטבע לא תקין. ב. 0.00781250 ג. 0.1678
- (6) א. 0.05 ב. 0.022
- (7) א. 0.0228 ב. 0.0918 ג. 0.9082 ד. i. α, β יקטנו. ii. α לא משתנה, β קטנה.
- (8) א. 0.055 ב. 0.383
- (9) א. 0.0188 ב. טעות מסוג I ג. 0.8944 ד. העוצמה תגדל.
- (10) א. 0.2981 ב. 0.3974 ג. קטן.
- (11) א. 0.0113 ב. 0.0495
- (12) חוקר א'.
- (13) א. ראה סרטון וידאו. ב. 0.1428 ג. 0.566
- (14) להלן טבלת טעויות, ממנה ניתן להסיק שכלל 2 עדיף.

β	α	כלל
0.865	$\frac{1}{3}$	1
0.368	$\frac{1}{3}$	2

מבוא להסתברות וסטטיסטיקה

פרק 46 - הלמה של ניימן פירסון

תוכן העניינים

1. כללי 228

הלמה של ניימן פירסון:

רקע:

שיטה זו עוזרת לנו לבנות מבחנים בעלי עוצמה מקסימאלית עבור α נתונה. הלמה של ניימן פירסון אומרת שמבחן בעל עוצמה מקסימלית מתקבל כאשר אזור הדחיה שלו כולל את התוצאות שעבורן יחס הנראות הוא הגבוה ביותר.

נגדיר את יחס הנראות: $x_1, x_2, \dots, x_n$ - תוצאות הניסוי.

ההשערות: $H_0: x_i \sim p_0$, $H_1: x_i \sim p_1$.

$$\lambda(x_1, x_2, \dots, x_n) = \frac{p_{H1}(x_1, x_2, \dots, x_n)}{p_{H0}(x_1, x_2, \dots, x_n)}$$

המשמעות של יחס הנראות היא פי כמה H_1 יותר סבירה מ- H_0 .

עבור משתנים שמתפלגים בדיד:

שלב א: עבור כל תוצאות המדגם האפשריים מחשבים את הסיכויים בהנחת השערת האפס ובהנחת ההשערה האלטרנטיבית.

שלב ב: מחלקים את הסיכויים באופן הבא: $\lambda(x_1, x_2, \dots, x_n) = \frac{p_{H1}(x_1, x_2, \dots, x_n)}{p_{H0}(x_1, x_2, \dots, x_n)}$.

ומקבלים את יחס הנראות.

שלב ג: מסדרים את תוצאות הניסוי על פי סדר יורד מהתוצאה שמניבה את ערך יחס הנראות הגבוה ביותר עד התוצאה שמניבה את ערך יחס הנראות הנמוך ביותר.

שלב ד: מכניסים את התוצאה שמניבה את יחס הנראות הגבוה ביותר לאזור הדחיה ובודקים תחת השערת האפס מהי רמת המובהקות המתקבלת. צוברים את התוצאות לפי העיקרון שהוצג עד שרמת המובהקות לא תעלה על ה- α הרצויה.

דוגמה (הפתרון בהקלטה):

איכות של מוצר מסווגת ל-4 רמות איכות: מצוין, טוב, בינוני וירוד. להלן התפלגות טיב המוצר בשני מפעלים:

מפעל / איכות	מצוין	טוב	בינוני	ירוד
"היוצר"	0.6	0.15	0.25	0
"שמשון"	0.1	0.2	0.3	0.4

בוחרים ממשלוח מוצר באקראי, אך לא יודעים מאיזה מפעל המשלוח הגיע. על סמך בדיקת האיכות מנסים להכריע האם מדובר במפעל "היוצר" (השערת האפס) או במפעל "שמשון" (השערה אלטרנטיבית). צרו כלל הכרעה לפי הלמה של ניימן פירסון ברמת מובהקות שלא תעלה על 20%.

עבור משתנים שמתפלגים רציף:

שלב א: בונים את פונקציית הצפיפות המשותפת בהנחת השערת האפס ובהנחת ההשערה האלטרנטיבית.

שלב ב: מחלקים את שלב א באופן הבא: $\lambda(x_1, x_2, \dots, x_n) = \frac{f_{H1}(x_1, x_2, \dots, x_n)}{f_{H0}(x_1, x_2, \dots, x_n)}$

ומקבלים את פונקציית יחס הנראות.

שלב ג: מזהים את האזור עבורו יחס הנראות הוא הגבוה ביותר.

שלב ד: לפי ההתפלגות של השערת האפס מוצאים את הערכים הקריטיים באזור שנקבע בסעיף הקודם כך שרמת המובהקות תהיה ה- α שנקבעה מראש.

דוגמה (פתרון בהקלטה):

נתון: $X \sim \exp(\lambda)$

ההשערות הן: $H_0: \lambda = 1$, $H_1: \lambda = 2$

מצאו מבחן בעל עוצמה מקסימלית ברמת מובהקות של 5% על סמך תצפית בודדת.

שאלות:

- (1) איכות של מוצר מסווגת ל-5 רמות איכות: מצוין, טוב, בינוני, ירוד ופסול. להלן התפלגות טיב המוצר בשני מפעלים:

מפעל / איכות	מצוין	טוב	בינוני	ירוד	פסול
"היוצר"	0.4	0.2	0.1	0.1	0.2
"שמשון"	0.1	0.2	0.3	0.4	0

בוחרים משלוח מוצר באקראי, אך לא יודעים מאיזה מפעל המשלוח הגיע. על סמך בדיקת האיכות מנסים להכריע האם מדובר במפעל "היוצר" (השערת האפס) או במפעל "שמשון" (השערה אלטרנטיבית).

- א. חשבו את יחס הנראות עבור כל תוצאות המדגם האפשריים.
 ב. צרו כלל הכרעה לפי הלמה של ניימן פירסון ברמת מובהקות שלא תעלה על 25%.

ג. מהי עוצמת המבחן שיצרת בסעיף הקודם?

- (2) מטבע הוטל 3 פעמים ומתבוננים במספר הפעמים שהתקבלה התוצאה ראש. נסמן ב- p את הסיכוי בהטלה בודדת לקבל את התוצאה ראש.

ההשערות הן: $H_0: p = 0.5$, $H_1: p = 0.25$.

מצאו מבחן ברמת מובהקות שלא תעלה על 30% עם עוצמה מקסימלית. מהי העוצמה?

- (3) בכד א' 7 כדורים לבנים ו-8 שחורים. בכד ב' 8 כדורים לבנים ו-7 שחורים. אדם בוחר כד וממנו מוציא באקראי 4 כדורים ללא החזרה. הוא מתבונן במספר הכדורים הלבנים שהוצאו ומודיע לך את המספר המתקבל. יש לבנות כלל הכרעה על סמך המספר המתקבל שיכריע האם מדובר בהוצאה מכד א' (השערת האפס) או מכד ב' (השערה אלטרנטיבית).
 א. בנו כלל הכרעה בעל עוצמה מקסימלית ברמת מובהקות שלא תעלה על 10%.
 ב. מהי רמת המובהקות של כלל ההכרעה שבנית בסעיף הקודם?
 ג. מה הסיכוי לטעות מסוג שני של כלל ההכרעה שבנית?

- (4) בצרור מפתחות 5 מפתחות שרק אחד פותח את הדלת. על סמך מספר הניסיונות לפתיחת הדלת יש להחליט האם הניסיונות נעשו ללא החזרה (השערת האפס) או עם החזרה (השערה אלטרנטיבית) של המפתחות לצרור.
 מצאו מבחן לפי הלמה של ניימן פירסון ברמת מובהקות שלא עולה על 0.25.

(5) מספר תאונות הדרכים בכביש 4 מתפלג פואסונית עם קצב של תאונה ביממה. לאחרונה התעורר החשד שתוחלת מספר התאונות בכביש עלתה לקצב של שתי תאונות ביממה. דגמו 4 ימים אקראיים וקיבלו את מספר התאונות הבאות ליממה: 3, 3, 0, 1.

- נסחו את הבעיה ובנו מבחן MP עם: $\alpha \leq 0.1$.
- מהי מסקנתך ומהי הטעות האפשרית במסקנה?
- מה הסיכוי להכריע שכיום קצב תאונות הדרכים בכביש מספר 4 עלה לשתי תאונות ביממה שאכן כך הדבר באמת? פתרו על סמך המבחן של סעיף א'.

(6) התפלגות זמן ההמתנה לקופה בסופרמרקט מתפלג מעריכית. בעל הסופרמרקט טוען שתוחלת זמן ההמתנה היא 5 דקות אך הלקוחות חושדים שהתוחלת גבוהה יותר ושווה ל-10 דקות.

- רשמו את השערות המחקר וחשבו את פונקציית יחס הנראות על סמך זמן המתנה של לקוח אקראי לקופה.
- מה כיוון אזור הדחייה של השערת האפס?
- מצאו את אזור הדחייה עבור רמת מובהקות של 5%.
- בהמשך לסעיף הקודם, מה הסיכוי להכריע לטובת בעל הסופרמרקט בטעות?

(7) יהי X תצפית בודדת מפונקציית הצפיפות הבאה, כאשר θ פרמטר חיובי:

$$f_{\theta}(x) = \begin{cases} 2\theta x + 1 - \theta & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad \frac{n!}{r!(n-r)!}$$

ההשערות הן: $H_0: \theta = 0$, $H_1: \theta = 1$.

- הוכיחו שמבחן MP (עוצמה מקסימלית) עם רמת מובהקות α יהיה: $C = \{X > 1 - \alpha\}$.

- הוכיחו שהעוצמה של המבחן שנמצאה היא: $2\alpha - \alpha^2$.
- מצאו את כלל ההכרעה והעוצמה עבור: $\alpha = 0.05$.

(8) מחשב חניון "אתרים" רושם את זמן כניסת כל מכונית לחניון. ישנו חשד שעקב תקלה המחשב מבצע את הרישום לכל מכונית שניה. נסמן ב- X את הזמן בדקות בין רישום לרישום.

אם הרישום הוא תקין, ההתפלגות היא מעריכית: $f_0(x) = \lambda \cdot e^{-\lambda x}$.

אם הרישום הוא לפי החשד, פונקציית הצפיפות היא: $f_1(x) = \lambda^2 \cdot x \cdot e^{-\lambda x}$.
 כאשר מדובר באותו פרמטר λ .

א. רשמו את ההשערות.

ב. מצאו מבחן בעל עוצמה מקסימלית כדי לבדוק את ההשערות.

ג. פתרו עבור רמת מובהקות של 5% ו- $\lambda = 1$.

ד. רשמו את איזו הדחייה עבור שתי תצפיות אקראיות של X .

(9) יהי $X_1, \dots, X_n$ מדגם מקרי מהתפלגות בעלת פונקציית הצפיפות הבאה, כאשר θ

$$f_\theta(x) = \begin{cases} \theta e^{1-\theta x} & x \geq \frac{1}{\theta} \\ 0 & \text{otherwise} \end{cases} \quad \text{פרמטר חיובי}$$

א. מצאו מבחן בעל עוצמה מקסימלית לבדיקת ההשערות: $H_0: \theta = 1$

כנגד: $H_1: \theta = 2$, על סמך תצפית בודדת (ברמת מובהקות 5%).

ב. מהי עוצמת המבחן שמצאת?

ג. כעת נשנה את הערך תחת האלטרנטיבה ל-3 במקום 2.

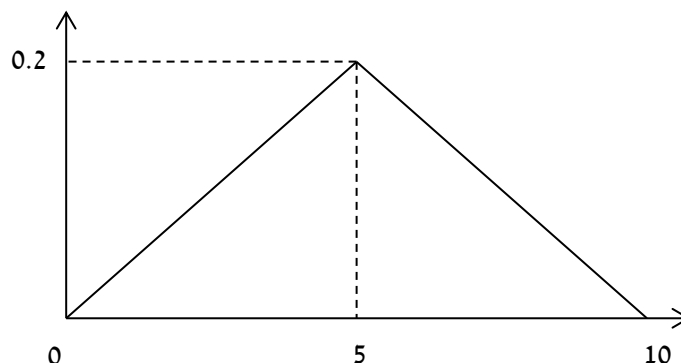
בחרו בתשובה הנכונה ונמקו:

i. אפשר לומר ללא חישוב נוסף שהעוצמה תגדל.

ii. אפשר לומר ללא חישוב נוסף שהעוצמה תקטן.

iii. יש לחשב כדי להחליט.

(10) נתונה פונקציית הצפיפות הבאה שנשמנה ב- $f_1(x)$:



$$f_0(x) = \begin{cases} 0.1 & 0 \leq x \leq 10 \\ 0 & \text{else} \end{cases} \quad \text{כמו כן נתון ש:}$$

- א. עבור השערות: $H_0: f = f_0$ כנגד: $H_1: f = f_1$, מצאו את צורת אזור הדחיה של מבחן בעל עוצמה מקסימלית, על סמך תצפית בודדת.
- ב. בהינתן: $\alpha = 0.05$, מצאו את כלל הכרעה מתאים בעל עוצמה מקסימלית.

(11) X הוא משתנה רציף המוגדר בין 0 ל-20.

$$\text{להלן השערות מחקר: } H_1: f(X) = \frac{e^{-\frac{x}{4}}}{4(1-e^{-5})} \quad , \quad H_0: X \sim U(0,20)$$

יש לבנות מבחן בעל עוצמה מקסימלית ברמת מובהקות של 10% על סמך הממוצע 100 תצפיות אקראיות.

תשובות סופיות:

(1) א. להלן טבלה: ב. ראה סרטון. ג. 0.7.

המפעל	ירוד	בינוני	טוב	מצוין	פסול
H_0	0.1	0.1	0.2	0.4	0.2
H_1	0.4	0.3	0.2	0.1	0
$\gamma(x)$	4	3	1	0.25	0

(2) $C = \{X = 0\}$, עוצמה $\frac{27}{64}$.

(3) א. $C = \{X = 4\}$. ב. 0.0256. ג. 0.9487.

(4) $C = \{X = 1 \text{ or } X \geq 6\}$.

(5) א. נדחה את השערת האפס אם מספר התאונות הכולל ב-4 הימים יהיה לפחות 8 תאונות. ב. נקבל את השערת האפס (טעות מסוג שני). ג. 0.5471.

(6) א. H_0 : התוחלת של זמן ההמתנה 5 דקות.

H_1 : התוחלת של זמן המתנה 10 דקות.

פונקציית יחס הנראות: $0.5e^{0.1x}$.

ב. $c = \{x \geq k\}$. ג. $c = \{x \geq 14.98\}$. ד. 0.776.

(7) א. שאלת הוכחה. ב. שאלת הוכחה. ג. $c = \{x > 0.95\}$, 0.0975.

(8) א. השערת האפס: שומר הלילה רושם כל מבקר אשר ניכנס לבניין. ההשערה האלטרנטיבית: שומר הלילה מדלג ברישום על מבקר, רושם אחד כן ואחד לא.

ב. אזור הדחייה יהיה היכן שיחס הנראות הינו גבוה לכן הוא יהיה מאינסוף ועד ערך k מסוים.

ג. נדחה את השערת האפס אם: $x > 2.996$. ד. $x_1 \cdot x_2 > k$.

(9) א. נדחה את H_0 עבור: $\frac{1}{2} \leq x \leq 1.0513$. ב. חישוב עוצמת המבחן: 0.668.

ג. i. אפשר לומר ללא חישוב נוסף שהעוצמה תגדל.

(10) א. $c = \{|x - 5| \leq a\}$. ב. $c = \{4.75 \leq x \leq 5.25\}$.

(11) כלל ההכרעה הוא נדחה את השערת האפס אם: $\bar{D} \leq 9.26$.

מבוא להסתברות וסטטיסטיקה

פרק 47 - מבוא לבדיקת השערות על פרמטרים

תוכן העניינים

235	1. הקדמה
239	2. סוגי טעויות

הקדמה:

רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטיקה. בבדיקת השערות על פרמטרים נעבוד לפי השלבים הבאים:

שלב א: נוהה את הפרמטר הנחקר.

שלב ב: נרשום את השערות המחקר.

השערת האפס המסומנות ב- H_0 .

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

השערה אלטרנטיבית (השערת המחקר) המסומנת ב- H_1 .

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

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

שלב ד: נרשום את כלל ההכרעה. בתהליך של בדיקת השערות יוצרים כלל שנקרא כלל הכרעה. הכלל יוצר אזורי שנקראים:

1. **אזור דחייה:**

דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.

2. **אזור קבלה:**

קבלה של השערת האפס ודחייה של האלטרנטיבה. כלל ההכרעה מתבסס על איזשהו סטטיסטי. אזור הדחייה מוכתב על ידי סיכון שלוקח החוקר מראש שנקרא רמת מובהקות ומסומן ב- α .

שלב ה: בתהליך יש ללכת לתוצאות המדגם ולחשב את הסטטיסטי המתאים ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה.

שלב ו: להסיק מסקנה בהתאם לתוצאות המדגם.

דוגמה (פתרון בהקלטה):

משרד הבריאות פרסם שמשקל ממוצע של תינוקות ביום לידתם בישראל 3300 גרם. משרד הבריאות רוצה לחקור את הטענה שנשים מעשנות בזמן ההיריון יולדות תינוקות במשקל נמוך מהממוצע. במחקר השתתפו 20 נשים מעשנות בהריון. להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:

$$n = 20, \bar{X} = 3120, S = 280$$

- א. מהי אוכלוסיית המחקר?
- ב. מה המשתנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

שאלות:

בשאלות הבאות, ענו על הסעיפים הבאים:

- א. מהי אוכלוסיית המחקר?
- ב. מה המשתנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

- (1) ממוצע הציונים בבחינת הבגרות באנגלית הנו 72 עם סטיית תקן 15 נקודות. מורה טוען שפיתח שיטת לימוד חדשה שתעלה את ממוצע הציונים. משרד החינוך החליט לתת למורה 36 תלמידים אקראיים. ממוצע הציונים של אותם תלמידים לאחר שלמדו בשיטתו היה 75.5.
- (2) לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 סמ"ק במדגם בגודל 25.
- (3) במשך שנים אחוז המועמדים שהתקבל לפקולטה למשפטים היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. מחקר מעוניין לבדוק האם השנה מקשים על הקבלה לפקולטה למשפטים.
- (4) בחודש ינואר השנה פורסם שאחוז האבטלה במשק הוא 8% במדגם עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים. רוצים לבדוק ברמת מובהקות של 5% האם כיום אחוז האבטלה הוא כמו בתחילת השנה.

תשובות סופיות:

- (1) א. נבחנים בבגרות באנגלית.
 ב. ציון.
 ג. ממוצע הציונים בשיטת לימוד חדשה.
 ד. $H_0: \mu = 72$
 $H_1: \mu > 72$
- (2) א. משקאות בבקבוק של חברה מסוימת.
 ב. נפח משקה בסמ"ק.
 ג. ממוצע נפח המשקה בבקבוק.
 ד. $H_0: \mu = 500$
 $H_1: \mu < 500$
- (3) א. מועמדים לפקולטה למשפטים.
 ב. משתנה דיכוטומי (התקבל, לא התקבל).
 ג. אחוז הקבלה.
 ד. $H_0: p = 0.25$
 $H_1: p < 0.25$
- (4) א. אזרחים בוגרים במשק.
 ב. משתנה דיכוטומי (מובטל, עובד).
 ג. אחוז האבטלה כיום.
 ד. $H_0: p = 0.08$
 $H_1: p \neq 0.08$

סוגי טעויות:

רקע:

בתהליך של בדיקת השערות יוצרים כלל שניקרא כלל הכרעה.
הכלל יוצר אזורים שנקראים:

1. אזור דחייה – דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.
2. אזור קבלה – קבלה של השערת האפס ודחייה של האלטרנטיבה.

כלל ההכרעה מתבסס על איזשהו סטטיסטי.
בתהליך יש ללכת לתוצאות המדגם ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה וכך להגיע למסקנה – המסקנה היא בעירבון מוגבל כיוון שהיא תלויה בכלל ההכרעה ובתוצאות המדגם. אם נשנה את כלל ההכרעה אז אנחנו יכולים לקבל מסקנה אחרת. אם נבצע מדגם חדש אז אנחנו עלולים לקבל תוצאה אחרת. לכן יתכנו טעויות במסקנות שלנו:

מציאות	הכרעה		
		H_0	H_1
	H_0	טעות מסוג 1 אין טעות	טעות מסוג 2 אין טעות

הגדרת הטעויות:

טעות מסוג ראשון: להכריע לדחות את H_0 למרות שבמציאות H_0 נכונה.

טעות מסוג שני: להכריע לקבל את H_0 למרות שבמציאות H_1 נכונה.

דוגמה (פתרון בהקלטה):

אדם חשוד בביצוע עבירה ונתבע בבית המשפט.
אילו סוגי טעויות אפשריות בהכרעת הדין?

שאלות:

- (1) לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 סמ"ק במדגם בגודל 25. בסופו של דבר הוחלט להכריע לטובת חברת המשקאות.
- א. רשמו את השערות המחקר.
 ב. מה מסקנת המחקר?
 ג. איזו סוג טעות יתכן וביצעו במחקר?
- (2) במחקר על פרמטר מסוים הוחלט בסופו של דבר לדחות את השערת האפס.
- א. האם ניתן לדעת אם בוצע טעות במחקר?
 ב. מה סוג הטעות האפשרית?
- (3) לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטיית תקן 0.4. ישנה טענה שכיום ממוצע מספר הילדים במשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה. על סמך תוצאות המדגם נקבע שלא ניתן לקבוע שבאופן מובהק תוחלת מספר הילדים למשפחה קטנה כיום.
- א. מהי אוכלוסיית המחקר?
 ב. מה המשתנה הנחקר?
 ג. מה הפרמטר הנחקר?
 ד. מה השערות המחקר?
 ה. מה מסקנת המחקר?
 ו. מהי סוג הטעות האפשרית במחקר?

תשובות סופיות:

- (1) א. $H_0: \mu = 500$
 ב. לא דחינו את H_0 .
 ג. טעות מסוג שני.
- (2) א. לא ניתן לדעת.
 ב. טעות מסוג ראשון.
- (3) א. משפחות כיום.
 ב. מס' הילדים.
 ג. תוחלת מספר הילדים למשפחה כיום.
 ד. $H_0: \mu = 2.3$
 $H_1: \mu < 2.3$
 ה. לא לדחות את H_0 . ו. טעות מסוג שני.

מבוא להסתברות וסטטיסטיקה

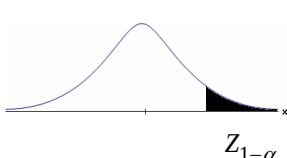
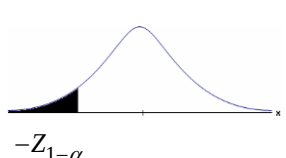
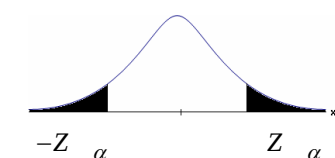
פרק 48 - בדיקת השערות על תוחלת (ממוצע)

תוכן העניינים

1. בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסיה ידועה 241
2. סיכוי לטעויות ועוצמה (ששונות האוכלוסיה ידועה) 245
3. בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסיה לא ידועה 251

בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסייה ידועה:

רקע:

השערת האפס: השערה אלטרנטיבה:	$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
תנאים:	1. σ ידועה 2. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0 :	$Z_{\bar{x}} > Z_{1-\alpha}$  $Z_{1-\alpha}$ דוחים את H_0	$Z_{\bar{x}} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ דוחים את H_0	$Z_{\bar{x}} < -Z_{\frac{\alpha}{2}}$ או $Z_{\bar{x}} > Z_{\frac{\alpha}{2}}$  $-Z_{\frac{\alpha}{2}}$ $Z_{\frac{\alpha}{2}}$ דוחים את H_0

$$Z_{\bar{x}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

סטטיסטי המבחן:

חלופה אחרת לכלל הכרעה:

$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$	נדחה H_0 אם מתקיים:

דוגמה:

יבול העגבניות מתפלג נורמלית עם תוחלת של 10 טון לדונם וסטיית תקן של 2.5 טון לדונם בעונה. משערים ששיטת זיבול חדשה תעלה את תוחלת היבול לעונה מבלי לשנות את סטיית התקן. נדגמו 4 חלקות שזוּבְלו בשיטה החדשה. היבול הממוצע שהתקבל היה 12.5 טון לדונם. בדקו את ההשערה ברמת מובהקות של 1%.

פיתרון:

אוכלוסייה: עגבניות.

המשתנה: $X =$ יבול העגבניות בטון לעונה.

הפרמטר: $\mu =$ תוחלת היבול בשיטת הזיבול החדשה.

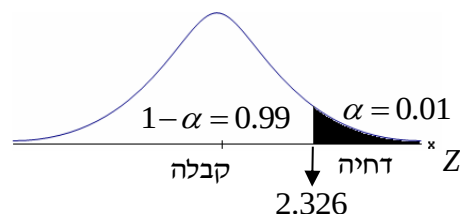
השערות:
 $H_0: \mu = 10$
 $H_1: \mu > 10$

תנאים:

1. $X \sim N$.

2. $\sigma = 2.5$.

כלל הכרעה:



נדחה את H_0 אם $Z_{\bar{X}} > 2.326$

תוצאות: $n = 4$, $\bar{x} = 12.5$

סטטיסטי המבחן: $Z_{\bar{X}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

נציב: $Z_{\bar{X}} = \frac{12.5 - 10}{\frac{2.5}{\sqrt{4}}} = 2 < 2.326$

מסקנה:

לא נדחה H_0 (נקבל H_0).

ברמת מובהקות של 1% לא נוכל לקבל את הטענה ששיטת הזיבול החדשה מעלה את תוחלת היבול של העגבניות.

שאלות:

- (1) ממוצע הציונים בבחינת הבגרות באנגלית הנו 72 עם סטיית תקן 15 נקודות. מורה טוען שפיתח שיטת לימוד חדשה שתעלה את ממוצע הציונים. משרד החינוך החליט לתת למורה 36 תלמידים אקראיים. ממוצע הציונים של אותם תלמידים לאחר שלמדו בשיטתו היה 75.5. בהנחה שגם בשיטתו סטיית התקן תהיה 15 מה מסקנתכם ברמת מובהקות של 5%?
- (2) לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 סמ"ק במדגם בגודל 25. א. מה מסקנתכם ברמת מובהקות של 2.5%? ב. האם ניתן לדעת מה תהיה המסקנה עבור רמת מובהקות הגבוהה מ-5%?
- (3) מהנדס האיכות מעוניין לבדוק אם מכונה מכיילת (מאופסת). המכונה כוונה לחתוך מוטות באורך 50 ס"מ. לפי נתוני היצרן סטיית התקן בחיתוך המוטות היא 0.5 ס"מ. במדגם של 50 מוטות התקבל ממוצע אורך המוט 50.93 ס"מ. מה מסקנתכם ברמת מובהקות של 5%?
- (4) המשקל הממוצע של הספורטאים בתחום ספורט מסוים הוא 90 ק"ג, עם סטיית תקן 8 ק"ג. לפי דעת מומחים בתחום יש צורך בהורדת המשקל ובשימוש בדיאטה מסוימת שצריכה להביא להורדת המשקל. לשם בדיקת יעילות הדיאטה נלקח מדגם מקרי של 50 ספורטאים ובתום שנה של שימוש בדיאטה התברר שהמשקל הממוצע במדגם זה היה 84 ק"ג. יש לבדוק בר"מ של 10%, האם הדיאטה גורמת להורדת המשקל.
- (5) לפי מפרט נתון, על עובי בורג להיות 4 מ"מ עם סטיית תקן של 0.2 מ"מ. במדגם של 25 ברגים העובי הממוצע היה 4.07 מ"מ. קבעו ברמת מובהקות 0.05, האם עובי הברגים מתאים למפרט. הניחו כי עובי של בורג מתפלג נורמלית וסטיית התקן של עובי בורג היא אכן 0.2 מ"מ.
- (6) במחקר נמצא שתוצאה היא מובהקת ברמת מובהקות של 5% מה תמיד נכון? בחרו בתשובה הנכונה.
 - א. הגדלת רמת המובהקות לא תשנה את מסקנת המחקר.
 - ב. הגדלת רמת המובהקות תשנה את מסקנת המחקר.
 - ג. הקטנת רמת המובהקות לא תשנה את מסקנת המחקר.
 - ד. הקטנת רמת המובהקות תשנה את מסקנת המחקר.

(7) חוקר ערך מבחן דו צדדי ברמת מובהקות של α והחליט לדחות את השערת האפס.

אם החוקר היה עורך מבחן צדדי ברמת מובהקות של $\frac{\alpha}{2}$ אזי בהכרח:

- א. השערת האפס הייתה נדחית.
- ב. השערת האפס הייתה לא נדחית.
- ג. לא ניתן לדעת מה תהיה מסקנתו במקרה זה.

(8) שני סטטיסטיקאים בדקו השערות: $H_0: \mu = \mu_0$ כנגד $H_1: \mu > \mu_0$,

עבור שונות ידועה ובאותה רמת מובהקות.

שני החוקרים קבלו אותו ממוצע במדגם אך לחוקר א' היה מדגם בגודל 100 ולחוקר ב' מדגם בגודל 200.

- א. אם חוקר א' החליט לדחות את H_0 , מה יחליט חוקר ב'? נמקו.
- ב. אם חוקר א' יחליט לא לדחות את H_0 , מה יחליט חוקר ב'? נמקו.

תשובות סופיות:

(1) נקבל H_0 , בר"מ של 5% לא נקבל את הטענה של המורה ששיטת הלימוד שלו מעלה את ממוצע הציונים.

(2) א. נדחה H_0 , בר"מ של 2.5% נקבל את תלונת אגודת הצרכנים בדבר הפחתת נפח המשקה בבקבוק.

ב. הגדלנו את רמת המובהקות לכן אנחנו נשארים בדחייה של H_0 והמסקנה לא תשתנה.

(3) נדחה H_0 , בר"מ של 5% נקבע שהמכונה לא מאופסת.

(4) נדחה H_0 , בר"מ של 0.1 נקבל את הטענה שהדיאטה יעילה ומפחיתה את המשקל הממוצע.

(5) נקבל H_0 , בר"מ של 0.05 נכריע שתוחלת עובי הבורג מתים למפרט.

(6) א'.

(7) ג'.

(8) א. לדחות. ב. לא ניתן לדעת.

סיכוי לטעויות ועוצמה (ששונות האוכלוסייה ידועה):

רקע:

מציאות	הכרעה		
		H_0	H_1
	H_0	אין טעות	טעות מסוג 1
	H_1	טעות מסוג 2	אין טעות

הגדרת הסתברויות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):

$$\alpha = P(H_0 \text{ נכונה} \mid \text{לדחות את } H_0) = P_{H_0}(H_0 \text{ לדחות})$$

הסיכוי לבצע טעות מסוג 2:

$$\beta = P(H_0 \text{ לקבל את } H_0 \mid \text{לדחות את } H_0) = P_{H_1}(H_0 \text{ לקבל})$$

רמת בטחון:

$$(\alpha - 1) = P(H_0 \text{ לקבל את } H_0 \mid \text{לדחות את } H_0) = P_{H_0}(H_0 \text{ לקבל})$$

עוצמה:

$$(\beta - 1) = \pi = P(H_0 \text{ לדחות את } H_0 \mid \text{לדחות את } H_0) = P_{H_1}(H_0 \text{ לדחות})$$

התהליך לחישוב סיכוי לטעות מסוג שני:

השערת האפס: השערה אלטרנטיבה:	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
תנאים:	1. σ ידועה 2. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0 :	$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
חישוב β :	$P_{\mu_1} \left(\bar{X} < \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right)$	$P_{\mu_1} \left(\bar{X} > \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right)$	$P_{\mu_1} \left(\mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$

התפלגות ממוצע המדגם: $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

התקנון:

דוגמה:

בתחילת השנה חשבון הטלפון הסלולארי הממוצע לאדם היה 200 ₪ עם סטיית תקן של 80 ₪ לחודש. בעקבות כניסתן של חברות טלפון סלולארית חדשות מעוניינים לבדוק האם כיום ממוצע חשבון הטלפון הסלולארי פחת. לצורך בדיקה דגמו באקראי 36 אנשים וחשבון הטלפון הסלולארי שלהם היה 150 ₪ בממוצע לחודש.

א. רשמו את השערות המחקר ובנו כלל הכרעה במונחי חשבון ממוצע מדגמי ברמת מובהקות של 5%.

ב. מה מסקנתכם? איזה סוג טעות אפשרית במסקנה?

ג. נניח שבמציאות כיום החשבון הממוצע הוא 160 ₪. מה הסיכוי לבצע טעות מסוג שני?

ד. אם נקטין את רמת המובהקות מסעיף א', כיצד הדבר ישפיע על התשובה מסעיף ג'?

פתרון:

א. אוכלוסייה: משלמי חשבון טלפון סלולאר כיום.

המשתנה: X = חשבון הטלפון החודשי בשקלים.

הפרמטר: μ .

השערות:
 $H_0: \mu = 200$
 $H_1: \mu < 200$

תנאים:

1. $\mu = 200$.

2. $n = 36$.

$$\bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, K = \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\alpha = 0.05$$

$$Z_{1-\alpha} = Z_{0.95} = 1.645$$

$$\text{נציב: שקלים } K = 200 - 1.645 \cdot \frac{80}{\sqrt{36}} = 178.07$$

כלל ההכרעה: דחה את H_0 אם שקלים $\bar{X} < 178.07$.

ב. ברמת מובהקות של 5% נכריע שאכן ממוצע חשבון הטלפון הסלולרי פחת מתחילת השנה.

ג. השערות:
 $H_0: \mu_0 = 200$
 $H_1: \mu < 200$

כלל ההכרעה: נדחה את H_0 אם $\bar{X} < 178.07$.

$$H_1: \bar{X} \sim N\left(160, \frac{80^2}{36}\right)$$

$$Z = \frac{178.07 - 160}{\frac{80}{\sqrt{36}}} = 1.36$$

$$\beta = P_{H_1} \left(\overset{\text{לקבל את}}{H_0} \right) = P_{H_1} (\bar{X} > 178.07) = 1 - \phi(136) = 1 - 0.9131 = 0.0869$$

ד. הקטנת α מגדילה את β .

שאלות:

(1) נתון ש: $X \sim N(\mu, \sigma^2 = 1)$.

- להלן השערות של חוקר לגבי הפרמטר μ : $H_0: \mu = 5$, $H_1: \mu = 7$. מעוניינים ליצור כלל הכרעה המתבסס על הסמך תצפית בודדת כך שרמת המובהקות תהיה 5%.
- א. עבור אילו ערכים של X שידגם נדחית השערת H_0 ?
- ב. מה הסיכוי לבצע טעות מסוג שני?
- ג. אם במדגם התקבל ש- $X = 6.9$ מה תהיה המסקנה ומה הטעות האפשרית?

- (2) לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטיית תקן 0.4. מעוניינים לבדוק אם כיום ממוצע מספר הילדים למשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה.
- א. רשמו כלל הכרעה במונחי ממוצע מדגם קריטי ברמת מובהקות של 5%.
- ב. בהמשך לסעיף א' מה תהיה המסקנה ומהי הטעות האפשרית במסקנה?
- ג. אם באמת ממוצע מספר הילדים במשפחה פחת לכדי 2.1 מהי העצמה של הכלל מסעיף א'?

- (3) להלן נתונים על תהליך של בדיקת השערות על תוחלת:

$$H_0: \mu = 200, H_1: \mu \neq 200, \sigma = 30, n = 225$$

- א. רשמו כלל הכרעה במונחי ממוצע מדגם קריטי וברמת מובהקות של 10%.
- ב. בהמשך לסעיף א', מהי העצמה אם התוחלת שווה ל-195?
- ג. הסבירו, ללא חישוב, איך העצמה תשתנה אם רמת המובהקות תהיה 5%?

- (4) מפעל לייצור צינורות מייצר צינור שקוטרו מתפלג נורמלית עם תוחלת של 50

מ"מ וסטית תקן של 6 מ"מ. במחלקת ביקורת האיכות דוגמים בכל יום 81 צינורות ומוודדים את קוטרים, בכדי לבדוק, בעזרת מבחן סטטיסטי, האם מכונת הייצור מכוילת כנדרש או שקוטר הצינורות קטן מהדרוש.

- א. רשמו את ההשערות ואת כלל ההכרעה ברמת מובהקות של 5%.
- ב. אם ביום כלשהו מכונת הייצור התקלקלה והיא מייצרת את הצינורות בקוטר שתוחלתו 48 מ"מ בלבד (סטית התקן לא השתנתה), מה ההסתברות שהתקלה לא תתגלה בביקורת האיכות? כיצד נקראת הסתברות זו?
- ג. הסבירו ללא חישוב כיצד התשובה לסעיף ב' תשתנה אם רמת המובהקות תגדל.
- ד. הסבירו ללא חישוב כיצד התשובה לסעיף ב' תשתנה אם התוחלת האמיתית היא 47 ולא 48 מ"מ.

- (5) להלן השערות של מחקר: $H_0: \mu = 50$, $H_1: \mu = 58$. מעוניינים לדגום 100 תצפיות. ידוע שסטיית התקן של ההתפלגות הינה 20.
- א. בנו כלל הכרעה שהסיכוי לטעות מסוג שני בו הוא 10%. מהי רמת המובהקות?
- ב. כיצד הייתה משתנה רמת המובהקות אם (כל סעיף בפני עצמו)?
- i. סטיית התקן הייתה יותר גדולה.
- ii. הסיכוי לטעות מסוג שני גדול יותר.

השאלות שלהלן הן שאלות רב-ברירה, בחרו בתשובה הנכונה ביותר:

- (6) אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו אזי:
- א. הסיכוי לטעות מסוג ראשון גדל.
- ב. העוצמה של המבחן גדלה.
- ג. הסיכוי לטעות מסוג שני גדל.
- ד. תשובות א' ו-ב' נכונות.
- (7) חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן:
- א. השערת האפס נכונה.
- ב. השערת האפס נדחתה.
- ג. השערת האפס לא נדחתה.
- ד. אף אחת מהתשובות לא נכונה בהכרח.

- (8) מה המצב הרצוי לחוקר המבצע בדיקת השערה:

$1 - \beta$	α
גדולה	א. גדולה
קטנה	ב. גדולה
גדולה	ג. קטנה
קטנה	ד. קטנה

- (9) נערך שינוי בכלל ההחלטה של בדיקת השערה מסוימת ובעקבותיו אזור דחיית H_0 קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה מכך:
- א. הן α , והן $1 - \beta$, יקטנו.
- ב. α יישאר ללא שינוי ואילו $1 - \beta$ יגדל.
- ג. α יגדל ואילו $1 - \beta$ יקטן.
- ד. הן α והן $1 - \beta$ יגדלו.

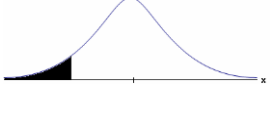
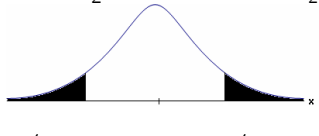
- (10) ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניח שלחץ הדם בקרב עיתונאים גבוה יותר מהממוצע באוכלוסייה. הוא לקח מדגם של 60 עיתונאים וקיבל ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק שלחץ הדם בקרב העיתונאים אינו גבוה יותר. מה הטעות האפשרית שהרופא עושה?
- טעות מסוג ראשון.
 - טעות מסוג שני.
 - טעות מסוג שלישי.
 - אין טעות במסקנתו.

תשובות סופיות:

- (1) א. מעל 6.645. ב. 0.3594.
ג. דחינו את H_0 , תתכן טעות מסוג ראשון.
- (2) א. נדחה H_0 אם $\bar{X} < 2.24$. ב. נדחה H_0 ג. 1.
- (3) א. נדחה H_0 אם $\bar{X} > 203.29$ או $\bar{X} < 196.71$. ב. 0.8051. ג. תקטן.
- (4) א. נדחה H_0 אם $\bar{X} < 48.9$. ב. 0.0885. ג. תקטן. ד. תקטן.
- (5) א. 0.0033. ב. i. רמת המובהקות הייתה קטנה.
ב. ii. רמת המובהקות הייתה גדלה.
- (6) ד'.
- (7) ג'.
- (8) ג'.
- (9) א'.
- (10) ב'.

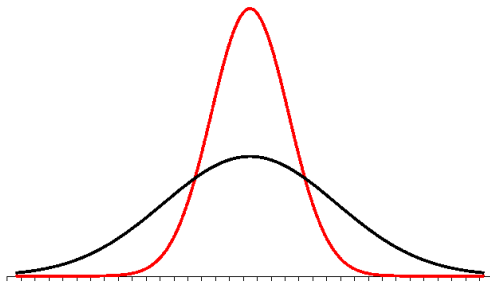
בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסייה לא ידועה:

רקע:

השערת האפס: השערה אלטרנטיבה:	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
תנאים:	1. σ אינה ידועה 2. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0 :	$t_{\bar{x}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha, n-1}$ - דוחים את H_0	$t_{\bar{x}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha, n-1}$ - דוחים את H_0	$t_{\bar{x}} < -t_{1-\frac{\alpha}{2}}^{(n-1)}$ או $t_{\bar{x}} > t_{1-\frac{\alpha}{2}}^{(n-1)}$  $-t_{1-\frac{\alpha}{2}, n-1}$ $t_{1-\frac{\alpha}{2}, n-1}$ - דוחים את H_0
חלופה לכלל ההכרעה: נדחה H_0 אם מתקיים:	$\bar{X} > \mu_0 + t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} < \mu_0 - t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$ או $\bar{X} < \mu_0 - t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$

$$t_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} : \text{סטטיסטי המבחן}$$

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



התפלגות T:

הינה התפלגות סימטרית פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים. התפלגות T תלויה במושג שנקרא דרגות חופש.

דרגות החופש הן: $df = n - 1$.

ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה. כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה (פתרון בהקלטה):

מפעל קיבל הזמנה לייצור משטחים בעובי של 0.1 ס"מ. כדי לבדוק האם המפעל עומד בדרישה נדגמו 10 משטחים ונמצא שהעובי הממוצע הוא 0.104 עם אומדן לסטיית תקן 0.002 ס"מ.

א. מהן השערות המחקר?

ב. מה ההנחה הדרושה לצורך פתרון?

ג. בדוק ברמת מובהקות של 5%.

שאלות:

(1) משך זמן ההחלמה בלקיחת אנטיביוטיקה מסוימת הוא 120 שעות בממוצע עם סטיית תקן לא ידועה. מעוניינים לבדוק האם אנטיביוטיקה אחרת מקטינה את משך זמן ההחלמה. במדגם של 5 חולים שלקחו את האנטיביוטיקה האחרת התקבלו זמני ההחלמה הבאים: 90, 95, 100, 80, 125 שעות. מה מסקנתכם ברמת מובהקות של 5%. מהי ההנחה הדרושה לצורך הפתרון?

(2) משרד הבריאות פרסם שמשקל ממוצע של תינוקות ביום היוולדם בישראל 3300 גר'. משרד הבריאות רוצה לחקור את הטענה שנשים מעשנות בזמן ההיריון יולדות תינוקות במשקל נמוך מהממוצע. במחקר השתתפו 20 נשים מעשנות בהריון. להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:

$$n = 20$$

$$\bar{x} = 3120$$

$$S = 280$$

מה מסקנתכם ברמת מובהקות של 5% מה יש להניח לצורך פתרון?

(3) ציוני מבחן אינטליגנציה מתפלגים נורמלית. בארה"ב ממוצע הציונים הוא 100. במדגם שנעשה על 23 נבחנים ישראלים, התקבל ממוצע ציונים 104.5 וסטיית התקן המדגמית 16. האם בישראל ממוצע הציונים שונה מארה"ב? הסיקו ברמת מובהקות של 5%.

(4) באוכלוסייה מסוימת נדגמו 10 תצפיות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{10} X_i = 750$$

$$\sum_{i=1}^{10} (X_i - \bar{X})^2 = 900$$

נתון שההתפלגות היא נורמלית.

בדוק ברמת מובהקות של 5% האם התוחלת של ההתפלגות שונה מ-80.

- (5) ליאור ורוני העלו את אותן השערות על ממוצע האוכלוסייה. כמו כן הם התבססו על אותן תוצאות של מדגם. ליאור השתמש בטבלה של התפלגות Z . רוני השתמשה בטבלה של התפלגות t . מה נוכל לומר בנוגע להחלטת המחקר שלהם? בחר בתשובה הנכונה.
- אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.
 - אם רוני תדחה את השערת האפס אז גם בהכרח ליאור.
 - שני החוקרים בהכרח יגיעו לאותה מסקנה.
 - לא ניתן לדעת על היחס בין דחיית השערת האפס של שני החוקרים.

- (6) נתון ש: $X \sim N(\mu, \sigma^2)$ כמו כן נתונות ההשערות הבאות: $H_0: \mu = \mu_0$. $H_1: \mu < \mu_0$. חוקר בדק את ההשערות הללו על סמך מדגם שכלל 10 תצפיות. σ^2 לא הייתה ידועה לחוקר. החוקר החליט לדחות את השערת האפס ברמת מובהקות של 5% לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תצפיות ושקלל את תוצאות אלה גם למדגם כך שכלל עכשיו 15 תצפיות. בחר בתשובה הנכונה:
- א. כעת בברור הוא ידחה את השערת האפס.
 - ב. כעת הוא דווקא יקבל את השערת האפס.
 - ג. כעת לא ניתן לדעת מה תהיה מסקנתו.

תשובות סופיות:

- (1) נדחה H_0 .
- (2) נדחה H_0 .
- (3) נקבל H_0 .
- (4) נקבל H_0 .
- (5) ב'.
- (6) ג'.