סטודנטים יקרים

לפניכם ספר תרגילים בקורס סטטיסטיקה להנדסה אזרחית הספר הוא חלק מקורס חדשני וראשון מסוגו בארץ בנושא זה, המועבר ברשת האינטרנט On-line.

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

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

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

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

GOOL
בשביל התירגול

אנו מאחלים לכם הצלחה מלאה בבחינות

צוות האתר GooL

גוּל זה בּוּל .בִּשְׁבִילךָ!

תוכן

פרק 1 - בעיות בסיסיות בהסתברות

רקע :

ניסוי מקרי : תהליך לו כמה תוצאות אפשריות. התוצאה המתקבלת נודעת רק לאחר ביצוע התהליך.

למשל : תוצאה בהטלת קובייה , מזג האוויר בעוד שבועיים .

מרחב מדגם : כלל התוצאות האפשריות בניסוי המקרי :

בהטלת קובייה : $\{1,2,3,4,5,6\}$.

מזג האוויר בעוד שבועיים : { נאה, שרבי, מושלג, גשום, מעונן חלקית, אביך }

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

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

לקבל תוצאה זוגית : $B=\{2,4,6\}$

גודל מרחב המדגם : מספר התוצאות האפשריות במרחב המדגם :

בהטלת הקובייה : $|\Omega|=6$

גודל המאורע : מספר התוצאות האפשריות במאורע עצמו.

בהטלת הקובייה : $|A|=2$ $|B|=3$

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

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

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

הסתברות במרחב מדגם אחיד :

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

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

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

הסתברות במרחב לא אחיד :

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

להלן התפלגות הציונים בכיתה מסוימת :

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

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

ב. מה ההסתברות שתלמיד אקראי שניבחר בכיתה יכשל?

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

הסתברות למאורע משלים :

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

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

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

תרגילים:

1. מהאותיות E, F ו-G יוצרים מילה בת 2 אותיות לא בהכרח בת משמעות.
   א. הרכב את כל המילים האפשריות.
   ב. רשום את המקרים למאורע:
   A- במילה נמצאת האות E.
   B- במילה האותיות שונות.
   ג. רשום את המקרים למאורע $\bar{A}$.

2. מטילים זוג קוביות.
   א. רשום את מרחב המדגם של הניסוי. האם המרחב מדגם הוא אחיד?
   ב. רשום את כל האפשרויות למאורעות הבאים:
   A- סכום התוצאות 7.
   C- מכפלת התוצאות 12.
   ג. חשב את הסיכויים למאורעות שהוגדרו בסעיף ב.

3. בוחרים באקראי ספרה מבין הספרות 0-9.
   א. מה ההסברות 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 שנספרו עבור כל משפחה בישוב מסוים:

מספר משפחות	מספר מקלטים
22	0
28	1
18	2
22	3
10	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	מספר מכוניות
20	0
40	1
100	2
30	3
10	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. מטילים מטבע רגיל 3 פעמים. בצד אחד של המטבע מוטבע עץ ובצד השני פלי.

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

ב. רשום את כל האפשרויות למאורעות הבאים:

A- התקבל פעם אחת עץ.

D- התקבל לפחות פלי אחד.

ג. מהו המאורע המשלים ל –D.

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

פתרונות:

שאלה 2

ג. הסיכוי ל-A: $\frac{1}{6}$

הסיכוי ל-B: $\frac{1}{9}$

שאלה 3

א. 0.4

ב. 0.4

ג. 0.5

שאלה 4

א. 0.22

ב. 0.78

ג. 0.32

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

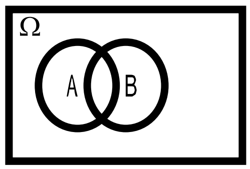
רקע:

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.



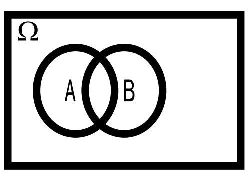
בהטלת קובייה , למשל, לקבל לפחות 5 : $A = \{5, 6\}$

לקבל תוצאה זוגית : $B = \{2, 4, 6\}$

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

פעולת איחוד :

נותנת את כל האפשריות שנמצאות לפחות באחת מהמאורעות. הסימון הוא : $A \cup B$ נותנת את אשר נימצא ב-A או B. כלומר, לפחות אחד מהמאורעות קורה.



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

לקבל תוצאה זוגית : $B = \{2, 4, 6\}$

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

<u>דוגמה (הפתרון נמצא בהקלטה)</u>

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

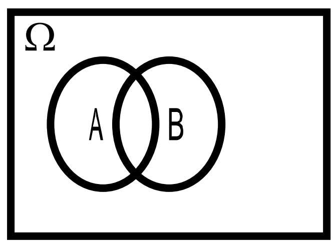
א. מה ההסתברות שלו לעבור את המבחן בסטטיסטיקה בלבד?

ב. מה ההסתברות שלו להיכשל בשני המבחנים?

ג. מה ההסתברות לעבור לפחות מבחן אחד?

<u>נוסחת החיבור לשני מאורעות</u> :

$$p(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$$



<u>חוקי דה מורגן לשני מאורעות:</u>

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

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

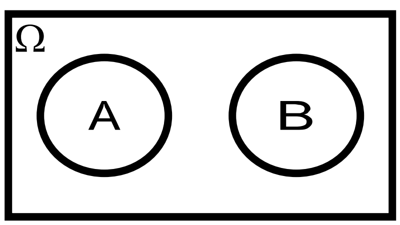
$$P(A \cup B) = 1 - P(\overline{A} \cap \overline{B})$$

שיטת ריבוע הקסם:

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

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

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



$A \bigcap B = \{\,\}$

$P(A \bigcap B) = 0$
$P(A \bigcup B) = P(A) + P(B)$

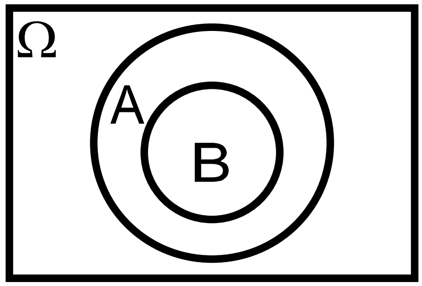
למשל , בהטלת קובייה

לקבל לפחות 5 : $A = \{5, 6\}$

לקבל 3 : $B = \{3\}$

$A \bigcap B = \{\,\}$

מאורעות מכילים :



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

קשר זה מסומן באופן הבא : $\boxed{B \subset A}$

$A \cap B = B$

$A \cup B = A$

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

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

למשל :

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

$B = \{2,4\}$

תרגילים:

1. מהאותיות E, F ו-G יוצרים מילה בת 2 אותיות לא בהכרח בת משמעות.

נגדיר את המאורעות הבאים:

E- במילה נמצאת האות E.

F- במילה אותיות שונות.

א. רשום את כל האפשרויות לחיתוך A עם B.

ב. רשום את כל האפשרויות לאיחוד של A עם B.

2. תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ומבחן בסטטיסטיקה. נגדיר את המאורעות הבאים:

A- לעבור את המבחן בסטטיסטיקה.

B- לעבור את המבחן בכלכלה.

העזר בפעולות חיתוך, איחוד ומשלים בלבד כדי להגדיר את המאורעות הבאים וסמן בדיאגראמת וון את השטח המתאים:

א. התלמיד עבר רק את המבחן בכלכלה.

ב. התלמיד עבר רק את המבחן בסטטיסטיקה.

ג. התלמיד עבר את שני המבחנים.

ד. התלמיד עבר לפחות מבחן אחד.

ה. התלמיד נכשל בשני המבחנים.

ו. התלמיד נכשל בכלכלה.

3. נתבקשתם לבחור ספרה באקראי. נגדיר את A להיות הספרה שנבחרה היא זוגית. נגדיר את B להיות הספרה שנבחרה קטנה מ-5.

א. רשמו את כל התוצאות למאורעות הבאים:

$A =$

$B =$

$\overline{B} =$

$A \cap B =$

$A \cup B =$

ב. חשבו את ההסתברויות לכל המאורעות מהסעיף הקודם.

4. נסמן ב- Ω את מרחב המדגם וב- ϕ קבוצה ריקה.
נתון כי A הינו מאורע בתוך מרחב המדגם.
להלן מוגדרים מאורעות שפתרונם הוא Ωאו ϕאו A.
קבע עבור כל מאורע מה הפתרון שלו.

$\overline{\overline{A}}$

$A \cap \phi$

$A \cup \phi$

$A \cap \Omega$

$A \cup \Omega$

$A \cap \overline{A}$

$\overline{\phi}$

$A \cup \overline{A}$

5. הוגדרו המאורעות הבאים:
A= אדם שגובהו מעל 1.7 מטר
B=אדם גובהו מתחת ל-1.8 מטר
קבע את גובהם של האנשים הבאים:
א. $A \cap B$
ב. $A \cup B$
ג. $\overline{A} \cap B$
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.

השתמש בפעולות איחוד, חיתוך והשלמה לתיאור המאורעות הבאים:

א. אדם דובר את כל שלוש השפות.

ב. אדם דובר רק עברית.

ג. אדם דובר לפחות שפה אחת מתוך השפות הללו.

ד. אדם אינו דובר אנגלית.

ה. קבוצת התלמידים דוברי 2 שפות בדיוק (מהשפות הנ"ל).

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

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

ב. מה ההסתברות ששתי המפלגות תעבורנה את אחוז החסימה?

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

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

א. איזה אחוז מהעובדים הם גברים אקדמאיים?

ב. איזה אחוז מהעובדים הם גברים או אקדמאיים?

ג. איזה אחוז מהעובדים הם נשים לא אקדמאיות?

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

א. ששתי המניות תעלנה.

ב. שאף אחת מהמניות לא תעלנה.

ג. שמניה A בלבד תעלה.

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

A- בקובייה האדומה התקבלה התוצאה 4 ובשחורה 2.

B- סכום התוצאות משתי הקוביות 6.

C- מכפלת התוצאות בשתי הקוביות 10.

א. האם A ו-B מאורעות זרים?

ב. האם המאורע B מכיל את המאורע A?

ג. האם A ו-C מאורעות זרים?

ד. האם A ו-C מאורעות משלימים?

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

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

א. האם A ו-B מאורעות זרים?

ב. חשב את $p(\bar{A}\cap B)$

12. מטבע הוטל פעמיים. נגדיר את המאורעות הבאים:

A- קיבלנו עץ בהטלה הראשונה.

B- קיבלנו לפחות עץ אחד בשתי ההטלות.

איזו טענה נכונה?

א. A ו- B מאורעות זרים.

ב. A ו- B מאורעות משלימים.

ג. B מכיל את A.

ד. A מכיל את B.

13. בהגרלה חולקו 100 כרטיסים על 3 מהם רשום חופשה ועל 2 מהם רשום מחשב שאר הכרטיסים ריקים. אדם קיבל כרטיס אקראי.

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

ב. מה ההסתברות לא לזכות בפרס?

.14

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

א. חשב את הסיכוי ל - $P(A\cap B)$

ב. האם A ו- B מאורעות זרים?

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

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

מה הסיכוי למאורע A ומה ההסתברות למאורע B?

.16 קבע אילו מהטענות הבאות נכונות:

א. $A\bigcap B=B\bigcap A$

ב. $\overline{A\bigcup B}=A\bigcap B$

ג. $A\bigcap B\bigcap C=A\bigcap B\bigcap(C\bigcup B)$

ד. $\overline{A\bigcap B\bigcap C}=\overline{A}\bigcup\overline{B}\bigcup\overline{C}$

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

א. האם יתכן ש-0.4= $p(A\bigcup B)$?

ב. האם יתכן ש -0.6= $p(A\bigcup B)$?

ג. אם A ו- B זרים מה הסיכוי $p(A\bigcup B)$?

ד. אם A מכיל את B מה הסיכוי $p(A\bigcup B)$?

18. מתוך אזרחי המדינה הבוגרים ל- 30% חשבון בבנק הפועלים. ל28% 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. הוכח: $p(\bar{A} \cap \bar{B}) = 1 - P(A) - P(B) + P(A \cap B)$

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

פתרונות:

שאלה 7

א. 0.24

ב. 0.04

ג. 0.16

שאלה 8

א. 10%

ב. 50%

ג. 50%

שאלה 9

א. 0.2

ב. 0.3

ג. 0.3

שאלה 10

א. לא.

ב. כן.

ג. כן.

ד. לא.

שאלה 11

א. כן

ב. 0.3

שאלה 12

התשובה הנכונה ג

שאלה 13

א. 0.05

ב. 0.95

שאלה 14

א. 0.06

ב. לא זרים

ג. 0.43

שאלה 18

א.0.19

ב.0.05

ג. 0.31

ד. 0.46

ה.0.12

ו.0.41

ז.0.59

פרק 3 - הסתברות מותנית - במרחב מדגם אחיד

רקע:

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

ההסתברות של A בהינתן ש- B כבר קרה:

$$P(A|B)$$

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

למשל, (פתרון בהקלטה)

מטילים קובייה.
נגדיר:
A – התוצאה זוגית.
B – התוצאה גדולה מ-3.
נרצה לחשב את:
$P(A|B)$

<u>תרגילים:</u>

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

2. יוסי הטיל קובייה. מה הסיכוי שקיבל את התוצאה 4 אם ידוע שהתוצאה שהתקבלה זוגית ?

3. מטילים צמד קוביות.
נגדיר:
A – סכום התוצאות בשתי ההטלות הינו 7
B – מכפלת התוצאות 12
חשבו את $P(A|B)$.

4. הוטל מטבע פעמיים. ידוע שהתקבל לכל היותר ראש אחד, מה הסיכוי שהתקבלו שני ראשים ?

5. אדם הטיל זוג קוביות והתקבל שהתוצאות זהות. מה הסיכוי שלפחות אחת התוצאות 5?

6. אדם הטיל זוג קוביות והתקבל לפחות פעם אחת 4. מה הסיכוי שאחת התוצאות 5?

7. נבחרה משפחה בת שני ילדים. ידוע שאחד הילדים בן. מה ההסתברות שבמשפחה שני בנים בקרב הילדים?

8. נבחרה משפחה בת שלושה ילדים. נתון שהילד האמצעי בן. מה הסיכוי שיש בנות בקרב הילדים?

השאלות הבאות משלבות קומבינטוריקה:

9. בכיתה 6 בנים ו-7 בנות. נבחרו ארבעה ילדים מהכיתה.
אם ידוע שנבחרו 2 בנים ושתי בנות , מה הסיכוי שאלעד לא נבחר?

10. חמישה חברים יצאו לבית קולנוע והתיישבו זה ליד זה באקראי בכיסאות מספר 5 עד 9.
אם ידוע שערן ודין התיישבו זה ליד זה . מה ההסתברות שהם יושבים בכיסאות מספר 6 ו 7 ?

פתרונות:

שאלה 1

0.2

שאלה 2

1/3

שאלה 3

0.5

שאלה 4

0

שאלה 5

1/6

שאלה 6

2/11

שאלה 7

1/3

שאלה 8

3/4

שאלה 9

2/3

שאלה 10

1/4

פרק 4 - הסתברות מותנית - מרחב לא אחיד

רקע:

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

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

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

במכנה : הסיכוי למאורע שנתון שהתרחש :

למשל,

נלקחו משפחות שיש להם שתי מכוניות. ל- 30% מהמשפחות הללו המכונית הישנה יותר היא מתוצרת אירופה ואצל 60% מהמשפחות הללו המכונית החדשה יותר מתוצרת אירופה. כמו כן 15% מהמשפחות הללו שתי המכוניות הן מתוצרת אירופאית.

אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית? (פתרון בהקלטה)

תרגילים:

1. תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ומבחן בסטטיסטיקה:
נגדיר את המאורעות הבאים: A- לעבור את המבחן בסטטיסטיקה. B– לעבור את המבחן בכלכלה.
כמו כן נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8 והסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9. הסיכוי לעבור את שני המבחנים הנו 0.75. חשבו את הסיכויים למאורעות הבאים:
   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 שני המבחנים?

2. במדינה שתי חברות טלפון סלולארי "סופטי" ו"בלי". 30% מהתושבים הבוגרים רשומים אצל חברת "בלי". 60% מהתושבים הבוגרים רשומים אצל חברת "סופטי".
ל-15% מהתושבים הבוגרים אין טלפון סלולארי בכלל.
   א. איזה אחוז מהתושבים הבוגרים רשומים אצל שתי החברות?
   ב. נבחר אדם שרשום אצל חברת "סופטי", מה ההסתברות שהוא רשום גם אצל חברת "בלי"?
   ג. אם אדם לא רשום אצל חברת "בלי", מה ההסתברות שהוא כן רשום בחברת "סופטי"?
   ד. אם אדם רשום אצל חברה אחת בלבד, מה ההסתברות שהוא רשום בחברת "סופטי"?

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, מתוך העצמאיים 30 הם אקדמאיים.
   א. בנו טבלת שכיחות משותפת לנתונים.
   ב. נבחר אדם אקראי מהי 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 הבן אדם שנבחר הוא לא אקדמאי, מה ההסתברות שהוא שכיר?

5. 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, 5% מחזיקים בכל שלושת הכרטיסים הנ"ל.

א. אם לאדם יש ויזה , מה הסיכוי שאין לו כרטיס ישראכרט?

ב. אם לאדם שני כרטיסי אשראי , מה הסיכוי שאין לו כרטיס ישראכרט?

ג. אם לאדם לפחות כרטיס אשראי אחד, מה הסיכוי שאין לו כרטיס ישראכארט?

פתרונות:

שאלה 1

א. 0.833

ב. 0.9375

ג. 0.0625

ד. 0.5

ה. 0.789

שאלה 2

א. 5%

ב. 0.0833

ג. 0.786

ד. 0.6875

שאלה 3

א. 0.4

ב. $\frac{2}{3}$

ג. 0.25

ד. $\frac{6}{7}$

פרק 5 - דיאגרמת עצים, נוסחת בייס ונוסחת ההסתברות השלמה

רקע:

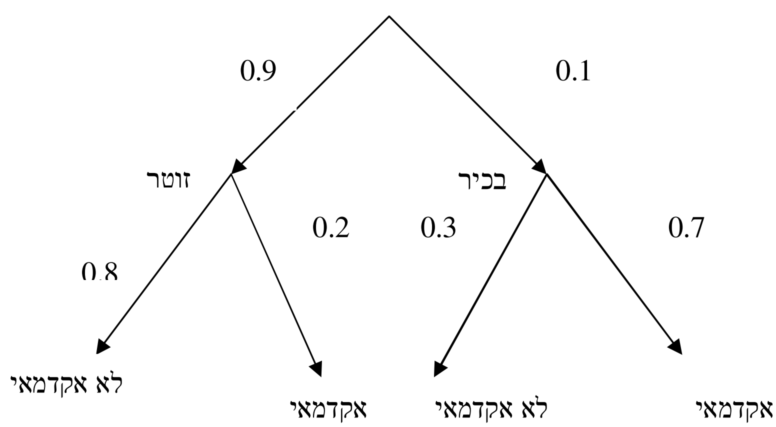
נשתמש בשיטה זו כאשר יש תרגיל שבו התרחשות המאורעות היא בשלבים , כך שכל תוצאה של כל שלב תלויה בשלב הקודם, פרט לשלב הראשון :

למשל,

בחברה מסוימת 10% מוגדרים בכירים והיתר מוגדרים זוטרים .

מבין הבכירים 70% הם אקדמאים ומבין הזוטרים 20% הם אקדמאים.

נשרטט עץ שיתאר את הנתונים , השלב הראשון של העץ אינו מותנה בכלום ואילו השלב השני מותנה בשלב הראשון.



כדי לקבל את הסיכוי לענף מסוים נכפיל את כל ההסתברויות על אותו ענף.

נבחר אדם באקראי מאותה חברה.

א. מה הסיכוי שהוא בכיר אקדמאי ?

$0.1*0.7=0.07$

ב. מה הסיכוי שהוא זוטר לא אקדמאי ?

$0.9*0.8=0.72$

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

ג. מה הסיכוי שהוא אקדמאי ?

$0.1*0.7+0.9*0.2=0.25$

ד. נבחר אקדמאי מה ההסתברות שהוא עובד זוטר?

מדובר כאן על שאלה בהסתברות מותנה ולכן נשתמש בעיקרון של הסתברות מותנה

$$P(zutar \mid academay) = \frac{0.9*0.2}{0.25} = \frac{0.18}{0.25} = 0.72$$

<u>נוסחת ההסתברות השלמה</u>

B מאורע כלשהו, $A_1,...,A_n$ חלוקה ממצה של Ω.

אזי: $P(B) = \sum_{i=1}^{n} P(A_i) \cdot P(B/A_i)$

<u>נוסחת בייס</u>

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

תרגילים:

1. בשקית סוכריות 4 סוכריות תות ו-3 לימון . מוציאים באקראי סוכרייה אם היא בטעם תות אוכלים אותה ומוציאם סוכרייה נוספת , אך אם היא בטעם לימון מחזירים אותה לשקית ומוציאים סוכרייה נוספת.

א. מה ההסתברות שהסוכרייה הראשונה שהוצאה בטעם תות והשנייה בטעם לימון ?

ב. מה ההסתברות שהסוכרייה השנייה בטעם לימון?

2. באוכלוסיה 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 אם נכשלתם במבחן מסוים אינכם ניגשים 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 ש68% 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 ש האנייה 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%. במחלות 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 לשאלה מסוימת את התשובה הוא p , אם 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 :

א. דולב, רונית, דולב

ב. רונית, דולב, רונית

בכל משחק מישהו חייב לנצח (אין תיקו). האדם ינצח בטורניר רק אם ינצח בשני משחקים ברציפות.
נתון שדולב שחקן טוב יותר מאשר רונית. איזו אפשרות עדיפה יותר על האדם כדי לנצח בטורניר?

פתרונות:

שאלה 1

א. 2/7

ב. 23/49

שאלה 2

א. 6%

ב. 58%

ג. 0.241

ד. 0.2

שאלה 3

א. 0.544

ב. 0.5

שאלה 4

א. 9%

ב. 0.09375

ג. 0.9722

שאלה 8

א. 0.0886

ב. 0.2889

ג. 0.3137

ד. 0.8778

שאלה 9

$$\frac{kp}{1+(k-1)p}$$

שאלה 10

אפשרות א עדיפה

פרק 6 - תלות ואי תלות בין מאורעות

רקע:

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

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

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

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 :

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

ב. מה הסיכוי להיכשל בשני הניסויים ?

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

<u>הרחבה</u> : <u>אי תלות בין n מאורעות</u>

n מאורעות $A_1,...,A_n$ הם בלתי תלויים אם ורק אם :

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

<u>תרגילים:</u>

1. נתון:

$$p(A)=0.2$$
$$P(B)=0.5$$
$$P(A\bigcup B)=0.6$$

האם המאורעות הללו בלתי תלויים?

2. תלמיד ניגש לשני מבחנים שהצלחתם לא תלויה זו בזו. הסיכוי שלו להצליח במבחן הראשון הוא 0.7 והשני 0.4 .

א. מה הסיכוי להצליח בשני המבחנים יחדו?

ב. מה הסיכוי שניכשל בשני המבחנים ?

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

א. מה ההסתברות ששניהם מובטלים?

ב. מה ההסתברות שלפחות אחד מהם מובטל?

4. מוצר צריך לעבור בהצלחה ארבע בדיקות בלתי תלויות לפני שיווקו, אחרת הוא נפסל ולא יוצא לשוק. הסיכוי לעבור בהצלחה כל אחת מהבדיקות הוא 0.8. בכל מקרה מבוצעות כל 4 הבדיקות.

א. מה הסיכוי שהמוצר יפסל?

ב. מה ההסתברות שהמוצר יעבור בהצלחה לפחות בדיקה אחת?

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

א. מה ההסתברות שכולם מובטלים במדגם?

ב. מה ההסתברות שלפחות אחד מהם מובטל?

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

7. הוכח אם

$$P(A/B)=P(B/A)$$

אז:

$$P(A)=P(B)$$

8. קבע אילו מהטענות הבאות נכונות. נמק!

א. אם $p(A\bigcup B)=p(A)\bullet p(B)$ אזי המאורעות בלתי תלויים.

ב. מאורע A כלול במאורע B. $P(A)>0$, $0<p(B)<1$ לכן $p(A/B)<p(A)$.

ג. A ו-B מאורעות זרים שסיכוייהם חיובים לכן הם מאורעות תלויים.

ד. A ו-B מאורעות תלויים שסיכוייהם חיובים לכן Aו- B מאורעות זרים.

ה. $P(\overline{A}\bigcap \overline{B})=1-P(A)-P(B)$ לכן A ו- B מאורעות זרים.

פתרונות :

שאלה 1

כן

שאלה 2

א. 0.28

ב. 0.18

שאלה 3

א. 0.0064

ב. 0.1536

שאלה 4

א. 0.5904

ב. 0.9984

שאלה 8

א. לא נכון

ב. לא נכון

ג. נכון

ד. לא נכון

ה. נכון

פרק 7 - שאלות מסכמות בהסתברות

1. נלקחו משפחות שיש להם שתי מכוניות. ל- 30% מהמשפחות הללו המכונית הישנה יותר היא מתוצרת אירופה ואצל 60% מהמשפחות הללו המכונית החדשה יותר מתוצרת אירופה. כמו כן 15% מהמשפחות הללו שתי המכוניות הן מתוצרת אירופאית.

א. מה ההסתברות שמשפחה אקראית בת שתי מכוניות תהיה ללא מכוניות מתוצרת אירופה?

ב. מה ההסתברות שלפחות מכונית אחת תהיה אירופאית?

ג. ידוע שלמשפחה יש מכונית אירופאית . מה ההסתברות שרק המכונית החדשה שלה היא מתוצרת אירופאית?

ד. אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית?

2. במדינת "שומקום " 50% מהחלב במרכולים מיוצר במחלבה א' 40% במחלבה ב' והיתר במחלבה ג'.3% מתוצרת מחלבה א' מגיעה חמוצה למרכולים ואילו במחלבה ב' 10%. כמו כן ידוע שבמדינת "שומקום" בסך הכול 7.5% מהחלב חמוץ.

א. איזה אחוז מהחלב שמגיע למרכול ממחלבה ג' חמוץ?

ב. אם נרכש חלב חמוץ במרכול. מה הסיכוי שהוא יוצר במחלבה ג?

ג. ברכישת חלב נימצא 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 בילוים עם מספר אפשרויות בילוי:

בהסתברות של 0.3 הם ייצאו לבאולינג

בהסתברות של 0.5 הם ייצאו לבית קפה

בהסתברות של 0.7 הם יצאו לפחות לאחד מהם, באולינג/קפה.

א. מה ההסתברות שהם יצאו רק לבאולינג?

ב. האם המאורעות "לצאת לבאולינג" לצאת לבית קפה" זרים?

ג. האם המאורעות "לצאת לבאולינג" לצאת לבית קפה" תלויים?

ד. מה ההסתברות שיום אחד הם יצאו רק לבאולינג וביום למחרת לא יצאו לאף אחד מהמקומות?

4. 70% מהנבחנים בסטטיסטיקה עוברים את מועד א׳. כל מי שלא עובר את מועד א׳ ניגש לעשות מועד ב׳, מתוכם 80% עוברים אותו. מבין אלה שנכשלים בשני המועדים 50% נרשמים לקורס מחדש, והיתר פורשים מהתואר.

א. מה הסיכוי שסטודנט אקראי עבר את הקורס?

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

ג. מה אחוז הסטודנטים שפורשים מהתואר?

ד. נבחרו 2 סטודנטים אקראיים רונית וינאי, מה ההסתברות שרונית עברה במועד א׳ ושינאי עבר במועד ב׳?

5. באוכלוסיה מסוימת 40% הם גברים והיתר הן נשים. מבין הגברים 10% מובטלים. בסך הכול 13% מהאוכלוסייה מובטלת.

א. מה אחוז האבטלה בקרב הנשים?

ב. נבחר אדם מובטל, מה ההסתברות שזו אישה?

ג. 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, ב- B את ההטלה השנייה ראש.

א. חשבו את הסיכויים למאורעות A ו- B.

ב. האם המאורע A ו- B בלתי תלויים?

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

7. ערן מעוניין למכור את רכבו, הוא 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.

א. מה אחוז האנשים מאלה שמעוניינים לרכוש רכב משומש יראו את 2 המודעות?

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

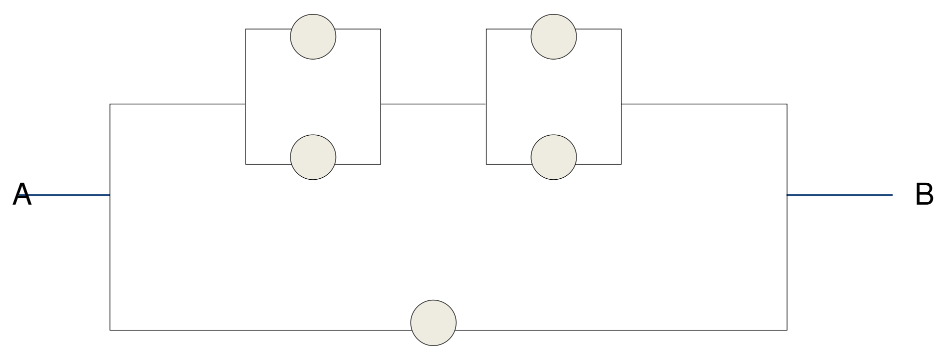
ג. האם המאורעות: "לראות את המודעה באינטרנט" ו"לראות את המודעה בעיתון" בלתי תלויים?

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

1. מה ההסתברות שאדם המעוניין לרכוש רכב משומש יתקשר לערן?

2. 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.

הוכח שהסיכוי שהמערכת תפעל:

$$P+(1-P)(2P-P^2)^2$$

פתרונות:

שאלה 1

א. 0.25

ב. 0.75

ג. 0.6

ד. 0.5

שאלה 2

א.0.2

ב. 0.267

ג. 0.524

ד. המאורעות תלויים.

שאלה 3

א. 0.2

ב. המאורעות אינם זרים.

ג. המאורעות הללו תלויים.

ד. 0.06

שאלה 4

א. 0.94

ב. 0.255

ג. 0.03

ד. 0.168

שאלה 5

א.15%

ב. 0.692

ג. לא זרים ותלויים.

שאלה 6

א. 0.65

ב. A ו-B תלויים.

ג. 0.5384

שאלה 7

א. 8%

ב. 0.733

ג. תלויים.

ד. 1. 0.478

2. 0.15

<u>פרק 8 - המשתנה המקרי הבדיד - פונקציית ההסתברות</u>

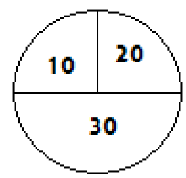
<u>רקע:</u>

משתנה מקרי בדיד : הנו משתנה היכול לקבל כמה ערכים בודדים בהסתברויות שונות.
מתארים את המשתנה המקרי על ידי פונקציית הסתברות.

פונקצית הסתברות : פונקציה המתאימה לכל ערך אפשרי של המשתנה את ההסתברות שלה.

סכום ההסתברויות על פונקציית ההסתברות חייב להיות 1.

למשל, בקזינו יש רולטה כמוראה בשרטוט :



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה בש״ח.
בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד (פתרון בהקלטה).

תרגילים:

1. ידוע שביישוב מסוים התפלגות מספר המכוניות למשפחה הוא:

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 C,B,A 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 והסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9. הסיכוי לעבור את שני המבחנים הנו 0.75. יהי X מספר המבחנים שהסטודנט עבר. בנה את פונקצית 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. בנה את פונקצית 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. הסיכוי שפרויקט ג' יצליח הינו 0.9. נתון שהצלחת כל פרויקט בלתי תלויה זו בזו. נגדיר את X להיות מספר הפרויקטים שיצליחו. בנה את פונקצית ההסתברות של X.

6. להלן פונקציית הסתברות של משתנה מקרי כלשהו :

$$P\left(X=k\right)=\frac{k}{A}$$
$$k=1,2...4$$

מצא 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.

פתרונות

שאלה 3

2	1	0	x
0.75	0.20	0.05	P(x)

שאלה 4

4	3	2	1	x
0.343	0.147	0.21	0.3	P(x)

שאלה 5

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

שאלה 6

10

פרק 9 - המשתנה המקרי הבדיד - תוחלת, שונות וסטיית תקן

רקע:

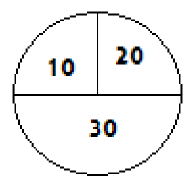
$$E(X) = \sum_i x_i P(x_i) = \mu$$
$$V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = \sigma^2$$

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

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

סטיית תקן- שורש של השונות. – הפיזור הממוצע הצפוי סביב התוחלת.

למשל, בקזינו רולטה כמוראה בשרטוט :



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה בש"ח.

x	10	20	30
P(x)	0.25	0.25	0.5

$$E(X) = 10\cdot 0.25 + 20\cdot 0.25 + 30\cdot 0.5 = 22.5 = \mu$$
$$V(X) = \sum_i (x_i - \mu)^2 P(x_i) = (10-22.5)^2\cdot 0.25 + (20-22.5)^2\cdot 0.25 + (30-22.5)^2\cdot 0.5$$
$$= 68.75 = \sigma^2$$

כדי לחשב את סטיית התקן נוציא שורש לשונות :

$$\sigma_x = \sqrt{V(X)} = \sqrt{68.75} = 8.29$$

תרגילים:

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

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

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 של הישוב. ל-40% חשבון בנק בסניף הלאומי של הישוב. ל20% מהתושבים הבוגרים אין חשבון בנק בישוב. יהי X מס' סניפי הבנק שלבוגר בישוב יש חשבון. חשב 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 וסטיית תקן 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 לא להשתמש בו שוב. נסמן ב-X את מספר הניסיונות עד שהדלת תפתח.

א. בנה את פונקצית ההסתברות של X.

ב. חשב את התוחלת והשונות של X.

5. נתונה פונקצית ההסתברות של המשתנה המקרי X :

x	2	4	6	8
P(x)		0.3		0.2

כמו כן נתון ש : $E(X)=4.2$

א. מצא את ההסתברויות החסרות בטבלה.

ב. חשב את $V(X)$.

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

7. להלן ההתפלגות של משתנה מקרי X .

P	X
¼	1
½	3
¼	K

מהו הערך K שייתן ערך מינימלי לשונות של X.

פתרונות:

שאלה 1

תוחלת : 2   שונות : 796

שאלה 3

ב . תוחלת : 3.36   סטיית תקן : 1.603

שאלה 4

א.

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

ב. תוחלת : 3

שונות 2

שאלה 5

א.

x	2	4	6	8
P(x)	0.4	0.3	0.1	0.2

ב. 5.16

שאלה 6

x	-5	0	5
P(x)	0.2	0.6	0.2

שאלה 7

2.33

פרק 10 - המשתנה המקרי הבדיד - טרנספורמציה לינארית

רקע

מצב שבו מבצעים הכפלה של קבועה ו או הוספה של קבוע על המשתנה המקורי. (כולל גם חלוקה של קבוע והחסרה של קבוע)

אם $Y = aX + b$

אזי:

$$E(Y) = aE(X) + b$$

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

$$\sigma_Y = |a|\sigma_x$$

שלבי העבודה:

1. נזהה שמדובר בטרנספורמציה ליניארית (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 :

$$E(X) = 22.5 = \mu$$

$$V(X) = 68.75 = \sigma^2$$

תרגילים:

1. סטודנט ניגש ל- 5 קורסים הסמסטר. נניח שכל קורס שסטודנט מסיים מזכה אותו ב-4 נקודות אקדמאיות. חשב 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 הינו 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 לעלות 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 $Y=7-X$.

   חשב את: $E(Y)$ ו- $V(Y)$.

5. אדם החליט לבטח את רכבו, שווי רכבו 100,000 ₪.

   להלן התביעות האפשריות והסתברותן:

   בהסתברות של 1/1000 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 פונקצית ההסתברות של X.

   ב. חשבו את התוחלת והשונות של גובה התביעה.

   ג. פרמיית הביטוח היא 4,000 ₪, מהי התוחלת ומהי השונות של רווח חברת הביטוח לביטוח הרכב הנ"ל?

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

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

כמו כן נתון שצפי מספר התשובות הנכונות בבחינה הוא 7.35.

א. השלימו את פונקציית ההסתברות.

ב. חשבו את השונות מספר התשובות הנכונות בבחינה.

ג. הציון בבחינה מחושב באופן הבא: כל שאלה נכונה מזכה ב-10 נקודות. לכל שאלה שגוייה, מופחתת נקודה. מהי התוחלת ומה השונות של הציון בבחינה?

7. להלן פונקצית הסתברות של משתנה מקרי כלשהו:

$$P\left(X=k\right)=\frac{k}{A}$$

$$k=1,2...4$$

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: $\frac{X}{2}-4$

פתרונות :

שאלה 1:

תוחלת: 14 שונות: 32

שאלה 2:

תוחלת: 8 שונות: 12

שאלה 3:

תוחלת: 13.2

סטיית תקן: 5.5

שאלה 4:

תוחלת: 3

שונות: 3

שאלה 6:

ב. $V(X)=1.8275$

שאלה 7 :

א. A=10

ב. $E(X)=3$
$V(X)=1$

ג. $E(X^3)=35.4$
$V(X^3)=616.84$

ד. $E(y)=-2.5$
$V(y)=0.25$

פרק 11 - תוחלת ושונות של סכום משתנים מקריים

רקע:

אם $X_1, X_2, ..., X_n$ משתנים מקרים אזי:

$$E(T) = E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

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

$$V(T) = V(X_1 + X_2 + ... + X_n) = V(X_1) + V(X_2) + ... + V(X_n)$$

למשל,

אדם משחק בשני משחקי מזל בלתי תלויים . תוחלת סכות 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?

<u>תרגילים:</u>

1. הרווח ממניה א׳ הוא עם תוחלת של 5 ושונות 10. הרווח ממניה ב׳ הוא עם תוחלת של 4 ושונות 5. 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 הזכיה במשחק הראשון.

Y= סכום הזכייה במשחק השני.

נתון:

$$E(x)=10 \qquad \sigma(X)=3$$
$$E(y)=12 \qquad \sigma(Y)=4$$

מהי התוחלת ומהי סטיית התקן של סכום הזכייה בשני המשחקים?

4. ברולטה הסיכוי לזכות ב- 30 ₪ הוא חצי וב-10 ₪ רבע כך גם ב- 20 ₪ . מה היא התוחלת והשונות של סכום הזכייה הכולל לאדם המשחק ברולטה 4 פעמים.

5. נתון משתנה מקרי בעל פונקציית ההסתברות הבאה:

$$P(X=K)=\begin{cases} \dfrac{A}{K-1} & K=2,3,4,5 \\ 0 & \text{אחרת} \end{cases}$$

א. מצא את ערכו של A.

ב. חשב את התוחלת והשונות של X.

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

פתרונות:

שאלה 1

תוחלת: 9

שונות : 15

שאלה 3

תוחלת :22

סטיית תקן: 5

שאלה 4

תוחלת : 90

שונות : 275

שאלה 5

א. $A=\frac{12}{25}=0.48$

ב. תוחלת 2.92

שונות 1.1136

ג. תוחלת 2.92n

שונות 1.1136n

פרק 12 - התפלגויות בדידות מיוחדות - התפלגות בינומית

רקע:

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

בהתפלגות בינומית חוזרים על אותו ניסוי ברנולי n פעמים באופן בלתי תלוי זה בזה.

מגדירים את X להיות מספר ההצלחות שהתקבלו בסך הכול.

נסמן ב p את הסיכוי להצלחה בניסוי בודד וב q את הסיכוי לכישלון בניסוי בודד.

ואז נגיד ש: $X \sim B(n, p)$.

<u>פונקציית ההסתברות של X</u>:

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

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

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

<u>תוחלת</u>: $E(X) = np$

<u>שונות</u>: $V(X) = npq$

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

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

<u>דוגמה :</u> (פתרון בהקלטה)

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

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

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

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

תרגילים:

1. במדינה 10% מהאוכלוסייה מובטלת. נבחרו 5 אנשים באקראי מאותה אוכלוסיה.
נגדיר את X להיות מספר המובטלים שהתקבלו במדגם.
א. מהי ההתפלגות של X?
ב. מה ההסתברות שיהיה בדיוק מובטל אחד?
ג. מה ההסתברות שכולם יעבדו במדגם ?
ד. מה ההסתברות ששלושה יעבדו במדגם ?
ה. מה ההסתברות שלפחות אחד יהיה מובטל ?
ו. מה תוחלת ומהי השונות של מספר המובטלים במדגם ?

2. על פי נתוני משרד התקשורת ל-70% מהאוכלוסייה יש סמארט-פון. נבחרו 10 אנשים באקראי. נגדיר את X כמספר האנשים שנדגמו עם סמארט-פון.

א. מהי ההתפלגות של X? הסבירו.
ב. מה ההסתברות שבמדגם ל-8 אנשים יש סמארט-פון?
ג. מה ההסתברות שבמדגם לפחות ל-9 יהיו סמארט-פון?
ד. 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.
א. מה ההסתברות שיפסיד בכל המכונות?
ב. מה ההסתברות שיזכה בדיוק בשתי מכונות?
ג. מה ההסתברות שיזכה ביותר כסף מה- 30 ₪ שהשקיע?
ד. 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 מהמדינה הנ"ל.
א. מה ההסתברות ש-5 מהם אקדמאים?
ב. 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 היא 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. מטילים מטבע הוגן 5 פעמים. נגדיר את X – מספר הפעמים שהתקבל עץ. חשבו את $E(x^2)$.

פתרונות :

שאלה 2:

ב. 0.2335

ג. 0.1493

ד. תוחלת : 7

סטיית תקן : 1.449

שאלה 3 :

א. 0.5314

ב. 0.0984

ג. 0.1143

ד. תוחלת : 18-

סטיית תקן : 14.697

שאלה 4:

א. 0.1789

ב. 2

שאלה 5 :

א. 0.1956

ב. 0.4253

שאלה 6 :

א. 0.85

ב. 0.182

ג. תוחלת : 82 נקודות

שונות : 91.8 נקודות

שאלה 7 :

א. 0.401

ב. תוחלת : 8.025

סטיית תקן : 2.193

שאלה 8 :

7.5

פרק 13 - התפלגויות בדידות מיוחדות - התפלגות גיאומטרית

רקע:

חוזרים באופן בלתי תלוי על אותו ניסוי ברנולי.

X – מוגדר להיות מספר הניסויים שבוצעו עד ההצלחה הראשונה כולל.

נסמן ב p את הסיכוי להצלחה בניסויי בודד וב- q את הסיכוי לכישלון בניסוי בודד.

$$X \sim G(p)$$

פונקציית ההסתברות:

$$P(X=k)=pq^{k-1} \qquad k=1,2,...\infty$$

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, $P(X=n+k)/X>k)=P(X=n)$.

$$P(X>k)=q^k$$

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?

<u>תרגילים:</u>

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:

א. שידגום 3 מוצרים.

ב. שידגום 4 מוצרים.

ג. שידגום 5 מוצרים.

ד. שידגום יותר מ-7 מוצרים.

ה. שידגום לא פחות מ-8 מוצרים.

2. צילום שמבוצע במכון הרנטגן "X-RAY" יתקבל תקין בהסתברות של 0.9. אדם נכנס למכון כדי להצטלם. הוא ייצא מהמכון רק כאשר יש בידו תצלום תקין.

א. מה ההסתברות שיצטלם בסך הכול 3 פעמים?

ב. מה ההסתברות שהצטלם יותר מ-4 פעמים?

ג. מה התוחלת ומה השונות של מספר הצילומים שייבצע?

ד. 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".

א. מה ההסתברות להטיל את המטבע לכל היותר 10 פעמים?

ב. מה ההסתברות להטיל את המטבע לכל היותר 5 פעמים אם ידוע שהמטבע הוטל לפחות 3 פעמים?

ג. אם ידוע שבשתי ההטלות הראשונות התקבלה התוצאה "פלי" מה ההסתברות שהאדם הטיל את המטבע 7 פעמים?

ד. מה תוחלת מספר הפעמים שהתקבלה התוצאה "פלי"?

4. 30% מהמכוניות בארץ הן בצבע לבן. בכל יום נכנסות לחניון 10 מכוניות אקראיות.

א. מה ההסתברות שביום מסוים בדיוק מחצית מהמכוניות בחניון יהיו לבנות?

ב. 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 ששיחק את המשחק בדיוק 10 פעמים?

ד. מהי סטיית התקן של מספר הפעמים שישחק את המשחק?

6. במאפייה מייצרים עוגת גבינה ועוגת שוקולד שנארזות באריזות אטומות. 40% מהעוגות הן עוגות גבינה והיתר עוגות שוקולד. התווית על האריזה מודבקת בשלב מאוחר יותר של הייצור. אדם נכנס למפעל ובוחר באקראי עוגה.

א. מה ההסתברות שייאלץ לבחור 5 עוגות עד שקיבל עוגות שוקולד?

ב. אם הוא דגם פחות מ-7 עוגות עד שיקבל עוגת שוקולד, מה ההסתברות שבפועל הוא דגם יותר מ-4 עוגות?

ג. האדם דוגם עוגות עד אשר הוא מוצא עוגת שוקולד ידוע שעוגת גבינה עולה לייצרן 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 עוגת הגבינה שדגם האדם?

<u>פתרונות :</u>

<u>שאלה 2 :</u>

א. 0.009

ב. 0.0001

ג. תוחלת : 1.111

שונות :0.1234

ד. תוחלת : 44.4

שונות : 308.5

<u>שאלה 3 :</u>

א. 0.999

ב. 0.875

ג. 0.03125

ד. 1

<u>שאלה 4:</u>

א. 0.1029

ב. 9.72

<u>שאלה 5:</u>

א. 0.06

ב. 0.7176

ג. 0.0729

ד. 9.487משחקים

<u>שאלה 6 :</u>

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

פרק 14 - התפלגויות בדידות מיוחדות - התפלגות אחידה

רקע:

התפלגות זו הנה התפלגות שבה לכל תוצאה יש את אותה הסתברות.

הערכים המתקבלים בהתפלגות הם החל מ- a ועד b בקפיצות של אחד.

$$X \sim U\left(a,b\right)$$

פונקציית ההסתברות:

$$P(X=K)=\frac{1}{b-a+1}$$

K= a, a+1,..,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?

תרגילים :

1. במשחק הלוטו 45 כדורים ממוספרים מ- 1 ועד 45 .
נתבונן במשתנה X המספר של הכדור הראשון שנשלף על ידי המכונה.
   א. חשבו את $P(X=2)$
   ב. חשבו את $P(X\le 30)$
   ג. חשבו את $P(X>4 \mid X\le 10)$
   ד. חשבו את $P(X=k)$

2. קוסם מבקש לבחור מספר שלם אקראי בין 1 ל- 100. בהנחה שאין כאן מניפולציות של הקוסם.
   א. מהי התוחלת ומהי סטיית התקן של המספר שיבחר?
   ב. הקוסם ביקש משישה אנשים לבחור מספר :
      1. מה ההסתברות ששלושה מהם יבחרו מספר הגדול מ 80?
      2. מה התוחלת ומהי סטיית התקן של סכום המספרים שהאנשים בחרו?

3. יהי X התוצאה בהטלת קובייה.
   א. מהי ההתפלגות של X?
   ב. מה התוחלת של X?
   ג. קובייה הוטלה 4 פעמים . מה התוחלת ומה השונות של סכום התוצאות ב- 4 ההטלות?

4. בכד 10 כדורים שרק אחד צבע אדום. אדם מוציא כדור ללא החזרה עד אשר מתקבל הכדור האדום. מה התוחלת ומהי השונות של מספר הכדורים שהוצאו?

5. יש לבחור מספר אקראי בי 1 ל- 50 כולל.
   א. מה הסיכוי שהמספר 4 יבחר?
   ב. מה הסיכוי שהמספר שיבחר גדול מ- 20 ?
   ג. אם נבחר מספר גדול מ- 20 מה ההסתברות שהוא קטן מ- 28?

6. הוכח שאם $X\sim U\left(a,b\right)$ אז מתקיים ש : $E(X)=\dfrac{a+b}{2}$.

פתרונות :

שאלה 1 :

א. תשובה : $\frac{1}{45}$

ב.תשובה : $\frac{30}{45}$

ג. תשובה : 0.6

שאלה 2

א. תוחלת : 50.5

סטיית התקן : 28.87

ב. 1. תשובה : 0.08192

ב. 2 תוחלת : 303 סטיית תקן :70.71

שאלה 4 :

תוחלת 5.5

שונות :8.25

פרק 15 - 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 בממוצע אירועים קורים ביחידת זמן.

$$X \sim pois(\lambda)$$

התפלגות פואסונית חייבת להופיע כנתון בשאלה ולכן לא יהיה צורך לזהותה .

פונקציית ההסתברות של ההתפלגות הפואסונית נתונה:

$$P(X = K) = \frac{e^{-\lambda} \cdot \lambda^{K}}{K!}$$

$$K = 0, 1, 2, \ldots \infty$$

התוחלת והשונות של ההתפלגות:

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

תכונות מיוחדות של ההתפלגות:

- בהתפלגות הזו הפרמטר λ פורפורציונלי 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.

א. מה ההסתברות שבדקה כלשהי תתקבל פניה 1?

ב. מה ההסתברות שבשתי דקות יגיעו 12 פניות?

ג. מה ההסתברות שבדקה אחת תגיע פניה 1 ובשתי דקות שלאחר מכן 12 פניות?

ד. מה התוחלת וסטיית התקן של מספר הפניות בדקה?

תרגילים:

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 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 (הנח שבחודש יש 4 שבועות) יהיה בדיוק שבוע אחד בו יהיו 3 תאונות דרכים קטלניות?

4. לחנות AMPM 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 מסוים 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?

פתרונות :

שאלה 1:

א. 0.0337

ב. 0.9933

ג. 0.1246

ד. 5

שאלה 2:

א. 0.084

ב. 0.099

ג. 0.151

שאלה 3:

א. 4

ב. 0.407

שאלה 5 :

א. 0.2388

ב. 0.2196

ג. 0.6948

שאלה 6 :

א. 16.7

ב. 0.0708

פרק 16 - המשתנה המקרי הבדיד - שאלות מסכמות

תרגילים:

1. נתון ש:

$$X \sim B(4,\frac{1}{2})$$

$$Y \sim B(10,\frac{1}{4})$$

א. חשב את התוחלת וסטיית התקן של X.

ב. $W = 2X - 4$, חשב את התוחלת וסטיית התקן של W.

ג. $T = X + Y$, חשב את התוחלת של T. האם ניתן לדעת מה סטיית התקן של T?

2. ערן משחק בקזינו בשתי מכונות הימורים. משחק אחד בכל מכונה (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. בנה את פונקציית ההסתברות של X.

ב. חשב 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 "בערוץ 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.
   א. מה ההסתברות למכור בחנות זו מוצר מתוצרת חוץ?
   ב. יהי X - מספר המוצרים שימכרו בחנות זו מפתיחתה ביום א׳ בבוקר, עד (וכולל) שלראשונה יימכר מוצר מתוצרת הארץ. מהי פונקצית ההסתברות של X?
   ג. מהי תוחלת מס׳ המוצרים **מתוצרת חוץ** שימכרו, עד שלראשונה יימכר מוצר מתוצרת הארץ?
   ד. 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 ״הצבי״ יימכר לחו״ל בסיכוי של 0.6.
   הסרט ״לעולם לא״ יימכר לחו״ל בסיכויי 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.
   א. בנו את פונקציית ההסתברות של מספר הסרטים שיימכרו לחו״ל.
   ב. מהי התוחלת והשונות של מספר הסרטים שיימכרו?
   ג. 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.
   א. נבחרה שקית ונתון שבשקית פחות מ- 3 סוכריות אדומות. מה ההסתברות שבשקית סוכריה אדומה אחת?
   ב. 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 התכונן 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 $V(X) = 1$. חשב $E(X-5)^2$.

10. הסיכוי לעבור מבחן נהיגה הינו P. בוחרים באקראי ארבעה נבחנים . ההסתברות ששניים מהם יעברו את מבחן הנהיגה גבוה פי 8/3 מהסיכוי שכל הארבעה יעברו את המבחן.
א. חשבו את ערכו של P.

תלמיד ניגש לבחינה עד אשר הוא עובר אותה.

ב. 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. לאילו ערכי P המשחק כדאי?

פתרונות :

שאלה 1:

א. תוחלת : 2

סטיית תקן : 1

ב. תוחלת : 0

סטיית תקן : 2

ג. תוחלת : 4.5

סטיית תקן : לא ניתן

שאלה 2:

א. 0.03

ב. תוחלת : 0.15, שונות 0.1875

ג. 0.1328

שאלה 3:

א.

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

ב. תוחלת : 3

שונות : 2

ג. תוחלת : 1.5

שונות 0.5

שאלה 4 :

א. 0.9975

ב. 0.0172

ג. 1.0025

<u>שאלה 5:</u>

א. 0.375

ג. 0.6

ד. 0.282

<u>שאלה 6:</u>

ב. תוחלת : 1.5

שונות 0.61

ג. תוחלת : 0

סטיית תקן : 4.68

<u>שאלה 7:</u>

א. 0.4348

ב. 0.0923

<u>שאלה 8 :</u>

א. 2.013

ב. 0.5256

ג. תוחלת : 8

שונות : 1.6

ד. תוחלת : 10.5

שונות 3.475

<u>שאלה 9 :</u>

10

<u>שאלה 10:</u>

א. 0.6

ב. 0.0384

ג. 0.0256

ד. תוחלת: 0.67

שונות: 1.11

ה. 0.24

<u>שאלה 12:</u>

ב. $\frac{1-2p^2}{p}$

ג. $0 < p < \sqrt{\frac{1}{2}}$

פרק 17 - 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.

פונקציית התפלגות מצטברת:

$$\boxed{F(t)=p(X\le t)=\int_{-\infty}^{t} f(x)dx}$$

כמו כן :

$$\boxed{p(X>t)=1-F(t)} \qquad \boxed{p(a<X<b)=F(b)-F(a)}$$

תוחלת של משתנה רציף :

$$\boxed{E(X)=\int_{-\infty}^{\infty} X\cdot f(x)dx=\mu}$$

שונות של משתנה רציף:

$$\boxed{V(X)=\int_{-\infty}^{\infty} X^2\cdot f(x)dx-\mu^2=\sigma^2}$$

תוחלת של פונקציה של X:

$$E\left(g(x)\right)=\int_{-\infty}^{\infty} g(x)f(x)\,dx$$

אחוזונים :

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

$$\boxed{p(X\le x_p)=p}$$

ריענון מתמטי:

נוסחאות לחישוב שטחים:

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

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

משוואת קו ישר:

y=mx+n

m = שיפוע.

n = נקודת החיתוך עם ציר הy.

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

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

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

נוסחאות - אינטגרלים

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

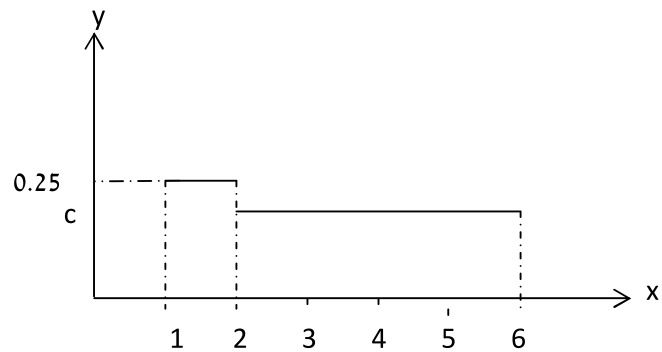
$$\int e^f\cdot f'dx=e^f+c \qquad \int\cos f\cdot f'dx=\sin(f)+c$$

$$\int\sin f\cdot f'dx=-\cos(f)+c \qquad \int\frac{f'}{\sqrt{f}}dx=2\sqrt{f}+c$$

$$\int\sqrt{f}\cdot f'dx=\frac{2}{3}f^{\frac{3}{2}}+c \qquad \int u\cdot v'dx=u\cdot v-\int u'\cdot v\,dx$$

תרגילים:

1. X הינו משתנה רציף עם פונקצית צפיפות כמוצג בשרטוטו:



א. מצא את ערכו של c.

ב. בנה את פונקצית ההתפלגות המצטברת.

ג. חשבו את ההסתברויות הבאות:

1. $P(x<4)$
2. $P(x>1.5)$
3. $P(1.5<x<5)$
4. $P(5<x<10)$

ד. מצא את החציון של המשתנה.

2. נתון משתנה מקרי רציף X שפונקצית הצפיפות שלו היא:

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

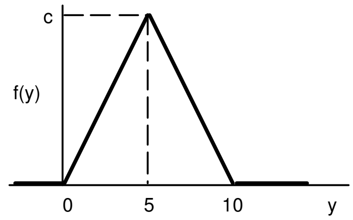
ידוע ש- $P(0<X<1)=1/4$.

א. מצאו במפורש את פונקצית הצפיפות של X.

ב. מצאו את החציון של X.

ג. מה הסיכוי ש- X קטן מ- 0.5 ?

3. נתונה פונקצית צפיפות של משתנה מקרי Y :



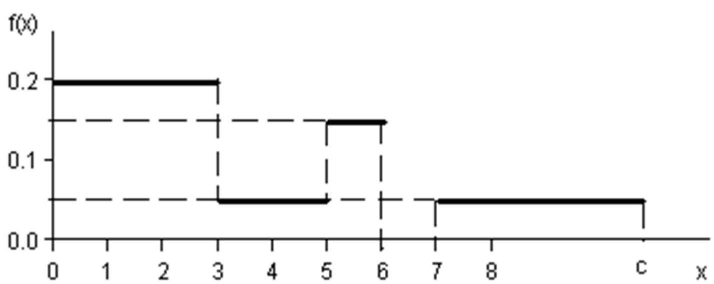
א. מצאו את c.

ב. מצאו את פונקצית ההתפלגות המצטברת של Y .

ג. חשבו את ההסתברויות : $P(Y = 7.0)$, $P(Y \leq 3.0)$, $P(7.5 \leq Y \leq 15.5)$ $P(\text{Y>4})$.

ד. מצאו את העשירון התחתון $y_{0.1}$, הרבעון התחתון $y_{0.25}$ והחציון של Y. הסיקו מהו העשירון עליון $y_{0.9}$.

4. נתונה פונקצית צפיפות של משתנה מקרי X :

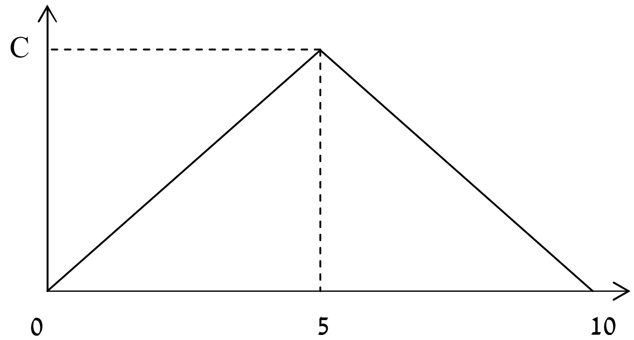


א. מצאו ערך c שעבורו תתקבל פונקצית צפיפות.

ב. מצאו את פונקצית ההתפלגות המצטברת .

ג. חשבו את ההסתברויות הבאות : $P(X \geq 4)$, $P(X \geq -2.0)$, $P(1.0 < X \leq 5.0)$

5. נתונה פונקצית הצפיפות הבאה :



א. מה ערכו של C?

ב. מצא אינטרוול (תחום) סימטרי סביב הערך 5 שהסיכוי ליפול בו הינו 0.5

6. נתונה פונקצית צפיפות $f(X)=\frac{2}{x}$ פונקציה זו מוגדרת מ-1 ועד K.

א. מצא את ערכו של K.

ב. בנה את פונקציית ההתפלגות המצטברת.

ג. חשב את הסיכוי ש X לפחות 1.5.

ד. מצא את העשירון התחתון של ההתפלגות.

ה. מה התוחלת של X?

7. נתונה פונקצית צפיפות הבאה : $f(X)=AX^2(10-X)$ $0<X<10$ A הינו קבוע חיובי.

א. מצא את A.

ב. חשב את $P(x>5 \mid x>2)$.

ג. מה התוחלת ומהי השונות של X ?

8. פונקצית הצפיפות של משתנה מקרי רציף X :

$$f(x) = 0.5 \cdot e^{2x}$$
$$-\infty \le X \le \ln(c)$$

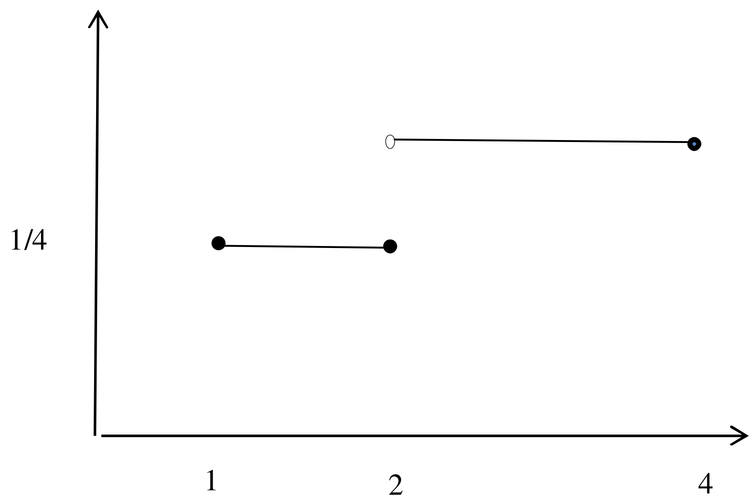
א. מצא את ערכו של c.

ב. מצא את פונקציית ההתפלגות המצטברת של ההתפלגות.

ג. חשב $P(X > 0)$.

ד. מהו הרבעון העליון של ההתפלגות?

9. נתונה פונקצית הצפיפות הבאה של משתנה מקרי X :



א. רשום את נוסחת פונקציית הצפיפות .

ב. בנה את פונקציית ההתפלגות המצטברת.

ג. מצא את החציון של ההתפלגות.

ד. חשב את התוחלת והשונות של המשתנה.

ה. חשב את $E(X^3)$

10. במפעל מייצרים מוצר A. זמן תהליך הייצור של המוצר בשעות הוא בעל פונקציית הצפיפות הבאה:

$$f(x)=6x(1-x) \qquad 0\le x\le 1$$

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 ההתפלגות המצטברת הבאה:

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

א. שרטט 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 ש90% מהלקוחות מחכים מתחתיו?

12. פונקצית הצפיפות של משתנה מקרי נתונה על ידי הנוסחה הבאה:

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

א. מצאו את b.

ב. חשבו את התוחלת של X.

ג. y הוא משתנה אינדיקטור המקבל את הערך 1 אם Xקטן מ-5 . מהי השונות של Y?

13. נתונה פונקציית הצפיפות הבאה:

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

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:

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

$a \le x \le b$

א. מצא את פונקציית ההתפלגות המצטברת.

ב. חשב את התוחלת והשונות של ההתפלגות.

ג. מצא את התוחלת של $\frac{1}{X}$.

פתרונות :

שאלה 1 :

א. $\frac{3}{16}$ ד. $3\frac{1}{3}$

שאלה 2 :

א. b=2 c=0.5

ב. 1.41

ג. 0.0625

שאלה 3 :

א. 0.2

ג. 0, 0.18, 0.125, 0.32

ד. העשירון התחתון : 2.24

הרבעון התחתון: 3.54

החציון : 5

העשירון העליון : 7.76

שאלה 4 :

א. 10

שאלה 5:

א. C=0.2

ב. 5 ± 1.46

שאלה 6 :

א. $e^{\frac{1}{2}}$

ג. 0.189

ד. 1.051

ה. 1.297

שאלה 7 :

א. 0.0012

ב.0.7067

ג. תוחלת : 6, שונות : 4

שאלה 8 :

א. 2

ג. 0.75

ד. 0.549

שאלה 9 :

ג. $2\frac{2}{3}$

ד. תוחלת : 2.625   שונות : 0.6927

ה. 23.4375

שאלה 10 :

א. $\frac{7}{27}$

ב. 0

ג. 3.704

שאלה 11 :

ב. 0.0498

ג. 0.6321

ד. 11.51

שאלה 12 :

א. $\frac{2}{3}$

ב. 5.22

ג. $\frac{2}{9}$

שאלה 13:

א. $\frac{1}{6}$

ג. 0.229

שאלה 14 :

ב. תוחלת :

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

השונות:

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

פרק 18 - 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 התרחשות המאורע הבא מתפלג מעריכית.

פונקציית הצפיפות של ההתפלגות היא :

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

פונקציית ההתפלגות המצטברת היא :

$F(t) = p(x \leq t) = 1 - e^{-\lambda t}$

התוחלת:

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

השונות:

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

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

דוגמה : (פתרון בהקלטה)

אורך חיי סוללה מתפלג מעריכית עם תוחלת של 8 שעות.

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

ב. מה סטיית התקן של אורך חיי הסוללה?

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

<u>תרגילים:</u>

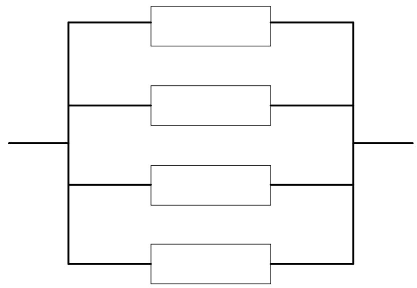
1. הזמן שלוקח במערכת עד שתקלה מתרחשת מתפלג מעריכית עם תוחלת של 0.5 שעה.
   א. מה הסתברות שהתקלה הבאה תתרחש תוך יותר מ-0.5 שעה?
   ב. מה ההסתברות שהתקלה הבאה תתרחש תוך פחות משעה?
   ג. מצא את הזמן החציוני להתרחשות תקלה במערכת.

2. הזמן שעובר בכביש מסוים עד להתרחשות תאונה מתפלג מעריכית עם תוחלת של 24 שעות.
   א. מהי סטית התקן של הזמן עד להתרחשות תאונה?
   ב. מה ההסתברות שהתאונה הבאה תתרחש תוך פחות מיממה?
   ג. 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.
   א. מה הסיכוי שעבודת סטודנט על המחשב תארך פחות מרבע שעה?
   ב. מה הסיכוי שעבודת סטודנט על המחשב תארך בין רבע שעה לחצי שעה?
   ג. אם סטודנט עובד על המחשב כבר יותר מ- 10 דקות, מה ההסתברות שמשך כל עבודתו יעלה על 30 דקות?
   ד. מהו הזמן שבסיכוי של 90% הסטודנט יעבוד פחות ממנו?

4. בממוצע מגיעים לחדר מיון 4 חולים בשעה בזרם פואסוני.
   א. שולה המזכירה הגיעה לחדר המיון. מה ההסתברות שזמן ההמתנה שלה לחולה הבא יהיה יותר מ- 20 דקות?
   ב. אם שולה המתינה יותר מרבע שעה לחולה הבא . מה ההסתברות שתמתין בסך הכל יותר מחצי שעה?
   ג. מה ההסתברות שבין החולה הראשון לשני יש להמתין יותר מרבע שעה ובין החולה השני לשלישי יש להמתין פחות מרבע שעה?

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 100שעות.

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?

<u>פתרונות:</u>

<u>שאלה 1 :</u>

א. 0.368

ב. 0.865

ג. 0.347

<u>שאלה 2:</u>

א. 24 שעות

ב. 0.632

ג. 0.135

<u>שאלה 3:</u>

א. 0.393

ב. 0.239

ג. 0.513

ד.69.08

<u>שאלה 4:</u>

א. 0.264

ב. 0.368

ג. 0.233

<u>שאלה 5:</u>

א. 0.8403

ב. A0.0588>K

פרק 19 - התפלגויות רציפות מיוחדות - התפלגות אחידה

רקע:

זו התפלגות שפונקציית הצפיפות שלה קבועה בין a לבין b.

X~U (a,b)

פונקציית הציפות :

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

$$a \le x \le b$$

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 : (הפתרון בהקלטה)

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

א. מה הסיכוי ש- X קטן מ-25?

ב. מה התוחלת והשונות של X?

תרגילים:

1. משך (בדקות) הפסקה בשיעור, X, מתפלג $U(13, 16)$.

א. מהי התוחלת ומהי סטית התקן של משך ההפסקה?

ב. מהי ההסתברות שהפסקה תמשך יותר מ- 15 דקות?

ג. מהי ההסתברות שמשך ההפסקה יסטה מהתוחלת בפחות מדקה?

2. רכבת מגיעה לתחנה בשעות היום כל עשר דקות . אדם הגיע לתחנה בזמן אקראי.

א. הסבר כיצד מתפלג זמן ההמתנה לרכבת?

ב. אם זמן ההמתנה לרכבת ארך יותר מ-5 דקות , מהי ההסתברות שבסך הכל האדם ימתין לרכבת פחות מ- 8 דקות?

ג. 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).

א. מה ההסתברות שמשקל הגלידה בגביע יהיה מעל 108 גרם?

ב. נתון שהגלידה בגביע עם משקל נמוך מ-107 גרם. מה ההסתברות שמשקל הגלידה יהיה מעל 105 גרם?

ג. מה העשירון העליון של משקל הגלידה בגביע?

פתרונות:

שאלה 1:

א. תוחלת: 14.5

שונות: 0.866

ב. 1/3

ג. 2/3

שאלה 2:

א. $X \sim U(0,10)$

ב. 0.6

ג. 10

שאלה 3:

א. 0.2

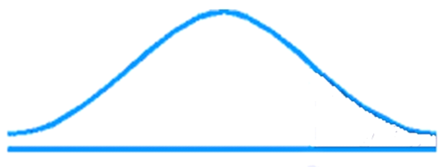
ב. $\frac{2}{7}$

ג. 109

פרק 20 - התפלגויות רציפות מיוחדות - התפלגות נורמלית

רקע:

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



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

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

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 הרלבנטים שמתחת לעקומה.
כדי לחשב שטחים אלה נמיר כל התפלגות נורמלית להתפלגות נורמלית סטנדרטית על ידי תהליך הנקרא תקנון.
התפלגות נורמלית סטנדרטית היא התפלגות נורמלית שהממוצע שלה הוא אפס וסטיית התקן היא אחת והיא תסומן באות Z.

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

תהליך התקנון מבוצע על ידי הנוסחה הבאה :

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

אחרי תקנון מקבלים ערך הנקרא ציון תקן.
ציון התקן משמעו בכמה סטיות תקן הערך סוטה מהממוצע.
לאחר חישוב ציון התקן של ערך מסוים נעזרים בטבלה של ההתפלגות הנורמלית הסטנדרטית לחישוב השטח הרצוי.

ובאופן כללי נתאר את הסכמה הבאה :

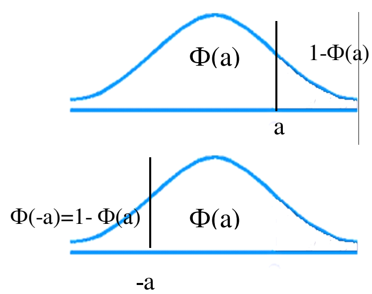
$$X \sim N(\mu,\sigma^2)$$

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

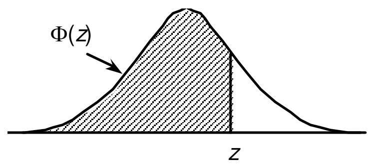
$$Z \sim N(0,1^2)$$

שימוש בטבלה

P



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
Φ(z)	0.90	0.95	0.975	0.99	0.995	0.999	0.9995	0.99995	0.999995

<u>דוגמה</u>: (הפתרון בהקלטה)

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

א. מה אחוז חפיסות השוקולד ששוקלות מתחת ל- 110 גרם?

ב. מה אחוז חפיסות השוקולד השוקלות מעל 110 גרם?

ג. מה אחוז חפיסות השוקולד השוקלות מתחת ל 92 גרם?

ד. מהו המשקל ש90% מהחפיסות בקו הייצור שוקלים פחות מהם?

תרגילים:

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

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

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

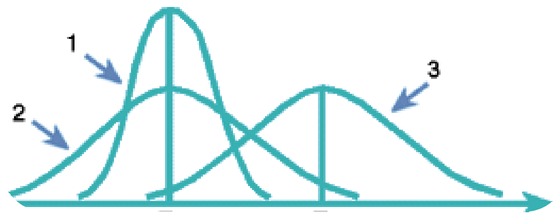
4. משקל תינוקות ביום היוולדם מתפלג נורמלית עם ממוצע של 3300 גרם וסטיית תקן 400 גרם.
   א. מצאו את העשירון העליון.
   ב. מצאו את האחוזון ה95.
   ג. מצאו את העשירון התחתון.

5. ציוני מבחן אינטיליגנציה מתפלג נורמלית עם ממוצע 100 ושונות 225 .
   א. מה העשירון העליון של הציונים במבחן האינטיליגנציה?
   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, נתון ש33% מהבקבוקים הם עם 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 המאון העליון של אורח 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 בינהן.



א.לאיזו התפלגות הממוצע הגבוה ביותר?

ב. במה מבין המדדים הבאים התפלגות 1 ו 2 זהות?

א. בעשירון העליון.

ב. בממוצע.

ג. בשונות.

ג. לאיזו התפלגות סטיית התקן הקטנה ביותר?

א. 1

ב. 2

ג. 3

ד. אין לדעת.

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

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

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

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

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

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 וסטיית 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?

פתרונות :

שאלה 1

א. 89.25%

ב. 2.28%

ג. 0

ד. 50%

שאלה 3

א. 26.43%

ב. 89.44%

ג. 39.44%

ד. 0.383

ה. 100%

שאלה 5

א. 119.2

ב. 80.8

ג. 112.6

ד. 87.4

שאלה 7

א. 500

ב. 100

ג. 0.3446

ד. 733

ה. 267

שאלה 8

א. 3

ב. בממוצע.

ג. 1

שאלה 9

א. 0.1587

ב. 0.0228

ג. 0.8563

ד. 0.3975

שאלה 10

א. 1925

ב. 0.2266

ג. 0.1587

שאלה 11

א. 0.1359

ב. 0.675

ג. 100

ד. 0.25

פרק 21- משתנה דו מימדי בדיד - פונקצית הסתברות משותפת

רקע:

התפלגות דו ממדית הינה התפלגות שדנה בשני משתנים.

נרצה כעת לבנות פונקציית הסתברות דו ממדית.

בפונקציה שכזו יש התפלגות של שני משתנים בו זמנית : X ו Y .

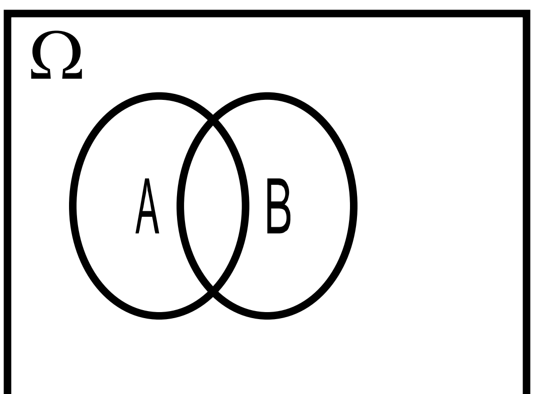
דוגמה:

תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ומבחן בסטטיסטיקה.
כמו כן נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8 והסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9.
הסיכוי לעבור את שני המבחנים הנו 0.75

יהי X- מספר הקורסים שהסטודנט עבר.

יהי Y- משתנה אינדיקטור המקבל את הערך אחד אם הסטודנט עבר את הבחינה בכלכלה ואפס אחרת.

בנו את <u>פונקציית ההסתברות המשותפת של X ו Y</u> .



נחשב את כל ההסתברויות המשותפות :

$$p(x=o,y=0)=0.05$$
$$p(x=o,y=1)=0$$
$$p(x=1,y=0)=0.15$$
$$p(x=1,y=1)=0.05$$
$$p(x=2,y=0)=0$$
$$p(x=2,y=1)=0.75$$

Y \ X	0	1	2
0	0.05	0.15	0
1	0	0.05	0.75

<u>שימו לב שסכום כל ההסתברויות בפונקציית ההסתברות המשותפת הוא 1.</u>

כעת נסכם את השורות ואת העמודות ונקבל את פונקציות הסתברות שוליות :

$Y \backslash X$	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 :

$$\boxed{p(x=k, y=l) = p(x=k)\cdot p(y=l)}$$

מספיק פעם אחת שהמשתנים אינם מקיימים תנאי זה אזי הם תלויים.

למשל , בדוגמה הזאת:

$$p(x=2, y=1) = 0.75 \neq p(x=2)\cdot p(y=1) = 0.75\cdot 0.8 = 0.6$$

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.

תרגילים :

1. אדם נכנס לקזינו עם 75 דולר . הוא ישחק במכונת מזל בה יש סיכוי של 03 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 מספר המשחקים שהאדם שיחק.
   א. בנו את פונקצית 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 פונקצית 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. השלם את ההסתברויות החסרות בטבלה.
   ב. האם X ו- Y תלויים ?
   ג. מצא 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 1/10. מהנדס הייצור בוחר באקראי חבילת מוצרים לשם בקורת איכות. יהיו: X– מספר המוצרים בחבילה, Y– מספר המוצרים הפגומים בחבילה.
   א. מה ההתפלגות של המשתנה Y בהינתן X הינו 3.
   ב. מה ההתפלגות של המשתנה Y בהינתן X הינו K כלשהו.
   ג. מהי תוחלת מספר המוצרים הפגומים בחבילות בנות 3 מוצרים? נמקו.
   ד. בנה את פונקצית ההסתברות המשותפת.

4. מתוך כד עם שלושה כדורים ממוספרים במספרים 2, 4, 8 שולפים באקראי שני כדורים **ללא החזרה**. נגדיר : X - המספר הקטן מבין השניים ; Y– המספר הגדול מבין השניים.

א. חשבו את ההתפלגות של (X, Y) .

ב. אם המספר המינימאלי שנבחר הוא 2 , מה הסיכוי שהמספר המקסימאלי 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.

ל-95% יש חשבון בלפחות אחד מהסניפים.

יהי X- מספר הסניפים בישוב אשר לתושב בוגר יש בהם חשבון.

יהי Y- משתנה אינדיקטור :

1- אם יש לתושב חשבון בסניף פועלים.

0- אחרת.

א. בנו את פונקציית ההסתברות המשותפת של Xו- Y.

ב. הוסיפו את פונקציית ההסתברות השולית.

ג. ידוע שלתושב בוגר חשבון בבנק פועלים, מה ההסתברות שיש לו חשבון בנק בסניף אחד בלבד?

פתרונות:

שאלה 1:

ב. 2.4

ג. התוחלת 1.348 השונות 0.575

שאלה 2:

ב. תלויים

ג. 0.125

שאלה 4:

ב. 0.5

ג. תוחלת 2

שאלה 5:

ג. 0.75

פרק 22 - 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.

-1 0 1

מקדם מתאם1- 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 לנארי מלא בו השיפוע a יהיה חיובי ואילו מתאם שלילי מלא אומר שקיים קשר לנארי מלא בו השיפוע a שלילי (מקדם מתאם 1-).

מתאם חיובי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט ואילו מתאם שלילי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לרדת אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט.

חישוב מקדם המתאם :

הנוסחה של מקדם המתאם היא : $\rho = \dfrac{\operatorname{cov}(x,y)}{\sigma_x \cdot \sigma_y}$

השונות המשותפת:

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

תכונות של השונות המשותפת :

1. $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$
2. $\operatorname{cov}(X,X) = Var(X)$

משתנים בלתי מתואמים :

משתנים בלתי מתואמים הם משתנים שמקדם המתאם שלהם אפס וכדי שדבר כזה יקרה השונות המשותפת צריכה להתאפס.
משתנים בלתי מתואמים הם משתנים שכלל אין בינם התאמה לינארית.
משתנים בלתי תלויים הם משתנים שאין בינם קשר ולכן הם גם בלתי מתואמים , אך משתנים בלתי מתואמים אינם בהכרח בלתי תלויים.

השפעת טרנספורמצייה לינארית על מקדם המתאם

$$\rho\left[(aX+b),(cY+d)\right]=\begin{cases}\rho(X,Y)\_\_\_if\ \_\_a\cdot c>0\\ -\rho(X,Y)\_\_if\ \_\_a\cdot c<0\end{cases}$$

כלומר , טרנספורמציה לינארית על שני משתנים לא משנה את עוצמת הקשר בינם 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 והסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9.
הסיכוי לעבור את שני המבחנים הנו 0.75
יהי X- מספר הקורסים שהסטודנט עבר.
יהי Y- משתנה אינדיקטור המקבל את הערך אחד אם הסטודנט עבר את הבחינה בכלכלה ואפס אחרת.
נחשב את מקדם המתאם :

$Y \backslash X$	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

2	1	0	x
0.75	0.20	0.05	P(x)

$$E(X)=\sum_i x_iP(x_i)=\mu=0*0.05+1*0.2+2*0.75=1.7$$

$$V(X)=\sum_i (x_i-\mu)^2P(x_i)=\sum_i x_i^2P(x_i)-\mu^2=0^2*0.05+1^2*0.2+2^2*0.75-1.7^2=0.31=\sigma^2$$

$$\sigma_x=\sqrt{V(X)}=\sqrt{0.31}=0.557$$

y	P_Y
0	0.2
1	0.8

$$E(y)=\sum_i y_iP(y_i)=0+0.8=0.8$$

$$V(y)=\sum_i (y_i-\mu_y)^2P(y_i)=\sum_i y_i^2P(y_i)-\mu_y{}^2=0+0.8-0.8^2=0.16=\sigma_y{}^2$$

$$\sigma_y=\sqrt{0.16}=0.4$$

$$E(xy)=0\cdot0\cdot0.05+0\cdot1\cdot0+1\cdot0\cdot0.15+1\cdot1\cdot0.05+2\cdot0\cdot0+2\cdot1\cdot0.75=1.55$$

$$\operatorname{cov}(x,y)=E(x\cdot y)-E(x)\cdot E(y)=1.55-1.7\cdot0.8=0.19$$

$$\rho=\frac{\operatorname{cov}(x,y)}{\sigma_x\cdot\sigma_y}=\frac{0.19}{0.557\cdot0.4}=0.853$$

כל קורס שהסטודנט מסיים מזכה אותו ב- 3 נקודות אקדמאיות.

מה יהיה מקדם המתאם בין נקודות הזכות שיצבור למשתנה Y? ?

תרגילים:

1. הסיכוי שסטודנט יעבור את מועד א בסטטיסטיקה הוא 0.8. אם הוא נכשל במועד א' הוא ניגש למועד ב' שם הסיכוי לעבור את המבחן מוערך להיות 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. ואז הסיכוי שלו לעבור את מועד ג' הוא 0.7.

נגדיר את X להיות מספר המבחנים אליהם ניגש הסטודנט.

נגדיר את Y להיות מספר הנבחנים שנכשל בהם.

א. בנו את פונקצית 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? הסבר ללא חישוב.

ה. חשבו את מקדם המתאם בין X לבין Y.

ו. האם המשתנים הם בלתי מתאומים?

2. מטילים 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 מתואמים?

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. מפזרים שלושה כדורים שונים בשלושה תאים.

נגדיר את המשתנים הבאים:

X= מספר הכדורים בתא הראשון.

Y= מספר הכדורים בתא השני.

א. בנו את פונקציית ההסתברות המשותפת.

ב. האם המשתנים בלתי מתאומים?

4. מטילים קובייה הוגנת פעמיים.

יהי: X= ההטלה הגדולה מבין שתי התוצאות

Y= מס' ההטלות בהן יצאה תוצאה זוגית.

א. מצא את פונקצית ההסתברות המשותפת של X ו-Y.

ב. חשבו את מקדם המתאם של X ו-Y.

ג. מצאו את ההתפלגות של Y בהינתן ש- X=2.

5. בבניין בן 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.

א. בנו את פונקצית ההסתברות המשותפת ואת פונקציות ההסתברות השולית.

ב. האם המשתנים מתואמים?

ג. מה מקדם המתאם בין X לבין Y?

ד. מה יהיה מקדם המתאם:

1. בין מספר הדירות המשופצות למספר הדירות הזוגיות שנדגמו.
2. בין מספר הדירות הזוגיות לדירות האי זוגיות שנדגמו.

ה. כל דירה משופצת עולה 2 מיליון שקלים, כל דירה לא משופצת עולה 1.5 מיליון שקלים. מה המתאם בין עלות הדירות שנדגמו למספר הדירות הזוגיות?

<u>פתרונות :</u>

<u>שאלה 1 :</u>

ג . 0.994

ה . 0.963

<u>שאלה 2 :</u>

ב. תלויים.

ג. מקדם המתאם : 0.5. מתואמים

ד. 0.25

ה. 0.5

<u>שאלה 3:</u>

ב. מתואמים

<u>שאלה 4 :</u>

ב. 0.252

<u>שאלה 5:</u>

ב. X ו-Y מתואמים.

ג. $\frac{2}{3}$

ד.1. $-\frac{2}{3}$

ד.2. (-1)

ה. $-\frac{2}{3}$

פרק 23 - המשתנה המקרי הדו ממדי - קומבינציות לנאריות

רקע:

תוחלת ושונות של סכום משתנים :

$$E(X+Y)=E(X)+E(Y)$$

$$V(X+Y)=V(X)+V(Y)+2\cdot COV(X,Y)$$

תוחלת ושונות של הפרש משתנים :

$$E(X-Y)=E(X)-E(Y)$$

$$V(X-Y)=V(X)+V(Y)-2\cdot COV(X,Y)$$

קומבינציות ליניאריות:

יוצרים משתנה חדש שהוא קומבינציה לינארית של שני משתנים אחרים :

$$W=(aX+b)+(cY+d)$$

$$COV\left[(aX+b),(cY+d)\right]=a\cdot c\cdot COV(X,Y)$$

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

$$V(W)=V\left((aX+b)+(cY+d)\right)=a^2V(X)+c^2V(Y)+2\cdot a\cdot c\cdot COV(X,Y)$$

דוגמה : (פתרון בהקלטה)

עבור שני משתנים מקריים נתון :

$\mu_X=80$
$\sigma_X=15$
$\mu_Y=70$
$\sigma_Y=20$
$COV(X,Y)=200$

- מצא את התוחלת והשונות של סכום המשתנים.
- מצא את התוחלת והשונות של Y-X.
- מצא את השונות ומה התוחלת של המשתנה $W=2X+3Y$

תרגילים:

1. נתונה פונקצית ההסתברות המשותפת הבאה:

Y\X	1	2	3	P(Y)
2		0.1	0.3	0.6
3	0.2		0.1	
P(X)				

א. השלם את ההסתברויות החסרות.

ב. האם המשתנים תלויים?

ג. האם המשתנים בלתי מתואמים?

ד. חשב את השונות המשותפת.

ה. חשב את התוחלת והשונות של סכום המשתנים.

ו. חשב 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 90 עם סטיית תקן 15. מקדם המתאם בין הציון הכמותי לציון המילולי הוא 0.8.

א. חשבו את השונות המשותפת בין הציון הכמותי לציון המילולי.

ב. חשבו את התוחלת והשונות של סכום הציונים בחלק הכמותי ובחלק המילולי.

ג. חשבו את התוחלת והשונות של הפרש הציונים בין החלק הכמותי לחלק המילולי.

ד. עלות הבחינה 2000 שקלים. הוחלט לזכות שקל עבור כל נקודה שנצברה בחלק המילולי ושני שקלים עבור כל נקודה שנצברה בחלק הכמותי. מהי התוחלת ומהי השונות של עלות הבחינה נטו (העלות לאחר הזיכוי)?

3. נתון: $Var(X-2Y)=2$.$Var(X+2Y)=3$. חשבו: $Cov(X,Y)$.

4. מטילים קובייה n פעמים.

נגדיר את המשתנים הבאים:

X=מספר הפעמים שהתקבלה התוצאה 6.

Y=מספר הפעמים שהתקבלה התוצאה 5

בטאו את השונות המשותפת באמצעות n.

פתרונות :

שאלה 1:

ב. תלויים

ג. מתואמים.

ד. ‎-0.1

ה. תוחלת : 4.4, שונות : 0.84

ו. תוחלת : ‎-0.4, שונות : 1.24

שאלה 2:

א. 240

ב. תוחלת: 190 שונות: 1105

ג. תוחלת: 10 שונות: 145

ד. תוחלת: 1710 שונות: 2785

שאלה 3:

‎-0.125

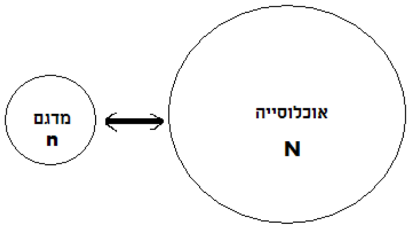
שאלה 4

$\frac{-n}{36}$

פרק 24 - 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 תהייה דגימה מקרית הכוונה לדגימה שבה לכל תצפית באוכלוסייה יש את אותו סיכויי להיכלל במדגם.

סטטיסטי – גודל המחושב על המדגם.

פרמטר – גודל המתאר את האוכלוסייה.

הסימונים לפרמטר וסטטיסטי הם שונים

למשל:

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

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?

<u>תרגילים :</u>

1. מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
   א. מי האוכלוסייה?
   ב. מה המשתנה?
   ג. מהם הפרמטרים?
   ד. מהו גודל המדגם?

2. להלן התפלגות מספר מקלטי הטלוויזיה למשפחה בישוב "העוגן".
   נגדיר את x להיות מספר המקלטים של משפחה אקראית.
   מתכננים לדגום מאוכלוסיה זו 4 משפחות ולהתבונן בממוצע מספר מקלטי הטלוויזיה במדגם.

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

   א. מיהי האוכלוסייה ומהו המשתנה הנחקר?
   ב. מהו הסטטיסטי שיילקח מהמדגם ומה סימונו?

3. נתון כי 20% מהשכירים במדינה הם אקדמאיים. נבחרו באקראי 10 שכירים באותה אוכלוסייה ומתכננים לפרסם את מספר האקדמאיים שנדגמו.
   א. מהי האוכלוסייה ?
   ב. מה המשתנה באוכלוסייה?
   ג. מהם הפרמטרים?
   ד. מהו הסטטיסטי?

פרק 25 - התפלגות הדגימה

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.

א. 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. מי אוכלוסיית המחקר? מהו המשתנה הנחקר?

ב. מהם הפרמטרים של האוכלוסייה?

ג. מה התוחלת ומהי סטית התקן של ממוצע המדגם?

<u>ב. דגימה מהתפלגות נורמאלית</u>

אם נדגום מתוך אוכלוסייה שהמשתנה בה מתפלג נורמאלית עם ממוצע μ ושונות σ^2 ממוצע המדגם גם יתפלג נורמאלית:

$$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$$

$$Z_{\bar{x}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

<u>דוגמה:</u> (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?

<u>ג. משפט הגבול המרכזי</u>

אם אוכלוסייה מתפלגת כלשהו עם ממוצע μ ושונות σ^2 אזי עבור מדגם מספיק גדול ($n \geq 30$)

ממוצע המדגם מתפלג בקירוב נורמאלית $\bar{x} \sim\succ N(\mu, \frac{\sigma^2}{n})$.

<u>דוגמה:</u> (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?

תרגילים :

1. מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.

א. מי האוכלוסייה?

ב. מה המשתנה?

ג. מהם הפרמטרים?

ד. מהו גודל המדגם?

ה. מהו תוחלת ממוצע המדגם?

ו. מהי טעות התקן?

2. להלן התפלגות מספר מקלטי הטלוויזיה למשפחה בישוב מסוים :

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

נגדיר את x להיות מספר המקלטים של משפחה אקראית.

א. בנו את פונקצית ההסתברות של x.

ב. חשבו את התוחלת, השונות וסטיית התקן של x.

ג. 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
א. מה ההסתברות שתינוק אקראי בעת הלידה ישקול פחות מ-3800 גרם?

נתון כי ביום מסוים נולדו 4 תינוקות.
ב. מה ההסתברות שהמשקל הממוצע שלהם יעלה על 4 ק"ג ?
ג. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-2.5 ק"ג?
ד. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת בלא יותר מ-50 גרם?
ה. הסבירו ללא חישוב כיצד התשובה לסעיף הקודם הייתה משתנה אם היה מדובר על יותר מ-4 תינוקות?

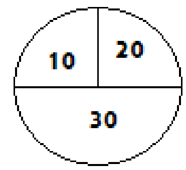
5. הגובה של המתגייסים לצה"ל מתפלג נורמאלית עם תוחלת של 175 ס"מ וסטיית תקן של 10 ס"מ. ביום מסוים התגייסו 16 חיילים.
א. מה ההסתברות שהגובה הממוצע שלהם יהיה לפחות 190 ס"מ?
ב. מה ההסתברות שהגובה הממוצע שלהם יהיה בדיוק 180 ס"מ?
ג. מה ההסתברות שהגובה הממוצע שלהם יסטה מתחולת הגבהים בפחות מ-5 ס"מ?
ד. מהו הגובה 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 פתרון הניחו שזמן הנסיעה לעבודה מתפלג נורמאלית.
א. מה ההסתברות שבמשך שבוע משך הנסיעה הממוצע יהיה מעל 33 דקות?
ב. מהו הזמן שבהסתברות של 90% ממוצע משך הנסיעה השבועי יהיה 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 מ755 סמ"ק?
ג. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה לפחות 755 סמ"ק?
ד. בקבוקיי 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 המסי' הבאים כמוראה בשרטוט :



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.
א. בנו את פונקצית 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. מטילים קובייה 50 פעמים בכל פעם מתבוננים בתוצאה של הקובייה. מה ההסתברות שהממוצע של התוצאות יהיה לפחות 3.72 ב- 50 ההטלות?

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. העזר במשפט הגבול המרכזי.

13. נתון משתנה מקרי בדיד בעל פונקצית ההסתברות הבאה :

2	4	6	8	X
¼	¼	¼	¼	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 בממוצע המדגם : $\bar{X}$

לכן $P(\bar{X} > \mu)$ יהיה : (בחר 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 ש :
(בחר בתשובה הנכונה)

א. $X \sim N(\mu,\frac{\sigma^2}{200})$

ב. $\mu \sim N(\mu,\frac{\sigma^2}{200})$

ג. $\bar{X} \sim N(\mu,\sigma^2)$

ד. $\bar{X} \sim N(\mu,\frac{\sigma^2}{200})$

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 :

(בחר 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 לגבי ממוצע המדגם :

$$E(\bar{x}) = \mu$$

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

פתרונות:

שאלה 2

א.

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

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

ג. $\mu_{\bar{x}} = 2.05 \ \sigma_{\bar{x}}^2 = 0.2369$

$\sigma(\overline{X}) = 0.486$

שאלה 3

$\mu_{\bar{x}} = 3.5$

$\sigma(\overline{X}) = 1.21$

שאלה 4

א. 0.8413

ב. 0.0013

ג. 0

ד. 0.1974

שאלה 6

א. 0.0465

ב. 27.71

ג. 0.2628

שאלה 7

א. 0

ב. 0.1587

ג. 0.1587

ד. 0.5

<u>שאלה 8</u>

א. 0.5468

ב. 0.6826

<u>שאלה 9</u>

א.

	10	20	30
P(x)	0.25	0.25	0.5

ב. התוחלת: 22.5

השונות: 68.75

ג. התוחלת: 22.5

השונות: 13.75

ד. 0.8997

<u>שאלה 10</u>

0.0475

<u>שאלה 11</u>

0.1814

<u>שאלה 12</u>

א. 0.9772

ב. 0.0228

ג. 271

<u>שאלה 14</u>

התשובה ב

<u>שאלה 15</u>

התשובה ד

<u>שאלה 16</u>

התשובה ג

<u>שאלה 17</u>

א. 2.429

ב. 0.25

התפלגות סכום תצפיות המדגם 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:

א. התוחלת והשונות של סכום התצפיות:

$$E(T)=n\mu$$
$$V(T)=n\sigma^2$$

ב. דגימה מתוך התפלגות נורמלית:

אם $X\sim N(\mu,\sigma^2)$ אזי $T\sim N(n\mu,n\sigma^2)$

$$Z=\frac{T-n\mu}{\sqrt{n\sigma^2}}$$

ג. משפט הגבול המרכזי :

אם x מתפלג כלשהו וידוע $E(X)=\mu$, $V(X)=\sigma^2$

אזי עבור מדגם מספיק גדול (לפחות 30)

$$T\ \sim\succ\ N(n\mu,n\sigma^2)$$

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.

א. מה התוחלת וסטיית התקן של סך המשכורות שיופקדו לסניף הבנק על ידי העובדים הללו?

ב. מה ההסתברות שלסניף יופקד פחות מ-780 אלף ₪ ע"י אותם עובדים? (0.1587)

תרגילים:

1. המשקל באוכלוסייה מסוימת מתפלג נורמאלית עם תוחלת של 60 ק״ג וסטיית תקן של 10 ק״ג.

   א. מה הסיכוי שאדם אקראי מהאוכלוסייה ישקול מתחת ל-65 ק״ג?

   ב. מה הסיכוי שהמשקל הממוצע של 4 אנשים אקראיים יהיה מתחת ל-65 ק״ג?

   ג. מה הסיכוי שהמשקל הכולל של 4 אנשים אקראיים יהיה מתחת ל-240 ק״ג?

2. נפח יין בבקבוק מתפלג נורמאלית עם תוחלת של 750 מ״ל וסטיית תקן של 20 מ״ל. אדם קנה מארז של 4 בקבוקי יין.

   א. מהי התוחלת ומהי סטיית התקן של נפח היין במארז?

   ב. את היין שבמארז האדם מזג לכלי שקיבולתו 3.1 ליטר. מה ההסתברות שהיין יגלוש מהכלי?

   ג. אם לא היה נתון שנפח היין מתפלג נורמאלית. האם התשובה לסעיף א הייתה משתנה? האם התשובה לסעיף ב הייתה משתנה?

3. בספר כלשהו 500 עמודים. קצב הקריאה הממוצע הוא עמוד אחד ב 4 דקות עם סטיית תקן של 1 דקות.

   א. מה ההסתברות לסיים את הפרק הראשון (40 עמודים) תוך שעתיים וחצי?

   ב. מהו האחוזון ה-95 לזמן סיום קריאת הספר?

4. במגדל נבנו 40 יחידות דיור. כמו כן נבנו 135 מקומות חנייה לבניין. להלן פונקצית ההסתברות של מספר המכוניות ליחידת דיור:

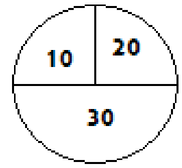
x	1	2	3	4	5
$P(X=x)$	0.1	0.2	0.3	0.25	0.15

נניח שמספר המכוניות ליחידת דיור בלתי תליות זו בזו ועם אותה פונקצית הסתברות לכל יחידת דיור (אין צורך בתיקון רציפות).

א. מהי ההסתברות שיהיה מקום בחניון המגדל לכל מכוניות הבניין ?

ב. בהינתן ויש מקום במגדל לכל המכוניות , מה הסיכוי שבפועל מספר המכוניות נמוך מ-130?

5. בקזינו ישנה רולטה עליה מסומנים המספרים הבאים :



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.

א. אם האדם משחק את המשחק 50 פעמים מה ההסתברות שבסך הכול יזכה בסכום של 1050 שקלים ומעלה?

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

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

חשבו את הסיכוי $P(\sum_i X_i \ge 115)$.

7. אורך חיי סוללה בשעות הוא בעל פונקצית הצפיפות הבאה :

$$f(x)=2x \quad 0<x<1$$

ברגע שסוללה מתרוקנת מחליפים אותה במידית בסוללה אחרת. כמה סוללות יש להחזיק במלאי אם רוצים שבסיכוי של 90% לפחות המלאי יספיק עבור 35 שעות לפחות?

פתרונות:

שאלה 1

א. 0.6915

ב. 0.8413

ג. 0.5

שאלה 2

א. תוחלת 3000 מ"ל וסטיית תקן 40 מ"ל

ב. 0.0062.

שאלה 4

א. 0.883

שאלה 5

א. 0.8997

ב. תוחלת : 1.111 שונות 0.1239

שאלה 7

56

התפלגות מספר ההצלחות במדגם - הקרוב הנורמלי להתפלגות הבינומית

רקע:

תזכורת על התפלגות בינומית

בפרק זה נדון על התפלגות מספר ההצלחות במדגם אקראי (תצפיות בלתי תלויות זו בזו).

מספר ההצלחות במדגם נסמן ב –Y.

מחלקים כל תצפית במדגם להצלחה או כישלון.

כעת מה שמשתנה מתצפית לתצפית הוא משתנה דיכוטומי (משתנה שיש לו שני ערכים).

תצפית

הצלחה כישלון

הסיכוי להצלחה יסומן עם הפרמטר p וכישלון יסומן ע"י הפרמטר $q=1-p$.

מבצעים מדגם אקראי בגודל n.

$$Y \sim B(n,p)$$

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

תוחלת : $E(y)=np$

שונות : $V(y)=npq$

<u>קירוב נורמלי עבור התפלגות בינומית</u>

אם לפנינו התפלגות בינומית : $Y \sim B(n,p)$ ומתקיים ש :

1. $n \cdot p \geq 5$

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

אז :
$$y \sim\succ N(np, npq)$$
$$Z_y = \frac{y-np}{\sqrt{npq}}$$

<u>תיקון רציפות:</u>

כאשר משתמשים בקירוב הנורמלי להתפלגות הבינומית יש לבצע תיקון רציפות .

הסיבה שעוברים כאן מהתפלגות בדידה להתפלגות נורמלית שהיא התפלגות רציפה.

על פי הכללים הבאים :

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

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

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

הערות:

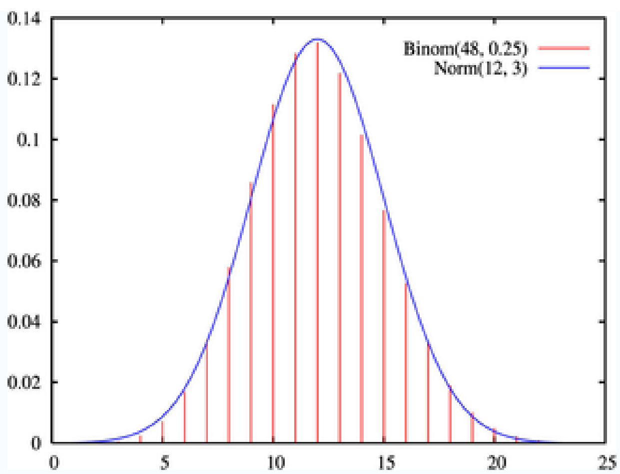
- התנאים למעבר מבינומי לנורמלי הם נזילים, כלומר משתנים ממרצה אחד לשני. התנאי שהצגתי כאן הוא הפופולרי ביותר:
  1. $n \cdot p \geq 5$
  2. $n \cdot (1-p) \geq 5$
- ישנם מרצים שנותנים את התנאי המחמיר הבא:
  1. $n \cdot p \geq 10$
  2. $n \cdot (1-p) \geq 10$
- וישנם מרצים שפשוט התנאי שהם נותנים הוא: $(n \geq 30)$.
- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור מהתפלגות בינומית לנורמלית.
- הערה נוספת היא לגבי תיקון רציפות. ישנם מרצים שלא מחייבים לבצע תיקון רציפות שהמדגמים גדולים (בדרך כלל מעל 100 תצפיות) אני בפתרונות שאציג תמיד אבצע תיקון רציפות במעבר מבינומי לנורמלי כיוון שכך הפתרון יהיה יותר מדויק (בכל מקרה שהמדגמים גדולים העניין זניח).

דוגמה: (הפתרון בהקלטה)

נתון שבקרב אוכלוסיית הנוער 25% זקוקים למשקפיים. נדגמו באקראי 48 בני נוער.

א. מה הסיכוי שבדיוק 14 מתוכם יהיו זקוקים למשקפיים?

ב. מה הסיכוי שלכל היותר 13 מתוכם זקוקים למשקפיים?



תרגילים:

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

א.מה ההסתברות ששלושה מהם אקדמאים?

ב. מה ההסתברות שלכל היותר אחד מהם אקדמאי?

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

2. במפעל 10% מהמוצרים פגומים. נלקחו 100 מוצרים באקראי מקו הייצור.

א.מה ההסתברות שנדגמו לפחות 6 מוצרים פגומים?

ב. מה ההסתברות שמספר המוצרים הפגומים יהיה לכל היותר 11 במדגם?

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

4. מטילים מטבע 50 פעמים.
א. מה ההסתברות לקבל לכל היותר 30 עצים?
ב. מה ההסתברות לקבל 28 עצים לפי התפלגות הבינומית ולפי הקירוב הנורמאלי?

5. במטוס מקום ל-400 נוסעים. נרשמו לטיסה 430 אנשים (overbooking). מנתונים סטטיסטיים ידוע שהסיכוי שאדם שנרשם לטיסה אכן יגיע הוא 0.9.
א. מה ההסתברות שלא יהיו מקומות ישיבה לכל האנשים שהגיעו לטיסה?
ב. מה צריך להיות גודל המטוס כדי שבסיכוי שלפחות 95% המטוס יספיק לכמות הנרשמים?

6. מפעל לייצור ארטיקים טוען ש הסיכוי שארטיק שהוא מייצר יהיה פגום הוא 0.01 . מוכר הזמין 1000 ארטיקים מהמפעל . מה ההסתברות שהמוכר יקבל לפחות 980 ארטיקים תקינים אם טענת המפעל מוצדקת ?

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

פתרונות:

<u>שאלה 1</u>

א. 0.201

ב. 0.3758

ג. התוחלת : 2, סטיית התקן : 1.2649

<u>שאלה 2</u>

א. 0.9332

ב. 0.6915

<u>שאלה 3</u>

0.1611

<u>שאלה 4</u>

א. 0.9406

<u>שאלה 5</u>

א. 0.015

<u>שאלה 6</u>

0.8544

<u>שאלה 7</u>

0.8643

פרק 26 - מושגים בסיסיים באמידה

רקע:

כזכור מהמפגש הקודם פרמטר הוא גודל המתאר את האוכלוסייה או התפלגות מסוימת.

כמו ממוצע הגבהים בקרב מתגייסים לצה"ל- μ .

כמו פרופורציית התומכים בממשלה בקרב אזרחי המדינה - p .

בדרך כלל הפרמטרים הם גדלים שאינם ידועים באמת , ולכן מבצעים מדגמים במטרה לאמוד אותם. אין אפשרות לחשב אותם הניסיון הוא בלהעריך כמה הם שווים ככל שניתן.

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

דוגמה: (פתרון בהקלטה)

בכנסת ה- 19 קיבלה מפלגת העבודה 15 מנדטים. בערוץ 10 ברגע סגירת הקלפיות העריכו את מספר המנדטים של המפלגה להיות 17 מנדטים וזאת על סמך תוצאות מדגם של הערוץ.

מה הפרמטר בדוגמה זו?

מהי טעות האמידה של ערוץ 10?

- $\hat{\theta}$ יהיה אומד חסר הטיה ל θ אם התוחלת של $\hat{\theta}$ תהיה שווה ל θ : $E(\hat{\theta})=\theta$
- טעות התקן של אומד היא סטיית התקן שלו , כלומר : $\sigma(\hat{\theta})=S.E$

להלן פרמטרים מרכזיים והאומדים שלהם :

ממוצע האוכלוסייה : μ

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

$E(\bar{x}) = \mu$ לכן $\bar{x}$ הינו אומר חסר הטיה ל μ .

כמו כן טעות תקן : $\sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}} = SE$

פרופורציה באוכלוסייה : p

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

$E(\hat{p}) = p$ לכן $\hat{p}$ הינו אומר חסר הטיה ל p .

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

שונות האוכלוסייה : σ^2

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

$E(S^2) = \sigma^2$ ולכן S^2 הינו אומד חסר הטיה ל σ^2 .

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$$

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

דוגמה: (פתרון בהקלטה)

נדגמו 10 משפחות בתל אביב ונבדק עבור כל משפחה מספר הילדים שלה. להלן התוצאות שהתקבלו:

2,1,3,2,1,4,5,2,1,3

אמדו באמצעות אומדים חסרי הטיה את הפרמטרים הבאים:

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

<u>תרגילים:</u>

1. מתוך 500 טירונים נמצאו 120 בעלי שברי הליכה. נתון שהסיכוי שטירון יהיה עם שבר הליכה הוא 0.25.
   א. מהי האוכלוסייה המוצגת בשאלה ? מהם הפרמטרים שלה?
   ב. מהי טעות התקן של האומד כשהמדגם בגודל 500?
   ג. מהו האומדן לפרמטר?
   ד. מהי טעות האמידה?

2. לפי נתוני היצרן מקרר צורך בממוצע 2400 וואט לשעה עם סטיית תקן של 500 וואט לשעה . במדגם של 25 מקררים של היצרן התקבל ממוצע של 2342 וואט לשעה.
   א.מהי האוכלוסייה המוצגת בשאלה ? מהם הפרמטרים שלה?
   ב.מהי טעות התקן של האומד?
   ג. מהו האומדן לפרמטר?
   ד. מהי טעות האמידה?

3. נדגמו עשרה מתגייסים לצה"ל. גובהם נמדד בס"מ. להלן התוצאות שהתקבלו :

   168, 184, 192, 171, 180, 177, 187, 168, 177 ו-175.

   א. מצא אומדן חסר הטיה לגובה הממוצע של מתגייסי צה"ל.

   ב. מצא אומדן חסר הטיה לשונות הגבהים של מתגייסי צה"ל.

   ג. מצא אומדן חסר הטיה לפרופורציות המתגייסים בגובה של לפחות 180 ס"מ.

4. נדגמו 20 שכירים באקראי. עבור כל שכיר נמדד השכר באלפי שקלים. להלן התוצאות שהתקבלו :

   $\sum_{i=1}^{20} X_i = 162$ $\sum_{i=1}^{20} X_i^2 = 1502.2$.

   א. אמדו את השכר הממוצע של השכירים במשק.

   ב. אמדו את סטיית התקן של שכר השכירים במשק.

5. במטרה לאמוד את ממוצע האוכלוסייה. דגמו תצפיות בלתי תלויות מהאוכלוסייה וחישבו את הממוצע שלהם. מהי טעות התקן?
   א. סטיית התקן של האוכלוסייה.
   ב. סטיית התקן של ממוצע האוכלוסייה.
   ג. סטיית התקן של המדגם.
   ד. סטיית התקן של ממוצע המדגם.

6. משקל הממוצע של אוכלוסייה מסוימת הוא 75 ק״ג עם שונות של 25 . אם יבחרו כל המדגמים האפשריים בגודל 10 מאוכלוסייה זו סטיית התקן של ממוצעי המדגמים תהייה:
   א.3
   ב.2.5
   ג. 1.581
   ד.אין מספיק נתונים לדעת.

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

8. $X_1, X_2, \ldots\ldots, X_{16}$ מדגם מקרי מתוך אוכלוסייה בעלת ממוצע μ לא ידוע ושונות $\sigma^2 = 64$. טעות התקן של האומד ל- μ היא:
   א. 16
   ב. 8
   ג. 4
   ד. 2

9. מהו אומד חסר הטיה?

א. אומד שערכו שווה לממוצע התפלגות הדגימה שלו.

ב. אומד שערכו שווה לערך הפרמטר באוכלוסייה.

ג. אומד שממוצע התפלגות הדגימה שלו שווה לערך הפרמטר באוכלוסייה.

ד. אומד שהסיכוי שערכו יהיה גבוה מערך הפרמטר באוכלוסייה שווה לסיכוי שיהיה נמוך ממנו.

פתרונות:

שאלה 3

א. 177.9
ב. 64.1
ג. 0.4

שאלה 4

א. 8.1
ב. 3.16

שאלה 5

התשובה היא ד.

שאלה 6

התשובה היא ג.

שאלה 7

התשובה היא ד.

שאלה 8

התשובה היא ד.

שאלה 9

התשובה היא ג.

פרק 27 - רווח סמך לתוחלת (ממוצע האוכלוסייה)

רווח סמך כששונות האוכלוסייה ידועה

רקע:

ממוצע המדגם הוא אומד לממוצע האוכלוסייה , אך לא באמת ניתן להבין ממנו על גודלו של ממוצע האוכלוסייה. ההסתברות שממוצע המדגם יהיה בדיוק כמו הממוצע האמתי הוא אפסי. מה שנהוג לעשות כדי לאמוד את ממוצע האוכלוסייה זה לבנות רווח סמך .

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

$1-\alpha$: נקרא רמת בטחון או רמת סמך.

כך ש : $P(A \le \mu \le B) = 1-\alpha$

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%.

מהי אוכלוסיית המחקר?

מה המשתנה באוכלוסייה?

מה הפרמטר שהחוקר רצה לאמוד?

מהו רווח הסמך?

מה אורך רווח הסמך?

מהי רמת הביטחון של רווח הסמך?

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

הפרמטר שנרצה לאמוד : μ

האומד נקודתי : $\bar{x}$

התנאים לבניית רווח הסמך :

1 $X \sim N$ או $n \geq 30$

2 σ^2 (שונות האוכלוסייה) ידועה

הנוסחה 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.

<u>שגיאת האמידה המקסימלית:</u>

$$\varepsilon = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

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

<u>דוגמה</u> : (פתרון בהקלטה)

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

<u>קשרים מתמטיים ברווח הסמך:</u>

- 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 $z_{1-\alpha/2}$ יותר גבוה, ורווח הסמך יותר ארוך.

תרגילים :

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.
   א. מי האוכלוסייה במחקר?
   ב. מה המשתנה הנחקר?
   ג. מה הפרמטר שאותו רוצים לאמוד?
   ד. מה רווח הסמך לפרמטר?
   ה. מהי רמת הסמך לפרמטר?
   ו. מה אורך רווח הסמך?
   ז. 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. בנה את רווח הסמך.

3. מעוניינים לאמוד את ממוצע אורך החיים של מכשיר. מנתוני היצרן ידוע שאורך החיים מתפלג נורמאלית עם סטיית תקן של 20 שעות. נדגמו 25 מכשירים ונמצא כי ממוצע אורך החיים שלהם היה 230 שעות.
   א. בנו רווח סמך ברמת סמך של 90% לאורך החיים הממוצע של מכשיר.
   ב. בנו רווח סמך ברמת סמך של 95% לאורך החיים הממוצע של מכשיר.
   ג. הסבר כיצד ומדוע השתנה רווח הסמך.

4. דגמו 200 עובדים מהמשק הישראלי. השכר הממוצע שלהם היה 9700 ₪. נניח שסטיית התקן של השכר במשק היא 3000 ₪.
   א. בנו רווח סמך ברמת סמך של % 95 לתוחלת השכר במשק.
   ב. מה ניתן לומר בביטחון של 95% על הסטייה המרבית בין ממוצע המדגם לתוחלת השכר?
   ג. מה היה צריך להיות גודל המדגם אם הינו רוצים להקטין את רווח הסמך ב50%?
   ד. 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 (99,105). 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%, מי מהמשפטים הבאים אינו 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. $\overline{X}=6$
   ג. $\overline{X}=51$
   ד. אורך רווח הסמך הינו 3.

10. איזה מהגורמים הבאים אינו משפיע על גודלו של רווח בר סמך, כאשר שונות האוכלוסייה ידועה? (בחר 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%. בנה 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-בטחון של 0.95 אם במדגם של 45 תצפיות התקבל: $\bar{x} = 74$.

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

פתרונות :

שאלה 2

$4920.6< \mu <4979.4$

שאלה 3

א. $223.42< \mu <236.58$

ב. $222.16< \mu <237.84$

שאלה 5

א. 102

ב. 3

ג. 0.9544

שאלה 6

א. $3.58< \mu <4.42$

ב. יקטן פי 2

ג. גדל

שאלה 7

א. 87

ב. 5

ג. 0.9544

שאלה 8

א. 139

ב. $21< \mu <25$

שאלה 9

התשובה היא : ב

שאלה 10

התשובה היא : ג

שאלה 11

התשובה היא : ד

קביעת גודל מדגם באמידת תוחלת עם שונות אוכלוסייה ידועה

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 בנוסחה הבאה :

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

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 האמתי לא יעלה על 0.5 ק״ג? (תשובה :87)

תרגילים:

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 $-\mu$
ברמת ביטחון של 0.95 כך שהאורך של הרווח יהיה 0.5σ . מהו גודל המדגם הנדרש?

<u>פתרונות :</u>

<u>שאלה 1</u>

780

<u>שאלה 2</u>

א. 139

ב. הדבר יקטין את ε פי 2.

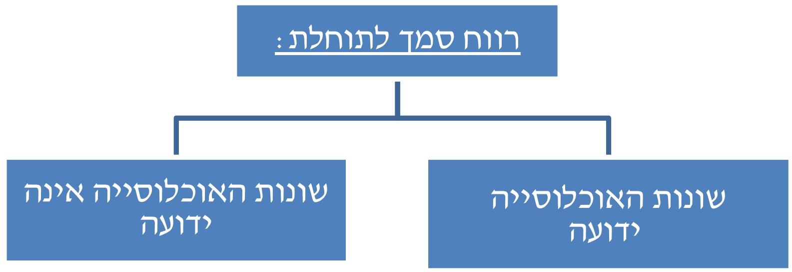
<u>שאלה 3</u>

$n = 62$

<u>רווח סמך לתוחלת (ממוצע האוכלוסייה) כששונות האוכלוסייה אינה ידועה</u>

<u>רקע:</u>

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



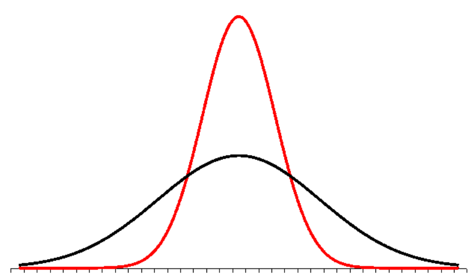
בפרק זה נעסוק במקרה ששונות האוכלוסייה(σ^2) אינה ידועה לנו .מקרה יותר פרקטי .

<u>התנאי :</u> $X \sim N$ או שהמדגם גדול

<u>רווח סמך :</u> $\overline{X} \pm t_{1-\alpha/2}^{(n-1)} \cdot \frac{S}{\sqrt{n}}$

<u>האומד לשונות</u> : $S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n-1}$

<u>התפלגות T :</u>



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.

<u>דוגמה :</u> (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'.

<u>פתרון :</u>

$4.39 < \mu < 5.51$

<u>תרגילים:</u>

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: 27,34,32,40,30.
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? (בחר בתשובה הנכונה)

א. סטטיסטיקאי א

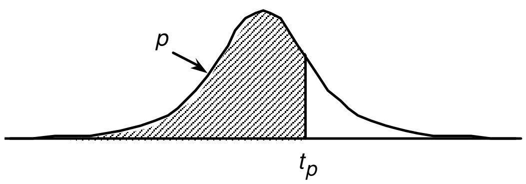
ב. סטטיסטיקאי ב

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?

נספח : טבלת התפלגות T



P

דרגות חופש	0.75	0.90	0.95	0.975	0.99	0.995	0.9995
1	1.000	3.078	6.314	12.709	31.821	63.657	636.619
2	0.816	1.886	2.920	4.303	6.965	9.925	31.598
3	0.765	1.638	2.353	3.182	4.541	5.841	12.941
4	0.741	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	1.476	2.015	2.571	3.365	4.032	6.859
6	0.718	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	1.415	1.895	2.365	2.998	3.499	5.405
8	0.706	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	1.319	1.714	2.069	2.500	2.807	3.767
24	0.685	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	1.303	1.684	2.021	2.423	2.704	3.551
60	0.679	1.296	1.671	2.000	2.390	2.660	3.460
120	0.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.291

פתרונות:

שאלה 1

א. $79.88 < \mu < 89.72$

שאלה 4

א. $96.63 < \mu < 107.37$

ב. $96.90 < \mu < 107.10$

שאלה 5

$3.149 < \mu < 3.351$

שאלה 8

90%

פרק 28 - בדיקת השערות כללית

רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטיקה.

בתהליך זה ישנן שתי השערות שנבדקות :

השערת האפס המסומנות ב- H_0

והשערה 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 .

למשל ,

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 שניקרא כלל הכרעה :

הכלל יוצר אזור שניקרא אזור דחייה (דחייה של השערת האפס כלומר קבלה של האלטרנטיבה) ואזור 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. נשנה את כלל ההכרעה אנחנו יכולים לקבל מסקנה אחרת . נבצע מדגם חדש אנחנו עלולים לקבל תוצאה אחרת.

לכן יתכנו טעויות במסקנות שלנו :

		הכרעה	
		H0	H1
מציאות	H0	אין טעות	טעות מסוג 1
	H1	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.

מה הן הטעויות האפשריות במחקר של התרופות? (בהקלטה)

נגדיר את ההסתברויות הבאות :

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות)

$\boldsymbol{\alpha}$=P(לדחות את H_0 | H_0 נכונה)= P_{H_0}(לדחות H_0)

הסיכוי לבצע טעות מסוג 2 :

$\boldsymbol{\beta}$ =P(לקבל את H_0 | H_1 נכונה)=P_{H_1}(לקבל H_0)

רמת בטחון :

$\boldsymbol{(1-\alpha)}$ =P(לקבל את H_0 | H_0 נכונה)= P_{H_0}(לקבלH_0)

עוצמה :

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

דוגמה: (פתרון בהקלטה)

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

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

ב. חשבו את הסיכוי לטעות מסוג שני והעוצמה של המבחן המוצע.

<u>תרגילים:</u>

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 $k=1,2,......,n$

   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	**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 "בנוני" או גרוע מכך נכריע שהמוצר בא ממפעל "שמשון", מה מהן ההסתברויות לסוגי הטעויות השונים?

ג. איזה כלל החלטה עדיף? נמק!

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 5 דרגות חופש)

$H_1: X \sim Z$ (התפלגות נורמאלית סטנדרטית)

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 6600 יחידות מופעלת מערכת ההתראה. נתון שעוצמת הזיהום מתפלגת נורמאלית.
   א. מה הסיכוי להתראת שווא? (באיזה סוג טעות מדובר)?
   ב. מה הסיכוי שבמצב חירום מערכת ההתראה לא תפעל? (באיזה סוג טעות מדובר)?
   ג. מה ההסתברות שאם המצב הוא מצב חירום מערכת ההתראה תפעל? (איך קוראים להסתברות זו)?
   ד. בסעיפים הבאים נשנה בכל סעיף נתון מסוים. כל סעיף עומד בפני עצמו, כיצד השינוי ישנה את הסיכוי לטעות מסוג ראשון ושני?
      1. המפעל יקנה עוד 4 חיישנים.
      2. מצב חרום מוגדר כעת בתוחלת של 7500 יחידות.
      3. מערכת ההתראה תופעל אם ממוצע של תשעת החיישנים יהיה מעל 6700.

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 העצמה בהנחה ובחברה 30% בנים?

9. זמן ההשפעה של משכך הכאבים "אופטלנוס" מתפלג נורמאלית עם תוחלת של 40 דקות וסטיית תקן של 12 דקות.

חברת התרופות המייצרת את התרופה מנסה לשפר את התרופה כך שתוחלת הזמן עד להשפעה תתקצר. לצורך כך, דגמו 25 מטופלים שיקבלו את התרופה "אופטלנוס פורטה", ממוצע זמן התגובה של המטופלים היה 34.5 דקות. חברת התרופות החליטה מראש שאם ממוצע הזמן עד להשפעה יהיה נמוך מ 35 דקות, היא תמשיך בתהליך שיווק "אופטלנוס פורטה".

א. מהי רמת המובהקות של המבחן המוצע?

ב. על סמך תוצאות המדגם. מהי המסקנה ומהי הטעות האפשרית במסקנה?

ג. מהי עצמת המבחן המוצע אם במציאות התרופה "אופטלנוס פורטה" מפחיתה את התוחלת לכדי 32דקות?

ד. כיצד תשתנה התשובה לסעיף ג' אם החברה הייתה מחליטה שהיא תמשיך בתהליך שיווק התרופה החדשה כאשר ממוצע המדגם יהיה נמוך מ-36 דקות?

פתרונות:

שאלה 2

$\alpha = \frac{1}{3}$ $\beta = \frac{2}{5}$

שאלה 3

$\alpha = 0.25$ $\beta = 0.5$

שאלה 4

א. $\alpha = 0.2$ $\beta = 0.8$

ב. $\alpha = 0.2$ $\beta = 0.3$

ג. כלל ב'

שאלה 5

ב. 0.00781

ג. 0.1678

שאלה 6

א. 0.05

ב. 0.022

שאלה 7

א. 0.0228

ב. 0.0918

ג. 0.9082

שאלה 8

א. 0.055

ב. 0.383

פרק 29 - בדיקת השערות על פרמטרים

הקדמה

רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטיקה.
בבדיקת השערות על פרמטרים נעבוד לפי השלבים הבאים :

שלב א: נזהה 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 שניקרא **כלל הכרעה :**
הכלל יוצר אזור שניקרא **אזור דחייה** (דחייה של השערת האפס כלומר קבלה של האלטרנטיבה) **ואזור קבלה** (קבלה של השערת האפס ודחייה של האלטרנטיבה). כלל ההכרעה מתבסס על איזשהו סטטיסטי .
אזור הדחיה 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 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:

$n = 20$

$\overline{X} = 3120$

$S = 280$

א. מהי אוכלוסיית המחקר?

ב. מה המשתנה הנחקר?

ג. מה הפרמטר הנחקר?

ד. מהן השערות המחקר?

תרגילים:

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.
   א. מהי אוכלוסיית המחקר?
   ב. מה המשתנה הנחקר?
   ג. מה הפרמטר הנחקר?
   ד. מהן השערות המחקר?

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. מהי אוכלוסיית המחקר?
   ב. מה המשתנה הנחקר?
   ג. מה הפרמטר הנחקר?
   ד. מהן השערות המחקר?

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.
   א. מהי אוכלוסיית המחקר?
   ב. מה המשתנה הנחקר?
   ג. מה הפרמטר הנחקר?
   ד. מהן השערות המחקר?

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.
   א. מהי אוכלוסיית המחקר?
   ב. מה המשתנה הנחקר?
   ג. מה הפרמטר הנחקר?
   ד. מהן השערות המחקר?

טעויות בבדיקת השערות

רקע:

בתהליך של בדיקת השערות יוצרים כלל שניקרא **כלל הכרעה** :

הכלל יוצר אזור שניקרא **אזור דחייה** (דחייה של השערת האפס כלומר קבלה של האלטרנטיבה) ו**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. נשנה את כלל ההכרעה אנחנו יכולים לקבל מסקנה אחרת . נבצע מדגם חדש אנחנו עלולים לקבל תוצאה אחרת.

לכן יתכנו טעויות במסקנות שלנו :

	הכרעה		
		H0	H1
מציאות	H0	אין טעות	טעות מסוג 1
	H1	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?

תרגילים:

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?

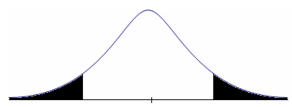
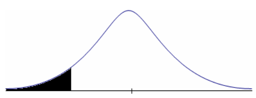
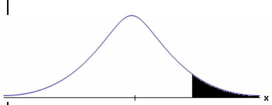
2. במחקר על פרמטר מסוים הוחלט בסופו של דבר לדחות את השערת האפס.
   א. האם ניתן לדעת אם בוצע טעות במחקר?
   ב. מה סוג הטעות האפשרית?

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?

פרק 30 - בדיקת השערות על תוחלת (ממוצע)

כאשר שונות האוכלוסיה ידועה

רקע:

השערת האפס : **השערת אלטרנטיבה:**	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$
תנאים:	1. σ ידועה 2. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0 :	$Z_{\bar{x}} > Z_{1-\frac{\alpha}{2}}$ או $Z_{\bar{x}} < -Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}}$ $Z_{1-\frac{\alpha}{2}}$ ■ - דוחים את 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

סטטיסטי המבחן :

$$Z_{\bar{X}} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

חלופה אחרת לכלל הכרעה :

נדחה H_0 אם מתקיים:	$\bar{X} > \mu_0 + Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$

דוגמה : (פתרון בהקלטה)

יבול העגבניות מתפלג נורמלית עם תוחלת של 10 טון לדונם וסטיית תקן של 2.5 טון לדונם בעונה. משערים ששיטת זיבול חדשה תעלה את תוחלת היבול לעונה מבלי לשנות את סטיית התקן. נדגמו 4 חלקות שזובלו בשיטה החדשה. היבול הממוצע שהתקבל היה 12.5 טון לדונם. בדוק את ההשערה ברמת מובהקות של 1%.

תרגילים:

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 תהייה 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 מכוילת (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? בחר בתשובה הנכונה.
א. הגדלת רמת המובהקות לא תשתנה 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: (בחר בתשובה הנכונה)

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.

פתרונות :

<u>שאלה 1:</u>

נקבל H_0

<u>שאלה 2:</u>

נדחה H_0

<u>שאלה 3:</u>

נדחה H_0

<u>שאלה 4:</u>

נדחה H_0

<u>שאלה 5:</u>

נקבל H_0

<u>שאלה 6:</u>

ב

<u>שאלה 7:</u>

ג

<u>שאלה 8:</u>

א. אותה מסקנה

ב. לא ניתן לדעת.

סיכוי לטעויות ועוצמה כאשר שונות האוכלוסייה ידועה

רקע:

	הכרעה		
		H0	H1
מציאות	H0	אין טעות	טעות מסוג 1
	H1	טעות מסוג 2	אין טעות

נגדיר את ההסתברויות הבאות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות)

P_{H_0} (לדחות H_0) $=(H_0$ נכונה | לדחות את $H_0)$P$=$**α**

הסיכוי לבצע טעות מסוג 2:

P_{H_1} (לקבל H_0) $=(H_1$ נכונה | לקבל את $H_0)$P$=$ **β**

רמת בטחון:

P_{H_0} (לקבל H_0) $=(H_0$ נכונה | לקבל את $H_0)$P$=$ **(1-α)**

עוצמה :

P_{H_1} (לדחות H_0) $=(H_1$ נכונה | לדחות את $H_0)$P$=$ **(1-β)=π**

<u>התהליך לחישוב סיכוי לטעות מסוג שני</u> :

השערת האפס : **השערה אלטרנטיבה:**	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
תנאים:	3. σ ידועה 4. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0 :	$\bar{X} > \mu_0 + Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$
חישוב β :	$P_{H_1}(\mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}})$	$P_{H_1}(\bar{X} > \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}})$	$P_{H_1}(\bar{X} < \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}})$

התפלגות ממוצע המדגם : $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

התקנון : $Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

<u>דוגמה</u> : (פתרון בהקלטה)

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'?

תרגילים:

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. רשום כלל הכרעה במונחי ממוצע מדגם קריטי וברמת מובהקות של 10%.

ב. בהמשך לסעיף א מהי העצמה אם התוחלת שווה ל- 195?

ג. הסבר ללא חישוב איך העצמה תשתנה אם רמת המובהקות תהייה 5%?

4. מפעל לייצור צינורות מייצר צינור שקוטרו מתפלג נורמלית עם תוחלת של 50 מ"מ וסטית תקן של 6 מ"מ. במחלקת ביקורת האיכות דוגמים בכל יום 81 צינורות ומודדים את קוטרם, בכדי לבדוק, בעזרת מבחן סטטיסטי, האם מכונת הייצור מכוילת כנדרש או שקוטר הצינורות קטן מהדרוש.
   א. רשום 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. הסבר ללא חישוב כיצד התשובה לסעיף ב תשתנה אם רמת המובהקות תגדל.
   ד. הסבר 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) ?
      1. סטיית התקן הייתה יותר גדולה .
      2. הסיכוי לטעות מסוג שני גדול יותר.

<u>השאלות שלהלן הן שאלות רב בררתיות. בחר בכל שאלה את התשובה הנכונה ביותר:</u>

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 מהתושבות לא נכונה בהכרח.

8. מה המצב הרצוי לחוקר המבצע בדיקת השערה :

	α	$1-\beta$
א.	גדולה	גדולה
ב.	גדולה	קטנה
ג.	קטנה	גדולה
ד.	קטנה	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 ?
א. טעות מסוג ראשון.
ב. טעות מסוג שני.
ג. טעות מסוג שלישי.
ד. אין טעות במסקנתו.

פתרונות :

שאלה 1:

א. מעל 6.645

ב. 0.3632

שאלה 2:

א. נדחה H_0 אם $\bar{X} < 2.24$

ב. נדחה H_0

ג. 1

שאלה 3:

א. נדחה H_0 אם $\bar{X} > 203.29$ או $\bar{X} < 196.71$

ב. 0.8051

ג. תקטן.

שאלה 4:

א. נדחה H_0 אם $\bar{X} < 48.9$

ב. 0.0885

ג. תקטן.

ד. תקטן.

שאלה 6:

ד

שאלה 7:

ג

שאלה 8:

ג

שאלה 9:

א

שאלה 10:

ב

קביעת גודל מדגם כששונות האוכלוסיה ידועה

רקע:

השערות המחקר הן :
$$H_0 : \mu = \mu_0$$
$$H_1 : \mu = \mu_1$$

סטיית התקן של האוכלוסייה ידועה σ ומעוניינים לבצע מחקר שרמת המובהקות לא תעלה על α והסיכוי לטעות מסוג שני לא יעלה על β.

הנוסחה הבאה נותנת את גודל המדגם הרצוי :

$$n \geq \left(\frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

דוגמה: (פתרון בהקלטה)

משרד החינוך מפעיל בגן חובה שיטת חינוך שפותחה בשנת 1995. לפי שיטת חינוך זו תוחלת הציון במבחן אוצר מילים לגיל הרך הוא 70. אנשי חינוך החליטו לבדוק שיטת חינוך שפותחה בהולנד הנותנת שם תוחלת ציון אוצר מילים של 80.
נניח שציוני מבחן זה מתפלגים נורמאלית עם $\sigma = 17$.
כדי לבדוק האם גם בישראל הפעלת שיטת החינוך ההולנדית תעבוד בגנים, רוצים לבנות מחקר ברמת מובהקות של 5%. כמו כן, מעוניינים שאם בהפעלת השיטה ההולנדית תוחלת הציונים תעלה לכדי 80, המחקר יגלה זאת בסיכוי של 90%. כמה ילדי גן חובה דרושים למחקר?

תרגילים:

1. במבחן אינטליגנציה הציונים מתפלגים נורמאלית עם סטיית תקן 8 וממוצע 100. פסיכולוג מעוניין לבדוק את הטענה שבאוכלוסיות במצב סוציו אקונומי נמוך תוחלת הציונים היא 95. אם מעוניינים לגלות את הטענה בהסתברות של לפחות 99% כשרמת המובהקות היא 5% מהו גודל המדגם הדרוש?

2. משרד התקשורת טוענים שאדם מדבר בממוצע 180 דקות בחודש בטלפון הסלולרי. חברות הטלפון הסלולרי טוענות שאינפורמציה זו אינה נכונה ואדם מדבר בממוצע פחות : כ-160 דקות. לצורך פתרון נניח שסטיית התקן של זמן השיחה החודשי ידוע ושווה ל-60 דקות. כמה אנשים יש לדגום כך שאם טענת משרד התקשורת נכונה נדחה אותה בסיכוי של 5% (איך קוראים להסתברות זאת?) כמו כן אם טענת חברות הטלפון הסלולרית נכונה המחקר יגלה זאת בסיכוי של 90% (איך קוראים להסתברות זאת?)

3. השערות המחקר הן : $H_0: \mu = \mu_0$, $H_1: \mu = \mu_1$

כמו כן נתון שהמשתנה מתפלג נורמלית עם סטיית התקן ידועה σ מעוניינים לבצע מחקר שרמת המובהקות לא תעלה על α והסיכוי לטעות מסוג שני לא יעלה על β.

הוכח שגוגל המדגם הרצוי לכך יהיה :

$$n \ge \left(\frac{(Z_{1-\alpha} + Z_{1-\beta}) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

פתרונות :

שאלה 1:

41

שאלה 2 :

78

שאלה 3:

הוכחה

מובהקות התוצאה (p-value) בבדיקת השערות על תוחלת עם שונות ידועה

רקע:

דרך נוספת להגיע להכרעות שלא דרך כלל הכרעה, היא דרך חישוב מובהקות התוצאה:

באמצעות תוצאות המדגם מחשבים את מובהקות התוצאה שמסומן ב- p_v .

את רמת המובהקות החוקר קובע מראש לעומת זאת, את מובהקות התוצאה החוקר יוכל לחשב רק אחרי שיהיו לו את התוצאות.

המסקנה של המחקר תקבע לפי העיקרון הבא:

אם $p_v \le \alpha$ דוחים את H_0

מובהקות התוצאה זה הסיכוי לקבלת תוצאות המדגם וקיצוני מתוצאות אלה בהנחת השערת האפס.

$p_v = P_{H_0}$ (לקבל את תוצאות המדגם וקיצוני)

אם ההשערה היא דו צדדית:

$p_v = 2\,P_{H_0}$ (לקבל את תוצאות המדגם וקיצוני)

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

השערת האפס: **השערה אלטרנטיבה:**	$H_0: \mu = \mu_0$ $H_1: \mu \ne \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
תנאים:	5. σ ידועה 6. $X \sim N$ או מדגם מספיק גדול		
p-value	אם $\bar{x} > \mu_0 \Leftarrow 2 \cdot P_{H_0}(\bar{X} \ge \bar{x})$ אם $\bar{x} < \mu_0 \Leftarrow 2 \cdot P_{H_0}(\bar{X} \le \bar{x})$	$P_{H_0}(\bar{X} \le \bar{x})$	$P_{H_0}(\bar{X} \ge \bar{x})$

כאשר בהנחת השערת האפס: $\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n})$

$$Z_{\bar{x}} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

<u>דוגמה:</u> (פתרון בהקלטה)

המשקל הממוצע של מתגייסים לצבא לפני 20 שנה היה 65 ק״ג. מחקר מעוניין לבדוק האם כיום המשקל הממוצע של מתגייסים גבוה יותר. נניח שמשקל המתגייסים מתפלג נורמאלית עם סטיית תקן של 12 ק״ג. במדגם של 16 מתגייסים התקבל משקל ממוצע של 71 ק״ג.

א. מהי מובהקות התוצאה?

ב. מה המסקנה אם רמת המובהקות היא 5% ואם רמת המובהקות היא 1%?

תרגילים:

1. לפניך השערות של מחקר :

$$H_0 : \mu = 70$$
$$H_1 : \mu > 70$$

המשתנה הנחקר מתפלג נורמלית עם סטיית תקן 20. במדגם מאותה אוכלוסייה התקבלו התוצאות הבאות :

$$n = 100$$
$$\bar{x} = 74$$

מהי מובהקות התוצאה?

2. השכר הממוצע במשק בשנת 2012 היה 8800 ₪ עם סטיית תקן 2000. במדגם שנעשה אתמול על 100 עובדים התקבל שכר ממוצע 9500 ₪ . מטרת המחקר היא לבדוק האם כיום חלה עליה בשכר. עבור אילו רמות מובהקות שיבחר החוקר יוחלט שחלה עליה בשכר הממוצע במשק?

3. אדם חושד שחברת ממתקים לא עומדת בהתחייבויותיה, ומשקלו של חטיף מסוים אותו הוא קונה מדי בוקר נמוך מ – 100 גרם. חברת הממתקים טוענת מצידה שהיא אכן עומדת בהתחייבויותיה. ידוע כי סטית התקן של משקל החטיף היא 12 גרם. האדם מתכוון לשקול 100 חפיסות חטיפים ולאחר מכן להגיע להחלטה. לאחר הבדיקה הוא קיבל משקל הממוצע של 98.5 גרם.

   א. רשמו את השערות המחקר.

   ב. מהי רמת המובהקות המינימלית עבורה דוחים את השערת האפס?

   ג. מהי רמת המובהקות המקסימלית עבורה נקבל את השערת האפס?

   ד. מה המסקנה ברמת מובהקות של 5?

4. מכונה לחיתוך מוטות במפעל חותכת מוטות באורך שמתפלג נורמאלית עם תוחלת אליה כוונה המכונה וסטיית תקן 2 ס"מ. ביום מסוים כוונה המכונה לחתוך מוטות באורך 80 ס"מ. אחראי האיכות מעוניין לבדוק האם המכונה מכוילת. לצורך כך נדגמו מקו הייצור 16 מוטות שנחתכו אורכן הממוצע היה 81.7 ס"מ.

   א. מהי רמת המובהקות המינימלית עבורה נכריע שהמכונה לא מכוילת?

   ב. אם נוסיף עוד תצפית שערכה יהיה 82 ס"מ , כיצד הדבר ישפיע על התשובה של הסעיף הקודם?

   ג. הכרע ברמת מובהקות של 5% האם המכונה מכוילת.

5. אם מקבלים בחישובים אלפא מינימלית (P value) קטנה מאוד, סביר להניח כי החוקר ידחה את השערת האפס בקלות. נכון ? לא נכון? נמק.

6. בבדיקת השערות התקבל שה- $p\text{-}value=0.02$.
מה תהיה מסקנת חוקר המשתמש ברמת מובהקות 1%? בחר בתשובה הנכונה.
   א. יקבל את השערת האפס בכל מקרה.
   ב. ידחה את השערת האפס מקרה.
   ג. ידחה את השערת האפס רק אם המבחן הנו דו צדדי.
   ד. לא ניתן לדעת כי אין מספיק נתונים.

7. מובהקות התוצאה (PV) היא גם : (בחר בתשובה הנכונה)
   א. רמת המובהקות המינימאלית לדחות השערת האפס.
   ב. רמת המובהקות המקסימאלית לדחיית השערת האפס.
   ג. רמת המובהקות שנקבעת מראש על ידי החוקר טרם קיבל את תוצאות המחקר.
   ד. רמת המובהקות המינימאלית לאי דחיית השערת האפס.

8. בבדיקת השערות מסוימת התקבל p value=0.0254 לכן (בחר בתשובה הנכונה) :
   א. ברמת מובהקות של 0.01 אך לא של 0.05 נדחה את H_0.
   ב. ברמת מובהקות של 0.01 ושל 0.05 לא נדחה את H_0.
   ג. ברמת מובהקות של 0.05 אך לא של 0.01 נדחה את H_0.
   ד. ברמת מובהקות של 0.01 ושל 0.05 נדחה את H_0.

פתרונות :

שאלה 1:

0.0228

שאלה 2:

עבור כל רמת מובהקות סבירה.

שאלה 3:

ב. 0.1056

ג. 0.1056

ד. נכריע שיש עמידה בהתחייבות של החברה.

שאלה 4:

א. 0.0006

ב. יקטן.

ג. נכריע שאין כיול.

שאלה 5:

נכון

שאלה 6:

תשובה :א

שאלה 7:

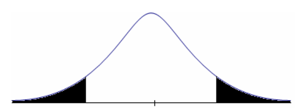
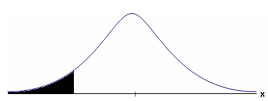
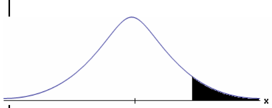
תשובה : א

שאלה 8:

תשובה : ג

בדיקת השערות על תוחלת (ממוצע) כאשר שונות האוכלוסייה אינה ידועה

רקע:

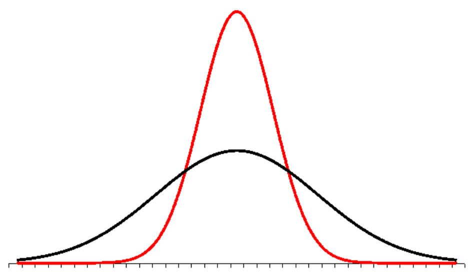
השערת האפס : **השערה אלטרנטיבה:**	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
תנאים:	7. σ אינה ידועה 8. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0 :	$t_{\bar{x}} > t_{1-\frac{\alpha}{2}}^{(n-1)}$ או $t_{\bar{x}} < -t_{1-\frac{\alpha}{2}}^{(n-1)}$  $-t_{1-\frac{\alpha}{2},n-1}$ $t_{1-\frac{\alpha}{2},n-1}$ - דוחים את H_0	$t_{\bar{x}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha,n-1}$ - דוחים את H_0	$t_{\bar{x}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha,n-1}$ - דוחים את H_0
חלופה לכלל הכרעה : נדחה H_0 אם מתקיים:	$\bar{X} > \mu_0 + t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$ או $\bar{X} < \mu_0 - t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 - t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$

סטטיסטי המבחן :

$$t_{\bar{x}} = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

$$S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} X^2{}_i - n\bar{X}^2}{n-1}$$

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.

<u>דוגמה:</u> (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%.

תרגילים:

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 מבארה"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.

א. אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.

ב. אם רוני תדחה את השערת האפס אז גם בהכרח ליאור.

ג. שני החוקרים בהכרח יגיעו לאותה מסקנה.

ד. 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:

$$H_0 : \mu = \mu_0$$
$$H_0 : \mu < \mu_0$$

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.

<u>פתרונות:</u>

שאלה 1:

נדחה H_0

שאלה 2:

נדחה H_0

שאלה 3:

נקבל H_0

שאלה 4:

נקבל H_0

שאלה 5:

התשובה היא : ב

שאלה 6:

התשובה היא : ג

מובהקות התוצאה (p-value) כאשר שונות האוכלוסייה לא ידועה

רקע:

נזכיר שהמסקנה של המחקר תקבע לפי העיקרון הבא :

אם $p_v \le \alpha$ דוחים את H_0

מובהקות התוצאה זה הסיכוי לקבלת תוצאות המדגם וקיצוני מתוצאות אלה בהנחת השערת האפס.

$p_v = P_{H_0}$ (לקבל את תוצאות המדגם וקיצוני)

אם ההשערה היא דו צדדית :

$p_v = 2P_{H_0}$ (לקבל את תוצאות המדגם וקיצוני)

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

השערת האפס : **השערה אלטרנטיבה:**	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$
תנאים:	9. σ אינה ידועה 10. $X \sim N$ או מדגם מספיק גדול		
p-value	אם $\bar{x} > \mu_0 \Leftarrow 2 \cdot P_{H_0}(\bar{X} \ge \bar{x})$ אם $\bar{x} < \mu_0 \Leftarrow 2 \cdot P_{H_0}(\bar{X} \le \bar{x})$	$P_{H_0}(\bar{X} \le \bar{x})$	$P_{H_0}(\bar{X} \ge \bar{x})$

$$t_{\bar{x}} = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

$$S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n-1}$$

$$d.f = n-1$$

<u>דוגמה :</u> (פתרון בהקלטה)

ממוצע זמן הנסיעה של אדם לעבודה הינו 40 דקות. הוא מעוניין לבדוק דרך חלופית שאמורה להיות יותר מהירה. לצורך כך הוא דוגם 5 ימים שבהם הוא נוסע בדרך החלופית. זמני הנסיעה שקיבל בדקות הם : 27,34,32,40,30 . הנח שזמן הנסיעה מתפלג נורמלית.

א. רשום את השערות המחקר.

ב. מצא חסמים למובהקות התוצאה.

ג. מה המסקנה ברמת מובהקות של 5% ?

<u>תרגילים :</u>

1. קו ייצור אריזות סוכר נארזות כך שהמשקל הממוצע של אריזות הסוכר צריך להיות אחד קילוגרם. בכל יום דוגמים מקו הייצור 5 אריזות במטרה לבדוק האם קו הייצור תקין. בבדיקה דגמו 5 אריזות סוכר ולהלן משקלן בגרמים :

   1008,1024,996,1005,997

   א. רשמו את השערות המחקר.

   ב. מהי מובהקות התוצאה? הצג חסמים.

   ג. מה המסקנה ברמת מובהקות של 5%?

2. חוקר בדק את הטענה כי פועלים העובדים במשמרת לילה איטיים יותר מפועלים העובדים ביום. ידוע כי משך הזמן הממוצע הדרוש לייצר מוצר מסוים ביום הוא 6 שעות. במדגם מיקרי של 25 פועלים שעבדו במשמרת לילה נמצא כי הזמן הממוצע לייצר אותו מוצר הוא 7 שעות עם סטית תקן של 3 שעות.

   מהי ה- α המינימלית שלפיה ניתן להחליט שאכן העובדים במשמרת לילה איטיים יותר ?

3. הגובה של מתגייסים לצה"ל מתפלג נורמלית . במדגם של 25 מתגייסים מדדו את הגבהים שלהם בס"מ והתקבלו התוצאות הבאות :

   $$\bar{x} = 176.2$$

   $$\sum (x_i - \bar{x})^2 = 2832$$

   מטרת המחקר היא לבדוק האם תוחלת הגבהים של המתגייסים גבוה מ-174 ס"מ באופן מובהק. מהי בקרוב מובהקות התוצאה ועל פיה מה תהיה המסקנה ברמת מובהקות של 6% ?

פתרונות :

שאלה 3:

נקבל H_0

<u>הקשר בין רווח סמך לבדיקת השערות על תוחלת</u>

<u>רקע:</u>

ניתן לבצע בדיקת השערות דו צדדית ברמת מובהקות α על μ :

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

על ידי בניית רווח סמך ברמת סמך של $1-\alpha$ ל μ :

אם μ_0 נופל ברווח ← נקבל את H_0

אם μ_0 לא נופל ברווח ← נדחה את H_0

<u>דוגמה:</u> (פתרון בהקלטה)

חוקר ביצע בדיקת השערות לתוחלת. להלן השערותיו :

$$H_0: \mu = 80$$
$$H_1: \mu \neq 80$$
$$\alpha = 5\%$$

החוקר בנה רווח סמך ברמה של 90% וקיבל : $79 < \mu < 84$.

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

תרגילים :

1. חוקר רצה לבדוק את ההשערות הבאות :

$$H_0 : \mu = 90$$
$$H_1 : \mu \neq 90$$

החוקר בנה רווח סמך לתוחלת ברמת סמך של 95% וקיבל את רווח הסמך הבא : $(87,97)$.
אם החוקר מעוניין לבצע בדיקת השערות ברמת מובהקות של 1% האם ניתן להגיע למסקנה ע״ס רווח הסמך? נמקו.

2. חוקר מעוניין לבדוק השפעת דיאטה חדשה על רמת הסוכר בדם. ידוע כי מספר מיליגרם הסוכר בסמ״ק דם הוא משתנה מקרי שמתפלג נורמלית עם סטיית תקן 10.4מ״ג. נלקח מדגם של 60 נבדקים שניזונו מדיאטה זו. נמצא כי ממוצע מספר המיליגרם סוכר היה 115.5 מ״ג לסמ״ק.
א. בנה רווח סמך ברמת סמך 95% לתוחלת רמת הסוכר בדם אצל הניזונים מדיאטה זו.
ב. ידוע שתוחלת רמת הסוכר בדם באוכלוסיה היא 90 מ״ג לסמ״ק. האם לדעתך ניתן להסיק על סמך תוצאת סעיף א שהדיאטה משפיעה על רמת הסוכר בדם? הסבר.

3. יצרן אנטיביוטיקה רושם על גבי התרופות שכמות הפנצלין היא 200 מ״ג לקפסולה.
משרד הבריאות ביצע מדגם של 8 קפסולות אקראיות מקו הייצור ומצא שבממוצע יש 196 מ״ג פנצילין לקפסולה עם סטיית תקן מדגמית של של 5 מ״ג. בהנחה וכמות הפנצלין בקפסולה מתפלגת נורמלית.
א. בנה רווח סמך ברמת סמך של 95% לממוצע כמות הפנצלין לקפסולה המיוצרת על ידי יצרן האנטיביוטיקה.
ב. בדוק ברמת מובהקות של 5% האם יש אמת באינפורמציה המסופקת על ידי היצרן.

פתרונות :

שאלה 1:

1. נקבל השערת H_0

שאלה 2:

א. $112.87 \le \mu \le 118.13$

ב. נכריע שהדיאטה משפיעה על תוחלת רמת הסוכר בדם.

שאלה 3:

א. $191.8 \le \mu \le 200.2$

ב. נכריע שיש אמת בפרסום.

פרק 31 - מדדי קשר - מדד הקשר הלינארי (פירסון)

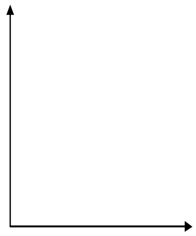
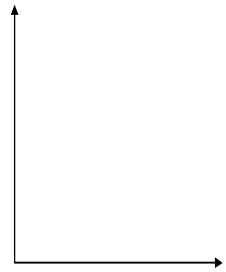
רקע:

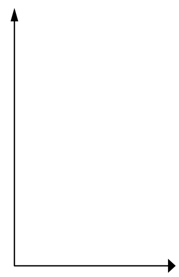
המטרה היא לבדוק האם קיים קשר (קורלציה, מתאם) של קו ישר בין שני משתנים כמותיים.
מבחינת סולמות המדידה קשר בין סולמות רווחים ומנה.
בדרך כלל, X הוא המשתנה המסביר (הבלתי תלוי) ו Y הוא המשתנה המוסבר (התלוי).למשל,
נרצה להסביר כיצד השכלה של אדם הנמדדת בשנות לימוד X– מסבירה את ההכנסה שלו Y.
במקרה זה שנות ההשכלה זהו המשתנה המסביר (או הבלתי תלוי) ואנחנו מעוניינים לבדוק כיצד שינויים בשנות ההשכלה של אדם יכולים להסביר את השינויים שלו בהכנסה , ולכן רמת ההכנסה זהו המשתנה המוסבר התלוי במשתנה המסביר אותו.
בשלב הראשון, נהוג לשרטט דיאגרמת פיזור. זו דיאגרמה שנותנת אינדיקציה ויזואלית על טיב הקשר בין שני המשתנים.
למשל, בבניין של 5 דירות בדקו את הנתונים הבאים: X - מס' חדרים בדירה. Y- מס' נפשות הגרות בדירה. להלן התוצאות שהתקבלו :

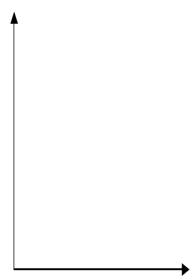
מס' דירה	X	Y
1	3	2
2	2	2
3	4	3
4	3	3
5	5	4

נשרטט מנתונים הללו דיאגרמת פיזור :

נתבונן בכמה מקרים של דיאגרמות פיזור וננתח אותן :

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

המדד בודק את כיוון הקשר (חיובי או שלילי).

ואת עוצמת הקשר (חלש עד חזק).

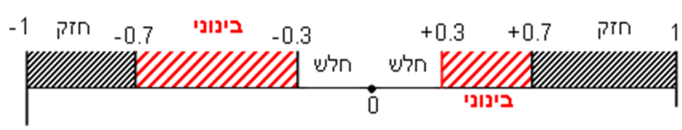
מקדם מתאם זה מקבל ערכים בין 1- ל 1.

מקדם מתאם1- או 1 אומר שקיים קשר לינארי מוחלט ומלא בין המשתנים שניתן לבטאו על ידי הנוסחה : $y = bx + a$.

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

מתאם חיובי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לעלות בערכו אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט ואילו מתאם שלילי חלקי אומר שככל שמשתנה אחד עולה לשני יש נטייה לרדת אבל לא קיימת נוסחה לינארית שמקשרת את X ל- Y באופן מוחלט.

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



מקדם המתאם יסומן באות –r .

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

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

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

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

מקדם המתאם הלינארי : $r_{xy}=\dfrac{\operatorname{cov}(x,y)}{s_x\cdot s_y}$

תרגילים:

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

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

א. שרטט דיאגראמת פיזור לנתונים. מה ניתן להסיק מהדיאגרמה על טיב הקשר ביו מספר החיסורים של תלמיד לציונו? מיהו המשתנה הבלתי תלוי ומיהו המשתנה התלוי?

ב. חשב את מדד הקשר של פירסון. האם התוצאה מתיישבת עם תשובתך לסעיף א'?

ג. הסבר ללא חישוב כיצד מקדם המתאם היה משתנה אם היה מתווסף תלמיד שהחסיר 4 פעמים וקיבל ציון 80?

2. במחקר רפואי רצו לבדוק האם קיים קשר בין רמת ההורמון X בדם החולה לרמת ההורמון Y שלו. לצורך כך מדדו את רמת ההורמונים ההלו עבור חמישה חולים.

להלן התוצאות שהתקבלו:

y	x
12	10
15	14
15	15
17	18
21	20

א. מה הממוצע של כל רמת הורמון?

ב. מהו מקדם המתאם בין ההורמונים? ומה משמעות התוצאה?

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

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

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

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

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

ב. מה המשמעות של התוצאה שקיבלת בסעיף א?

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

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

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

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

חשב את מדד הקשר הלינארי בין X ל-Y.

5. במוסד אקדמי ציון ההתאמה מחושב כך: מכפילים את הציון הממוצע בבגרות ב- 3 ומפחיתים 2 נקודות. ידוע שעבור 40 מועמדים סטיית התקן של ממוצע הציון בבגרות הייתה2. מה מקדם המתאם בין ציון ההתאמה לציון הממוצע בבגרות שלהם?

6. להלן רשימת טענות, לגבי כל טענה קבע נכון/לא נכון ונמק!

א. מתווך דירות המיר מחירי דירות מדולר לשקל. נניח שדולר אחד הוא 3.5 ₪. אם מתווך הדירות יחשב את מדד הקשר של פירסון בין מחיר הדירה בשקלים למחיר הדירה בדולרים הוא יקבל 1.

ב. לסדרה של נתונים התקבל $\bar{X}=\bar{Y}=6$ $S_X = S_Y = 1$ לכן מדד הקשר של פירסון יהיה 1.

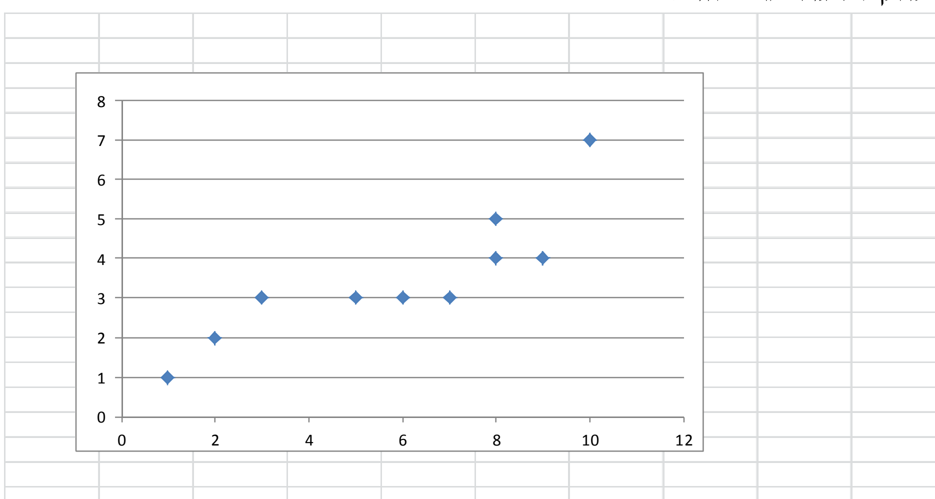
ג. אם השונות המשותפת של X ושל Y הינה 0 אז בהכרח גם מקדם המתאם של פירסון יהיה 0.

<u>שאלות אמריקאיות:</u>

7. נמצא שקיים מקדם מתאם שלילי בין הציון בעברית לציון בחשבון בבחינה לכן :
   א. הדבר מעיד שהציונים בכתה היו שליליים.
   ב. ככל שהציון של תלמיד יורד בחשבון יש לו נטייה לרדת בעברית.
   ג. ככל שהציון של תלמיד עולה בחשבון יש לו נטייה לרדת בעברית.
   ד. אף אחת מהתשובות לא נכונה.

8. נלקחו 20 מוצרים וניבדק ביום מסוים המחיר שלהם בדולרים והמחיר שלהם בש״ח (באותו היום ערך הדולר היה - 4.2 ₪) מהו מקדם המתאם בין המחיר בדולר למחיר בש״ח?
   א. 1
   ב. 0
   ג.4.2
   ד. לא ניתן לדעת.

9. להלן דיאגראמת פיזור :



מה יהיה מקדם המתאם בין שני המשתנים?

   א. 1
   ב. 0.85
   ג. 0.15
   ד. 0

פתרונות:

שאלה 1:

א. בהקלטה

ב. ‎-0.9325

שאלה 2:

א. $\bar{y} = 16$ $\bar{x} = 15.4$

ב. $r_{xy} = 0.96$

שאלה 3:

א : 0.8

שאלה 4:

0.8

שאלה 5:

1

שאלה 6:

א. נכון

ב. לא נכון

ג. נכון

שאלה 7:

התשובה : ג

שאלה 8:

התשובה : א

שאלה 9:

התשובה : ב

פרק 32 - מדדי קשר - השפעת טרנספורמציה לינאריות על מדד הקשר של פירסון

רקע:

טרנספורמציה לינארית בין אם נעשית על X ובין אם נעשית על y, או בין אם נעשית על שניהם, אינה משנה את עוצמת הקשר . היא עלולה רק לשנות את כיוונו אם השיפועים של שתי הטרנספורמציות שוני סימן.

$$r_{[(aX+b),(cY+d)]}=\begin{cases} r_{x,y} \_\_\_ if \_\_ a\cdot c>0 \\ -r_{x,y} \_\_ if \_\_ a\cdot c<0 \end{cases}$$

תרגילים:

1. מבחן בנוי מחלק כמותי ומילולי.
מקדם המתאם בין שני הציונים של שני החלקים הוא 0.9.
א. אם יעלו את כל הציונים בחלק המילולי ב- 20% , מה יהיה מקדם המתאם בין הציון המילולי החדש לציון הכמותי ובין הציון המילולי הישן לציון המילולי החדש?
ב. נגדיר משתנה חדש W להיות המרחק של הציון בחשיבה מילולית מהציון המקסימאלי בבחינה- 150. מצא את מקדם המתאם בין הציון המילולי ל- W ובין W ל-לציון הכמותי.

2. מקדם המתאם בין ההכנסה לבין ההוצאה של 10 משפחות חושב והתקבל 0.7 . אם חל גידול של 5% בהכנסת האוכלוסייה כולה וגידול של 7% בהוצאה שלה, אז מה יהיה מקדם המתאם בין ההכנסה החדשה להוצאה החדשה?

3. חברת "לקי" המייצרת גלידה החליטה לערוך מחקר לבדיקת הקשר בין מספר חבילות הגלידה הנמכרות ביום לבין הטמפרטורה באותו יום. נבדקו 10 ימים והתקבל מתאם לינארי 0.85.
חברת "לקי" דואגת להתחיל כל יום עם מלאי של 150 חבילות גלידה. בנוסף, מעוניינים כי הטמפרטורה תבוטא במעלות פרנהייט במקום במעלות צלסיוס. מה ערכו של מקדם המתאם בין מספר חבילות הגלידה שנשארות בסוף היום לבין הטמפרטורה במעלות פרנהייט?

הקשר בין מעלות צלסיוס (C°) למעלות פרנהייט (F°) נתון ע"י $F=\frac{9}{5}C+32$.

<u>בחר בתשובה הנכונה:</u>

א. 0.85
ב. -0.85
ג. 1
ד. לא ניתן לדעת.

4. מקדם המתאם בין X ל- Y הנו 0.4 כל ערכי ה-X הוכפלו ב- 2 לכן מקדם המתאם החדש בין שני המשתנים יהיה :

<u>בחר בתשובה הנכונה:</u>

א. 0.8
ב. 0.4
ג. -0.4
ד. לא ניתן לדעת.

פתרונות :

שאלה 1:

א. בין הציון המילולי הישן לחדש :1

בין הציון המילולי החדש לכמותי :0.9

ב. בין W ל ציון המילולי : 1-

בין W לציון הכמותי :0.9-

שאלה 2:

0.7

שאלה 3:

התשובה : ב

שאלה 4:

התשובה : ב

פרק 33 - מדדי קשר - רגרסיה ליניארית

רקע:

במידה וקיים קשר חזק בין שני המשתנים הכמותיים נהוג לבצע ניבויי. לבנות קו ניבויים הנקרא גם קו רגרסיה המנבא משתנה אחד על סמך האחר.

מדובר בקו שמנבא את Y על סמך X. השיטה למציאת הקו הנ"ל נקראת שיטת הריבועים הפחותים והקו המתקבל נקרא קו הרגרסיה או קו הניבויים או קו הריבועים הפחותים.

a - בעצם נותן את ערך Y כאשר X הנו אפס על גבי קו הניבויים. הוא ניקרא החותך של הקו.

b - הוא שיפוע הקו נותן בכמה בעצם Y משתנה כאשר X גדל ביחידה אחת על גבי קו הניבויים.

להלן המשוואות למציאת הפרמטרים של קו הרגרסיה:

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

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

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

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

תרגילים:

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

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

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

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

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

ב. מצא את קו הרגרסיה לניבוי ההוצאה של משפחה על סמך הכנסה שלה. הסבר את משמעות הפרמטרים של קו הרגרסיה.

ג. משפחת כהן הכניסה 15,000 ₪, מה ההוצאה הצפויה שלה?

1. נסמן ב-X את ההשכלה של אדם בשנות למוד. נסמן ב-Y את הכנסתו באלפי ₪. במחקר התקבלו התוצאות הבאות:

$$S_Y = 5 \qquad S_X = 2$$

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

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

א. חשב את מדד הקשר של פירסון בין ההשכלה להכנסה.

ב. מה ההכנסה הצפויה לאדם שהשכלתו 12 שנים?

ג. מה ההשכלה הצפויה לאדם שהכנסתו 10,000 ₪?

3. חוקר רצה לחקור את הקשר הקווי שבין הציון המבחן בסטטיסטיקה לבין מספר שעות ההכנה של הסטודנטים למבחן. במדגם של 100 סטודנטים שנבחנו בקורס נרשמו התוצאות הבאות: הציון הממוצע של הסטודנטים היה 65 עם סטיית תקן של 27. מספר שעות ההכנה הממוצע היה 30 עם סטיית תקן של 18. מקדם המתאם בין הציון לשעות ההכנה היה 0.8.

א. על פי משוואת הרגרסיה שעת הכנה נוספת משפרת את ציון המבחן ב?

ב. על פי משוואת הרגרסיה תלמיד שייגש למבחן ללא שעות הכנה כלל יקבל ציון ?

ג. מהו קו הרגרסייה לניבוי הציון לפי שעות ההכנה?

4. נתונים 2 משתנים Y,X . כמו כן נתון: X ממוצע = 1.5, שונות X = שונות Y = 4,וכן שקו הרגסיה של Y על בסיס X הינו $Y= -0.2X + 0.5$. חשב מהו מקדם המתאם בין X ל –Y?

פתרונות:

שאלה1:

א. 0.8

ב. $\tilde{Y}=0.8X+0.4$

ג. 12.4

שאלה 2:

א. 0.75

ב. 4.25 אלפי ש"ח

ג. 14.6 שנים

שאלה 3:

א. 1.2

ב. 29

ג. y=1.2x+29

שאלה 4:

-0.2

פרק 34 - מדדי קשר - רגרסיה - שונות מוסברת ושונות לא מוסברת

רקע:

המטרה ברגרסיה הנה להסביר את השונות של המשתנה התלוי.
למשל, להסביר את השונות של המשכורות באמצעות הוותק או להסביר את השוני בציונים באמצעות כמות החיסורים.

r^2 -נותן בעצם איזה חלק מהשונות של המשתנה התלוי מוסבר.
השונות המוסברת נקראת גם שונות ניבויים.
השונות הלא מוסברת נקראת גם שונות טעויות.

תרגילים :

1. נמצא קשר חיובי בעוצמה של 0.7 בין שטח דירה למחירה. כמו כן נתון שסטיית התקן של מחירי הדירות הינה 200.

   א. איזה אחוז מהשונות של מחירי הדירות מוסבר על ידי שטח הדירה?
   ב. איזה אחוז מהשונות של מחירי הדירות לא מוסבר על ידי שטח הדירה?
   ג. מהי השונות המוסברות ומהי השונות הלא מוסברת של מחירי הדירות?

2. להלן רשימת טענות, לגבי כל טענה קבע נכון/לא נכון ונמק!
   א. אם שונות הטעויות שווה ל-0 (השונות הלא מוסברת) אז מקדם המתאם של פירסון יהיה 1.
   ב. אם מקדם המתאם של פירסון בין שני משתנים הוא 1 אזי שונות הטעויות (השונות הלא מוסברת) תהיה 0.
   ג. אם השונות המשותפת של X ושל Y הינה 0 אז בהכרח גם מקדם המתאם של פירסון יהיה 0.

שאלות אמריקאיות:

בשאלות הבאות יש לבחור בתשובה הנכונה.

3. בקשר בין שני משתנים התקבל $r^2 = 0.64$ לכן :
   א. ללא יוצא מן הכלל ככל שערכי משתנה אחד עולה השני יעלה.
   ב. 64% מהשונות של משתנה אחד מוסבר על ידי המשתנה השני.
   ג. הקשר בין שני המשתנים הוא בעוצמה של 0.64.
   ד. כל התשובות נכונות.

4. אם מגדילים את r^2 מה ניתן לומר?
   א. אחוז השונות המוסברת יקטן
   ב. אחוז השונות המוסברת יגדל
   ג. אחוז השונות המוסברת יישאר ללא שינוי.
   ד. סטיית התקן משתנה
   ה. לא ניתן לדעת

5. בקורס מבוא לכלכלה ניתנו במשך השנה שני מבחנים : מבחן בסוף סימסטר א (X) ומבחן בסוף סימסטר ב (Y) . כאשר בנו את קו הרגרסיה של הציון במבחן סוף סמסטר ב לפי הציון במבחן סוף סמסטר א התקבלה שונות טעויות של 80 , ושונות ניבויים של 20 . לפי נתונים אלו מקדם המתאם בין הציון במבחן סוף סמסטר א לבין הציון במבחן סוף סמסטר ב הוא :

א. 0.44 .

ב. 0.44 - .

ג. עוצמת ההקשר הלינארי היא 0.44 , אך אין אפשרות לדעת את סימנה.

ד. אין אפשרות לחשב את מקדם המתאם.

ה. 0.35