

סטודנטים יקרים

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

הקורס באתר כולל פתרונות מלאים 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
לחויית לימודים יוצאת דופן וחדשה לחלוטין, היכנסו עכשיו לאתר
www.gool.co.il



אנו מאחלים לכם הצלחה מלאה בבחינות

צוות האתר GooL

גול, בשביל התירגול...

תוכן

פרק 1 - בעיות בסיסיות בהסתברות	4
פרק 2 - פעולות בין מאורעות (חיתוך ואיחוד), מאורעות זרים ומכילים	9
פרק 3 - הסתברות מותנית - מרחב לא אחיד	20
פרק 4 - דיאגרמת עצים, גוסחת בייס וגוסחת ההסתברות השלמה	24
פרק 5 - תלות ואי תלות בין מאורעות	30
פרק 6 - שאלות מסכמות בהסתברות	34
פרק 7 - המשתנה המקרי הבדיד - פונקציית ההסתברות	38
פרק 8 - המשתנה המקרי הבדיד - תוחלת, שונות וסטיית תקן	42
פרק 9 - המשתנה המקרי הבדיד - טרנספורמציה לינארית	46
פרק 10 - תוחלת ושונות של סכום משתנים מקריים	50
פרק 11 - התפלגויות בדידות מיוחדות - התפלגות בינומית	53
פרק 12 - התפלגויות בדידות מיוחדות - התפלגות גיאומטרית	58
פרק 13 - התפלגויות בדידות מיוחדות - התפלגות פואסונית	62
פרק 14 - המשתנה המקרי הבדיד - שאלות מסכמות	66
פרק 15 - המשתנה המקרי הרציף - התפלגויות כלליות (שימוש באינטגרלים)	72
פרק 16 - התפלגויות רציפות מיוחדות - התפלגות מעריכית	82
פרק 17 - התפלגויות רציפות מיוחדות - התפלגות נורמלית	86
פרק 18 - הסקה סטטיסטית - הקדמה	95
פרק 19 - התפלגות הדגימה	98
ממוצע המדגם ומשפט הגבול המרכזי	98
התפלגות סכום תצפיות המדגם ומשפט הגבול המרכזי	107
פרק 20 - מושגים בסיסיים באמידה	111
פרק 21 - רווח סמך לתוחלת (ממוצע האוכלוסייה)	118
רווח סמך כששונות האוכלוסייה ידועה	118
קביעת גודל מדגם באמידת תוחלת עם שונות אוכלוסייה ידועה	125
רווח סמך לתוחלת (ממוצע האוכלוסייה) כששונות האוכלוסייה אינה ידועה	128
פרק 22 - רווח סמך להפרש תוחלות ממדגמים בלתי תלויים	134
כששונות האוכלוסייה ידועות	134
כששונות האוכלוסייה אינן ידועות אך שוות והמדגמים בלתי תלויים	137
פרק 23 - רווח סמך לתוחלת ההפרש במדגם מזווג	140
פרק 24 - רווח סמך לשונות וסטיית תקן	143
פרק 25 - רווח סמך ליחס שונות	148
פרק 26 - תרגול מסכם ברווחי סמך	153
פרק 27 - בדיקת השערות על פרמטרים	158

161.....	טעויות בבדיקת השערות
163.....	פרק 28 - בדיקת השערות על תוחלת (ממוצע)
163.....	כאשר שונות האוכלוסייה ידועה
168.....	סיכוי לטעויות ועוצמה כאשר שונות האוכלוסייה ידועה
175.....	מובהקות התוצאה (p-value) בבדיקת השערות על תוחלת עם שונות ידועה
180.....	בדיקת השערות על תוחלת (ממוצע) כאשר שונות האוכלוסייה אינה ידועה
185.....	מובהקות התוצאה (p-value) כאשר שונות האוכלוסייה לא ידועה
189.....	ניתוח פלטי SPSS
194.....	הקשר בין רווח סמך לבדיקת השערות על תוחלת
197.....	פרק 29 - בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים
197.....	כשהשונות של האוכלוסייה ידועות
201.....	כשהשונות האוכלוסייה לא ידועות ומניחים שהן שוות
205.....	ניתוח פלטים
212.....	פרק 30 - בדיקת השערות על תוחלת ההפרשים במדגמים מזווגים (תלויים)
212.....	בדיקת השערות למדגמים מזווגים
218.....	מדגמים מזווגים - ניתוח פלטים
224.....	פרק 31 - הקשר בין רווח סמך לבדיקת השערות על הפרש תוחלות
228.....	פרק 32 - ניתוח שונות חד כיוונית

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 : $\{ \text{נאה, שרבי, מושלג, גשום, מעונן חלקית, אביך} \}$

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

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 ההסתברות שהספרה שנבחרה גדולה מ-5?
 - ב. מה ההסתברות שהספרה שנבחרה היא לכל היותר 3?
 - ג. מה ההסתברות שהספרה שנבחרה היא אי זוגית?
4. להלן התפלגות מספר מקלטי הטלוויזיה שנספרו עבור כל משפחה ביישוב מסוים:

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

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

5. להלן התפלגות מספר המכוניות למשפחה ביישוב "עדן" :

מספר משפחות	מספר מכוניות
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

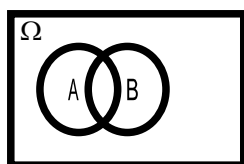
רקע:

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 : $A = \{5, 6\}$

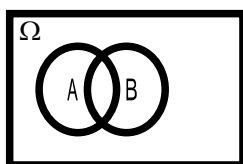
לקבל תוצאה זוגית : $B = \{2, 4, 6\}$

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

פעולת איחוד :

נותנת את כל האפשרויות שנמצאות לפחות באחת מהמאורעות. הסימון הוא : $A \cup B$ נותנת את

אשר נימצא ב-A או ב-B. כלומר, לפחות אחד מהמאורעות קורה.



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

לקבל תוצאה זוגית : $B = \{2, 4, 6\}$

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

דוגמה (הפתרון נמצא בהקלטה)

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

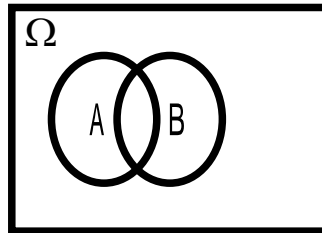
א. מה ההסתברות שלו לעבור את המבחן בסטטיסטיקה בלבד?

ב. מה ההסתברות שלו להיכשל בשני המבחנים?

ג. מה ההסתברות לעבור לפחות מבחן אחד?

נוסחת החיבור לשני מאורעות :

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



חוקי דה מורגן לשני מאורעות:

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

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

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

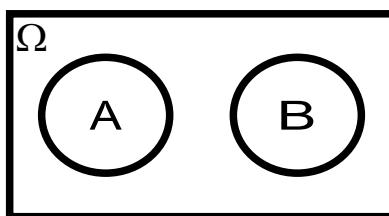
$$P(A \cup B) = 1 - P(\bar{A} \cap \bar{B})$$

שיטת ריבוע הקסם:

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

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

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



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

$$P(A \cap B) = 0$$

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

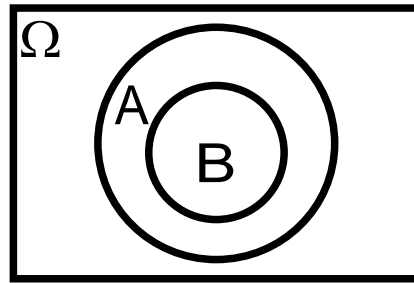
למשל, בהטלת קובייה

$$A = \{5, 6\} \quad \text{לקבל לפחות 5 :}$$

$$B = \{3\} \quad \text{לקבל 3 :}$$

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

מאורעות מכילים :



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

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

$$A \cap B = B$$

$$A \cup B = A$$

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

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

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 בדיאגרמת וון את השטח המתאים :
- א. התלמיד עבר רק את המבחן בכלכלה.
- ב. התלמיד עבר רק את המבחן בסטטיסטיקה.
- ג. התלמיד עבר את שני המבחנים.
- ד. התלמיד עבר לפחות מבחן אחד.
- ה. התלמיד נכשל בשני המבחנים.
- ו. התלמיד נכשל בכלכלה.
3. נתבקשתם לבחור ספרה באקראי. נגדיר את A להיות הספרה שנבחרה היא זוגית. נגדיר את B להיות הספרה שנבחרה קטנה מ-5.
- א. רשמו את כל התוצאות למאורעות הבאים :
- $A =$
- $B =$
- $\bar{B} =$
- $A \cap B =$
- $A \cup B =$
- 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 Ω או ϕ או A .

קבע עבור כל מאורע מה הפתרון שלו.

$=$
 $\bar{A}$
 $A \cap \phi$
 $A \cup \phi$
 $A \cap \Omega$
 $A \cup \Omega$
 $A \cap \bar{A}$
 $\bar{\phi}$
 $A \cup \bar{A}$

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. $\bar{A} \cap B$

ד. $\bar{A} \cup \bar{B}$

ה. $\bar{A}$

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 מפלגת "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?

$$P(A) = 0.3$$

$$P(B) = 0.25$$

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

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: $2 \cdot P(B \cap \bar{A}) = P(A \cap \bar{B}) = P(\bar{A} \cap \bar{B})$

מה הסיכוי למאורע A ומה ההסתברות למאורע B?

16. קבע אילו מהטענות הבאות נכונות:

א. $A \cap B = B \cap A$

ב. $\overline{A \cup B} = A \cap B$

ג. $A \cap B \cap C = A \cap B \cap (C \cup B)$

ד. $\overline{A \cap B \cap C} = \bar{A} \cup \bar{B} \cup \bar{C}$

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

א. האם יתכן ש- $p(A \cup B) = 0.4$?

ב. האם יתכן ש- $p(A \cup B) = 0.6$?

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

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

18. מתוך אזרחי המדינה הבוגרים ל- 30% חשבון בבנק הפועלים. ל-28% חשבון בבנק לאומי ול-15% חשבון בבנק מזרחי. כמו כן נתון כי 6% מחזיקים חשבון בבנק לאומי ובבנק הפועלים. ל-5% חשבון בבנק פועלים ומזרחי. ול-4% חשבון בבנק לאומי ומזרחי. כמו כן ל-1% מהאוכלוסייה הבוגרת חשבון בנק בשלושת הבנקים יחד.
- א. מה אחוז האזרחים להם חשבון בבנק לאומי בלבד?
- ב. מה ההסתברות שאזרח כלשהו יחזיק חשבון בבנק פועלים ולאומי אבל לא בבנק מזרחי?
- ג. מה ההסתברות שלאזרח יהיה חשבון בפועלים או במזרחי אבל לא בבנק לאומי?
- ד. מה אחוז האזרחים שיש להם חשבון בנק אחד בלבד?
- ה. מה אחוז האזרחים שיש להם בדיוק חשבון בשני בנקים בלבד?
- ו. מה ההסתברות שלאזרח בוגר אין חשבון בנק באף אחד מהבנקים הללו?
- ז. לאיזה אחוז מהאזרחים יש חשבון בנק בלפחות אחד מהבנקים הללו?

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

20. 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 - 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 "סופט" ו"בל". 30% מהתושבים הבוגרים רשומים אצל חברת "בל". 60% מהתושבים הבוגרים רשומים אצל חברת "סופט".
ל-15% מהתושבים הבוגרים אין טלפון סלולארי בכלל.
א. איזה אחוז מהתושבים הבוגרים רשומים אצל שתי החברות?
ב. נבחר אדם שרשום אצל חברת "סופט", מה ההסתברות שהוא רשום גם אצל חברת "בל"?
ג. אם אדם לא רשום אצל חברת "בל", מה ההסתברות שהוא כן רשום בחברת "סופט"?
ד. אם אדם רשום אצל חברה אחת בלבד, מה ההסתברות שהוא רשום בחברת "סופט"?
3. במכללה שני חניונים: חניון קטן וחניון גדול. בשעה 08:00 יש סיכוי של 60% שבחניון הגדול יש מקום, סיכוי של 30% שבחניון הקטן יש מקום וסיכוי של 20% שבשני החניונים יש מקום.
א. מה ההסתברות שיש מקום בשעה 08:00 רק בחניון הגדול של המכללה?
ב. ידוע שבחניון הקטן יש מקום בשעה 08:00, מה הסיכוי שבחניון הגדול יש מקום?
ג. אם בשעה 08:00 בחניון הגדול אין מקום, מה ההסתברות שבחניון הקטן יהיה מקום?
ד. נתון שלפחות באחד מהחניונים יש מקום בשעה 08:00, מה ההסתברות שבחניון הגדול יש מקום?
4. נלקחו 200 שכירים ו-100 עצמאים, מתוך השכירים 20 הם אקדמאיים, 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 "ישראלכרט", 20% מחזיקים כרטיס "אמריקן אקספרס", 15% מחזיקים כרטיס ויזה וגם ישראלכרט, 8% מחזיקים כרטיס ישראלכרט וגם אמריקן אקספרס ו-7% מחזיקים כרטיס ויזה וגם אמריקן אקספרס. כמו כן, 5% מחזיקים בכל שלושת הכרטיסים הנ"ל.
- א. אם לאדם יש ויזה , מה הסיכוי שאין לו כרטיס ישראלכרט?
- ב. אם לאדם שני כרטיסי אשראי , מה הסיכוי שאין לו כרטיס ישראלכרט?
- ג. אם לאדם לפחות כרטיס אשראי אחד, מה הסיכוי שאין לו כרטיס ישראלכרט?

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 4 - דיאגרמת עצים, נוסחת בייס ונוסחת ההסתברות השלמה

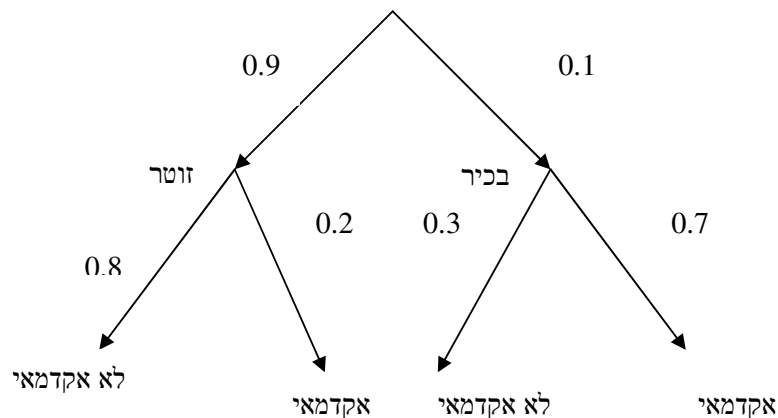
רקע:

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 ?

$$0.1 \cdot 0.7 + 0.9 \cdot 0.2 = 0.25$$

ד. נבחר אקדמאי מה ההסתברות שהוא עובד זוטרי?

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

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

נוסחת ההסתברות השלמה

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

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

נוסחת בייס

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

תרגילים:

1. בשקית סוכריות 4 סוכריות תות ו-3 לימון . מוציאים באקראי סוכרייה אם היא בטעם תות אוכלים אותה ומוציאים 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 ש 68% מהאוכלוסייה תחלה בשפעת בחורף.

א. מה אחוז החולים בשפעת בקרב האוכלוסייה הקשישה?

ב. נבחר אדם שלא חלה בשפעת, מה ההסתברות שהוא לא קשיש?

7. רדאר שנמצא על החוף צריך לקלוט אנייה הנמצאת ב-1 מ-4 האזורים A B C D.

אם האנייה נמצאת באזור A הרדאר מזהה אותה בסיכוי 0.8, סיכוי זה פוחת ב-0.1 ככל שהאנייה מתקדמת באזור.

כמו כן נתון שבהסתברות חצי האנייה נמצאת באזור D, בהסתברות 0.3 באזור C, באזור B היא

נמצאת בסיכוי 0.2, אחרת היא נמצאת באזור A.

א. מה הסיכוי 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 מיהו 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 5 - 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: $P(A \cap B) = P(A) \cdot P(B)$.

הוכחה לכך:

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

נשתמש בנוסחאות של מאורעות בלתי תלויים רק אם נאמר במפורש שהמאורעות בלתי תלויים

בתרגיל או שמהקשר אפשר להבין ללא צל של ספק שהמאורעות בלתי תלויים.

למשל,

חוקר מבצע שני ניסויים בלתי תלויים הסיכוי להצליח בניסוי הראשון הנו 0.7 והסיכוי להצליח

בניסוי השני הוא 0.4.

א. מה הסיכוי להצליח בשני הניסויים יחדו?

כיוון שהמאורעות הללו בלתי תלויים:

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

ב. מה הסיכוי להיכשל בשני הניסויים?

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

הרחבה: אי תלות בין n מאורעות

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

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

תרגילים:

1. נתון:

$$p(A) = 0.2$$

$$P(B) = 0.5$$

$$P(A \cup B) = 0.6$$

האם המאורעות הללו בלתי תלויים?

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

א. מה הסיכוי להצליח בשני המבחנים יחדו?

ב. מה הסיכוי שניכשל בשני המבחנים ?

3. במדינה מסוימת 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: $P(A \cup B) = 0.9$, $P(A|B) = 0.6$, $P(A \cap \bar{B}) = 0.3$. 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 $p(A \cup B) = p(A) \cdot p(B)$ 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 חיוביים לכן הם מאורעות תלויים.

ד. A ו-B מאורעות תלויים שסיכוייהם חיוביים לכן A ו-B מאורעות זרים.

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

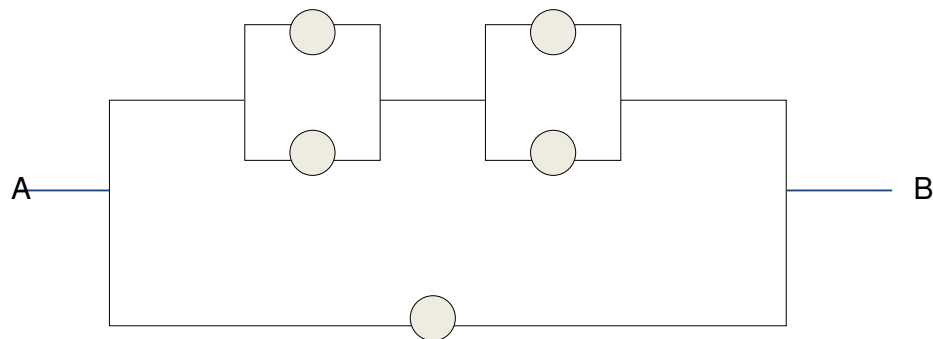
פרק 6 - שאלות מסכמות בהסתברות

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 בילויים 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

פרק 7 - המשתנה המקרי הבדיד - פונקציית ההסתברות

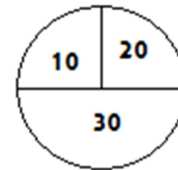
רקע:

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

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

סכום ההסתברויות על פונקציית ההסתברות חייב להיות 1.

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 מהישוב, 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 פונקציית ההסתברות של X .
4. הסיכוי לזכות במשחק מסוים הינו 0.3. אדם משחק את המשחק עד אשר הוא מנצח אך בכל מקרה הוא לא משחק את המשחק יותר מ- 4 פעמים. נגדיר את X להיות מספר הפעמים שהוא שיחק את המשחק. בנה את פונקציית ההסתברות של X .
5. חברה לניהול פרויקטים מנהלת 3 פרויקטים במקביל. הסיכוי שפרויקט א' יצליח הינו 0.7. הסיכוי שפרויקט ב' יצליח הינו 0.8. הסיכוי שפרויקט ג' יצליח הינו 0.9. נתון שהצלחת כל פרויקט בלתי תלויה זו בזו. נגדיר את X להיות מספר הפרויקטים שיצליחו. בנה את פונקציית ההסתברות של X .

6. להלן פונקציית ההסתברות של משתנה מקרי כלשהו :

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

$$k = 1, 2, \dots, 4$$

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 פונקציית 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 8 - המשתנה המקרי הבדיד - תוחלת, שונות וסטיית תקן

רקע:

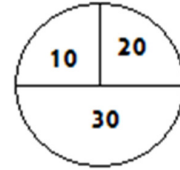
$$E(X) = \sum_i x_i P(x_i) = \mu$$
$$V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = \sigma^2$$

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.

30	20	10	x
0.5	0.25	0.25	P(x)

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

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

כדי לחשב את סטיית התקן נוציא שורש לשונות:

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

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. ל-20% 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 וסטית 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 פונקציית ההסתברות של X .

ב. חשב את התוחלת והשונות של X .

5. נתונה פונקציית ההסתברות של המשתנה המקרי X :

8	6	4	2	x
0.2		0.3		$P(x)$

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 פונקציית ההסתברות.

7. להלן ההתפלגות של משתנה מקרי X .

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

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.

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

ב. תוחלת : 3

שונות 2

שאלה 5

א.

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

ב. 5.16

שאלה 6

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

שאלה 7

2.33

פרק 9 - 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)

$$Y = aX + b \quad \text{אם}$$

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 פונקציית ההסתברות של X .
ב. חשבו את התוחלת והשונות של גובה התביעה.
ג. פרמיית הביטוח היא 4,000 ₪, מהי התוחלת ומהי השונות של רווח חברת הביטוח לביטוח הרכב הנ"ל?

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

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

כמו כן נתון שצפי מספר התשובות הנכונות בבחינה הוא 7.35.

א. השלימו את פונקציית ההסתברות.

ב. חשבו את השונות מספר התשובות הנכונות בבחינה.

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

7. להלן פונקציית הסתברות של משתנה מקרי כלשהו :

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

$$k = 1, 2, \dots, 4$$

א. מצא את ערכו של A .

ב. חשב את התוחלת והשונות של המשתנה הנחקר.

ג. חשב את $E(X^3)$

ד. חשב את התוחלת והשונות של המשתנה הבא : $\frac{X}{2} - 4$

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. $10 = A$

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 10 - תוחלת ושונות של סכום משתנים מקריים

רקע:

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

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

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

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

למשל,

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

תרגילים:

1. הרווח ממניה א' הוא עם תוחלת של 5 ושונויות 10. הרווח ממניה ב' הוא עם תוחלת של 4 ושונויות 5. ידוע שההשקעות של שתי המניות בלתי תלויות זו בזו.

מה התוחלת והשונויות 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 :

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

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

מהי התוחלת ומהי סטיית התקן של סכום הזכייה בשני המשחקים?

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

5. נתון משתנה מקרי בעל פונקציית ההסתברות הבאה :

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

0 אחרת

א. מצא את ערכו של A .

ב. חשב את התוחלת והשונויות של X .

ג. נלקחו n משתנים מקריים בלתי תלויים מההתפלגות הנ"ל. בטאו באמצעות n את תוחלת והשונויות 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

$$A = \frac{12}{25} = 0.48$$

א.

ב. תוחלת 2.92

שונות 1.1136

ג. תוחלת $n \cdot 2.92$

שונות $n \cdot 1.1136$

פרק 11 - 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: $X \sim B(n, p)$.

פונקציית ההסתברות של X :

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

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

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

$$E(X) = np \quad \text{תוחלת:}$$

$$V(X) = npq \quad \text{שונות:}$$

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

(1) חוזרים על אותו ניסוי ברנולי באופן בלתי תלוי זה בזה.

(2) חוזרים על הניסוי n פעמים.

(3) X – מוגדר כמספר ההצלחות המתקבלות בסך הכול.

דוגמה : (פתרון בהקלטה)

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

א. 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 אוכלוסייה. 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 הימורים 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 7 :

א. 0.401

ב. תוחלת : 8.025

סטיית תקן : 2.193

שאלה 3 :

א. 0.5314

ב. 0.0984

ג. 0.1143

ד. תוחלת : -18

סטיית תקן : 14.697

שאלה 8 :

7.5

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

פרק 12 - 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:

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

$$E(X) = \frac{1}{p} \quad \text{תוחלת:}$$

$$V(X) = \frac{q}{p^2} \quad \text{שונות:}$$

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?

תרגילים:

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. מטיילים 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 שישחק 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. התוויית 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?

פתרונות :

שאלה 2 :

- א. 0.009
ב. 0.0001
ג. תוחלת : 1.111
שונות : 0.1234
ד. תוחלת : 44.4
שונות : 308.5

שאלה 3 :

- א. 0.999
ב. 0.875
ג. 0.03125
ד. 1

שאלה 4 :

- א. 0.1029
ב. 9.72

שאלה 5 :

- א. 0.06
ב. 0.7176
ג. 0.0729
ד. 9.487 משחקים

שאלה 6 :

- א. 0.015
ב. 0.0215
ג. תוחלת $63\frac{1}{3}$, שונות $2777\frac{7}{9}$
ד. תוחלת $\frac{2}{3}$, שונות 1.054

פרק 13 - 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 במוצע אירועים קורים ביחידת זמן.

$$X \sim \text{pois}(\lambda)$$

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

פונקציית ההסתברות של ההתפלגות הפואסונית נתונה:

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

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

התוחלת והשונות של ההתפלגות:

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

תכונות מיוחדות של ההתפלגות:

- בהתפלגות הזו הפרמטר λ פרפורציונלי 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 בחלק 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 14 - 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 מכוונות הימורים. משחק אחד בכל מכוונה (במכוונה א' ובמכוונה ב'). הסיכוי

שלו לנצח במשחק במכוונה א' הינו 0.08 והסיכוי שלו לנצח רק במכוונה 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 פואסוני 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 פונקציית 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 "הצב" יימכר לחו"ל בסיכוי של 0.6.
- הסרט "לעולם לא" יימכר לחו"ל בסיכוי של 0.7.
- הסרט "מוות פתאומי" יימכר לחו"ל בסיכוי של 0.2.
- ידוע כי כל סרט עלה להפקה חצי מיליון שקלים. כמו כן, כל סרט הביא להכנסה של 200,000 שקלים מהטלוויזיה המקומית. במידה וסרט יימכר לחו"ל, כל סרט יימכר ב-600,000 שקלים.
- 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 התכוון רק לחלק א' של המבחן ובחלק זה בכל שאלה יש סיכוי של 0.8 שיענה נכון, בחלק השני לכל שאלה יש 4 תשובות כשרק אחת נכונה. בחלק זה הוא מנחש את התשובות.
- מהי ההסתברות שבחלק הראשון הוא יענה נכון על 7 שאלות בדיוק?
 - מהי ההסתברות שבחלק השני הוא יענה נכון על פחות מ-3 שאלות?
 - מה התוחלת ומהי השונות של מספר התשובות הנכונות בחלק הראשון?
 - מהי התוחלת ומהי השונות של מספר התשובות הנכונות בבחינה כולה?
9. יהי X משתנה מקרי המקיים $E(X) = 2$ וכן $V(X) = 1$. 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.
- חשבו את ערכו של P .
- תלמיד ניגש לבחינה עד אשר הוא עובר אותה.
- מה ההסתברות שיעבור את מבחן הנהיגה רק במבחן הרביעי?
 - מה ההסתברות שיאלץ לגשת לפחות לחמישה מבחנים בסך הכול?
 - מה התוחלת ומהי השונות של מספר המבחנים שבהם יכשל?
 - 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.

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

ב. תוחלת: 3

שונות: 2

ג. תוחלת: 1.5

שונות 0.5

שאלה 4 :

א. 0.9975

ב. 0.0172

ג. 1.0025

שאלה 5:

א. 0.375

ג. 0.6

ד. 0.282

שאלה 6:

ב. תוחלת : 1.5

שונות 0.61

ג. תוחלת : 0

סטיית תקן : 4.68

שאלה 7:

א. 0.4348

ב. 0.0923

שאלה 8 :

א. 2.013

ב. 0.5256

ג. תוחלת : 8

שונות : 1.6

ד. תוחלת : 10.5

שונות 3.475

שאלה 9 :

10

שאלה 10:

א. 0.6

ב. 0.0384

ג. 0.0256

ד. תוחלת : 0.67

שונות : 1.11

ה. 0.24

שאלה 12:

ב. $\frac{1-2p^2}{p}$

ג. $0 < p < \sqrt{\frac{1}{2}}$

פרק 15 - 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:

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

כמו כן:

$$p(a < X < b) = F(b) - F(a)$$

$$p(X > t) = 1 - F(t)$$

תוחלת של משתנה רציף:

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

שונות של משתנה רציף:

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

תוחלת של פונקציה של X:

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

אחוזונים:

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

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

ריענון מתמטי:

נוסחאות לחישוב שטחים:

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

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

משוואת קו ישר:

$$y = mx + n$$

m = שיפוע.

n = נקודת החיתוך עם ציר ה-y.

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

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

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

נוסחאות - אינטגרלים

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

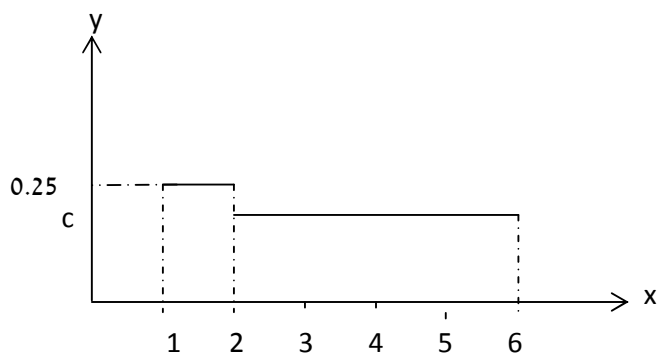
$$\int \cos f \cdot f' dx = \sin(f) + c$$

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

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

תרגילים:

1. X הינו משתנה רציף עם פונקציית צפיפות כמוצג בשרטוט:



א. מצא את ערכו של c .

ב. בנה את פונקציית ההתפלגות המצטברת.

ג. חשבו את ההסתברויות הבאות:

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 שפונקציית הצפיפות שלו היא:

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

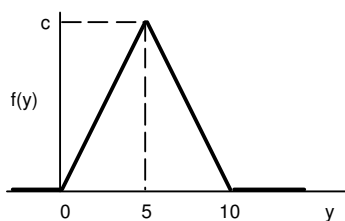
ידוע ש- $P(0 < X < 1) = 1/4$

א. מצאו במפורש את פונקציית הצפיפות של X .

ב. מצאו את החציון של X .

ג. מה הסיכוי ש- X קטן מ- 0.5 ?

3. נתונה פונקצית צפיפות של משתנה מקרי Y :



א. מצאו את c .

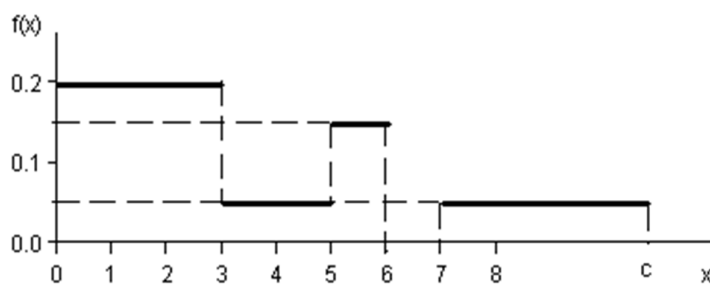
ב. מצאו את פונקצית ההתפלגות המצטברת של Y .

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

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 :

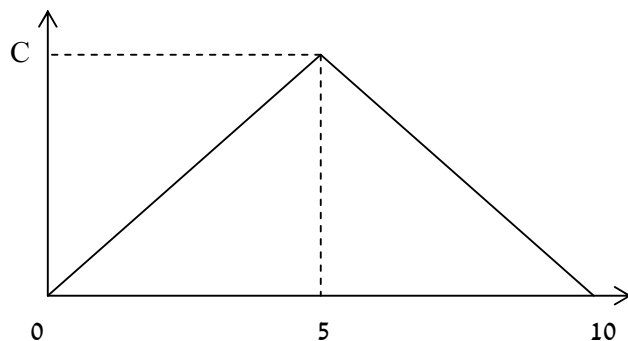


א. מצאו ערך c שעבורו תתקבל פונקצית צפיפות.

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.

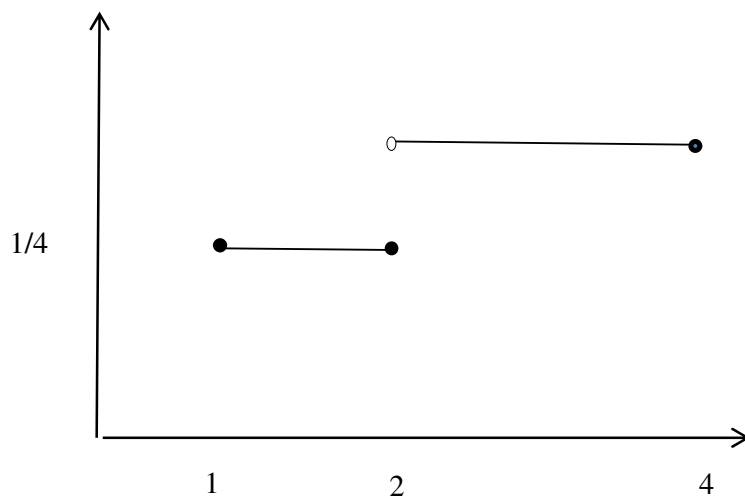
ב. חשב את $P(x > 5 | x > 2)$.

ג. מה התוחלת ומהי השונות של X?

8. פונקצית הצפיפות של משתנה מקרי רציף X :
 $f(x) = 0.5 \cdot e^{2x}$
 $-\infty \leq X \leq \ln(c)$

- 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 :

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

- א. מה ההסתברות שזמן הייצור של מוצר A אקראי יהיה קטן מ 20 דקות?
- ב. מה ההסתברות שזמן הייצור של מוצר A אקראי יהיה בדיוק חצי שעה?
- ג. נבחרו חמישה מוצרים אקראיים מסוג A. מה תוחלת מספר המוצרים שזמן הייצור שלהם יהיה גדול מ 20 דקות?

11. זמן ההמתנה בדקות של לקוח בתור למכולת השכונתית מתפלג עם פונקציית 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 ש-90% מהלקוחות מחכים מתחתיו?

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

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

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

13. נתונה פונקציית הצפיפות הבאה :

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

א. מצאו את ערכו של K.

ב. מצאו את פונקציית ההתפלגות המצטברת.

ג. חשבו $p(x > 2.5)$

14. להלן משתנה מקרי בעל פונקציית צפיפות הבאה :

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

$$a \leq x \leq b$$

א. 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. $\frac{1}{3}$

שאלה 2 :

א. $c=0.5$ $b=2$
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. $\frac{1}{e^2}$
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 14 :

ב. תוחלת :

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

השונות :

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

שאלה 13 :

א. $\frac{1}{6}$

ג. 0.229

פרק 16 - 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).

$$X \sim \exp(\lambda) \text{ כאשר } \lambda > 0$$

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:

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

פונקציית ההתפלגות המצטברת היא:

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

התוחלת:

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

השונות:

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

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

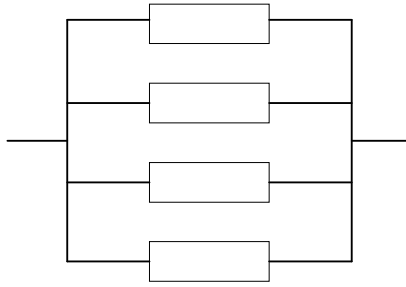
דוגמה: (פתרון בהקלטה)

- 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?

תרגילים:

1. הזמן שלוקח במערכת עד שתקלה מתרחשת מתפלג מעריכית עם תוחלת של 0.5 שעה.
 - א. מה ההסתברות 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 סטיית 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 100 שעות.

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

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

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

מה התנאי שבו יהיה כדאי להוסיף את הרכיב למערכת?

פתרונות:

שאלה 1 :

א. 0.368

ב. 0.865

ג. 0.347

שאלה 2:

א. 24 שעות

ב. 0.632

ג. 0.135

שאלה 3:

א. 0.393

ב. 0.239

ג. 0.513

ד. 69.08

שאלה 4:

א. 0.264

ב. 0.368

ג. 0.233

שאלה 5:

א. 0.8403

ב. $A0.0588 > K$

פרק 17 - 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.

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

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

כדי לחשב הסתברויות בהתפלגות נורמלית יש לחשב את השטחים הרלבנטיים 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:

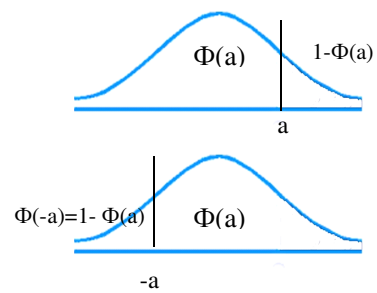
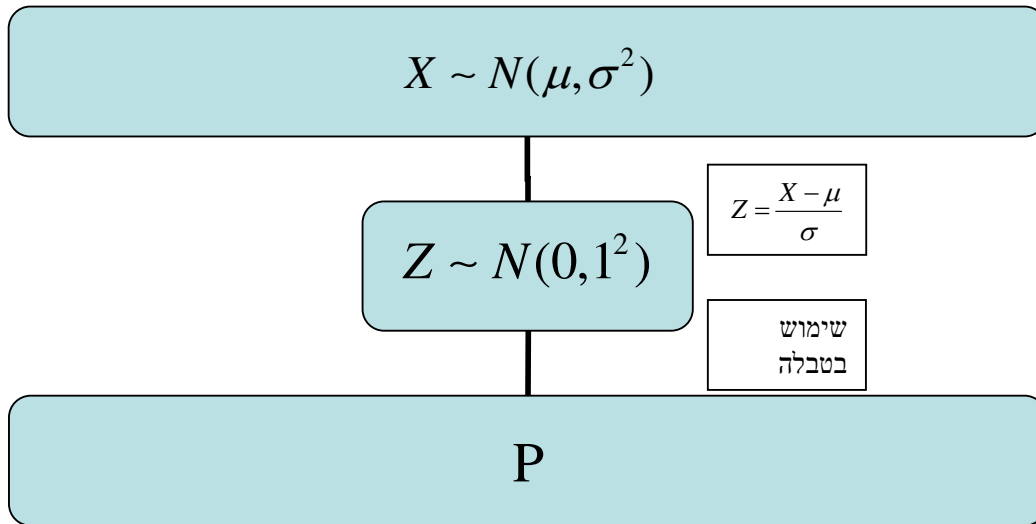
$$Z = \frac{X - \mu}{\sigma}$$

אחרי תקנון מקבלים ערך הנקרא ציון תקן.

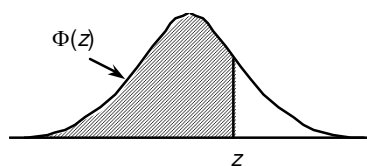
ציון התקן משמעו בכמה סטיות תקן הערך סוטה מהממוצע.

לאחר חישוב ציון התקן של ערך מסוים נעזרים בטבלה של ההתפלגות הנורמלית הסטנדרטית לחישוב השטח הרצוי.

ובאופן כללי נתאר את הסכמה הבאה :



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



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

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

דוגמה: (הפתרון בהקלטה)

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

- 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 ש-90% 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 וסטית תקן של 8 ק"ג.

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

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

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

5. ציוני מבחן אינטליגנציה מתפלג נורמלית עם ממוצע 100 ושונות 225 .
- א. מה העשירון העליון של הציונים במבחן האינטליגנציה?
 - ב. מה העשירון התחתון של ההתפלגות?
 - ג. מהו הציון ש- 20% מהנבחנים מקבלים מעליו?
 - ד. מהו האחוזון ה- 20?
 - ה. מהו הציון ש- 5% מהנבחנים מקבלים מתחתיו?
6. נפח משקה בבקבוק מתפלג נורמלית עם סטיית תקן של 20 מ"ל, נתון ש-33% 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 בבקבוק 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 סטיית בתקן של אורך חיי מכשיר?
 - ג. מה הסיכוי שמכשיר אקראי יחיה פחות מ- 460 שעות?
 - ד. מהו המאון העליון של אורך חיי מכשיר?
 - ה. 1% מהמכשירים בעלי אורך החיים הקצר ביותר נשלח למעבדה לבדיקה מעמיקה. מהו אורך החיים המקסימלי לשליחת מכשיר למעבדה?

8. להלן שלוש התפלגויות נורמליות של שלוש קבוצות שונות ששורטטו באותה מערכת צירים. ההתפלגויות מוספרו כדי להבדיל ביניהן.



- א. לאיזו התפלגות הממוצע הגבוה ביותר?
- ב. במה מבין המדדים הבאים התפלגות 1 ו 2 זהות?
 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 וסטיית תקן של 5 דקות.

- א. מה ההסתברות שמשך הנסיעה של האדם לעבודתו יהיה לפחות שלושת רבעי השעה?
- ב. אדם יצא לעבודתו בשעה 08:10 מביתו. הוא צריך להגיע לעבודתו בשעה 09:00. מה הסיכוי שיאחר לעבודתו?
- ג. אם ידוע שזמן נסיעתו לעבודה היה יותר משלושת רבעי השעה. מה ההסתברות שזמן הנסיעה הכולל יהיה פחות מ- 50 דקות?
- ד. מה הסיכוי שבשבוע (חמישה ימי עבודה) בדיוק פעם אחת יהיה זמן הנסיעה לפחות שלושת רבעי השעה?

10. ההוצאה החודשית לבית אב בעיר "טרירה" מתפלגת נורמלית עם ממוצע של 2000 דולר וסטית תקן של 300 דולר. בחרו באקראי 5 בתי אב. ההסתברות שלפחות אחד מהם מוציא בחודש מעל ל- T דולר היא 0.98976.
- א. מה ערכו של T ?
- ב. מה הסיכוי שההוצאה החודשית של בית אב בעיר תהיה לפחות סטיית תקן אחת מעל T ?
- ג. מסתבר שנפלה טעות בנתונים, ויש להוסיף 100 דולר להוצאות החודשית של כל בתי האב בעיר. לאור זאת, מה ההסתברות שההוצאה החודשית של בית אב נמוכה מ-1800 דולר?
11. אורך שיר אקראי המשודר ברדיו מתפלג נורמלית עם תוחלת של 3.5 דקות וסטית תקן של שלושים שניות.
- א. מה ההסתברות שאורך של שיר אקראי המנוגן ברדיו יהיה בין 3 ל 2.5 דקות?
- ב. מהו הטווח הבין רבעוני של אורך שיר המשודר ברדיו?
- ג. ביום מסוים מנוגנים 200 שירים ברדיו. כמה שירים מתוכם תצפה שיהיו באורך הנמוך מ 3.5 דקות?
- ד. בשעה מסוימת שודרו 8 שירים. מה ההסתברות שרבע מהם בדיוק היו ארוכים מ-4 דקות והיתר לא?

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

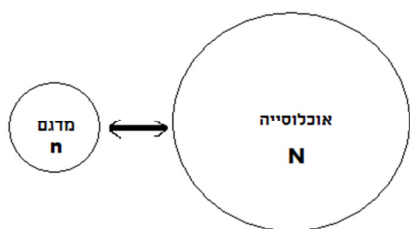
- א. 0.1359
ב. 0.675
ג. 100
ד. 0.25

פרק 18 - הסקה סטטיסטית - הקדמה

רקע:

אוכלוסייה – קבוצה שאליה מפנים שאלה מחקרית.

למשל, חברת תרופות שמעוניינת לפתח תרופה למחלת הסוכרת מתעניינת באוכלוסיית חולי הסוכרת בעולם.



מדגם – חלק מתוך האוכלוסייה.

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

במקרים רבים אין אפשרות לחקור את כל האוכלוסייה כיוון שאין גישה לכולה, היא גדולה מידי, או מוגבלים בזמן ובאמצעים טכניים ולכן מבצעים מדגם במטרה לבצע הסקה סטטיסטית מהמדגם לאוכלוסייה.

הדגימה בקורס 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:

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

פרמטר הוא גודל קבוע גם אם 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?

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?
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 19 - 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 או 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?

ב. דגימה מהתפלגות נורמאלית

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

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

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

דוגמה: (פתרון בהקלטה)

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

ג. משפט הגבול המרכזי

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

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

דוגמה: (פתרון בהקלטה)

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

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 הטלויזיה למשפחה בישוב מסוים :

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

נגדיר את x להיות מספר המקלטים של משפחה אקראית.

א. בנו את פונקציית 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 מתוחלת הגבהים בפחות מ-5 ס"מ?

ד. מהו הגובה שבהסתברות של 90% הגובה הממוצע של המדגם יהיה נמוך ממנו?

6. הזמן הממוצע שלוקח לאדם להגיע לעבודתו 30 דקות עם שונות של 16 דקות רבועות. האדם נוסע לעבודה במשך שבוע 5 פעמים. לצורך 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 מ-755 סמ"ק?

ג. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה לפחות 755 סמ"ק?

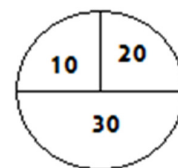
ד. בקבוקי היין שבארגז נמזגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?

8. משתנה מתפלג נורמאלי עם תוחלת 80 וסטיית תקן 4 .

א. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?

ב. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה שגודל המדגם הוא 16?
ג. הסבר את ההבדל בתשובות של שני הסעיפים.

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



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.

א. בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.

ב. מה התוחלת ומה השונות של סכום הזכייה?

ג. אם האדם ישחק את המשחק 5 פעמים מה התוחלת ומה השונות של ממוצע סכום הזכייה בחמשת המשחקים?

ד. אם האדם משחק את המשחק 50 פעם מה ההסתברות שבסה"כ יזכה ב-1050 ₪ ומעלה?

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

11. 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 פונקציית ההסתברות הבאה :

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

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

14. נתון ש $X \sim N(\mu, \sigma^2)$ דגמו 5 תצפיות מאותה התפלגות והתבוננו במוצע המדגם $\bar{X}$:

לכן $P(\bar{X} > \mu)$ יהיה : (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 ושונויות 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.

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

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

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

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

שאלה 3

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

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

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

שאלה 8

א. 0.5468

ב. 0.6826

שאלה 9

א.

30	20	10	
0.5	0.25	0.25	P(x)

ב. התוחלת: 22.5

השונות: 68.75

ג. התוחלת: 22.5

השונות: 13.75

ד. 0.8997

שאלה 10

0.0475

שאלה 11

0.1814

שאלה 12

א. 0.9772

ב. 0.0228

ג. 271

שאלה 14

התשובה ב

שאלה 15

התשובה ד

שאלה 16

התשובה ג

שאלה 17

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:

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

$$X \sim N(\mu, \sigma^2) \quad \text{אם}$$

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

ג. משפט הגבול המרכזי:

$$E(X) = \mu$$

$$V(X) = \sigma^2 \quad \text{אם } x \text{ מתפלג כלשהו וידוע}$$

אזי עבור מדגם מספיק גדול (לפחות 30)

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

דוגמה: (פתרון בהקלטה)

בעיר מסוימת המשכורת הממוצעת של עובד הינה 8000 ₪. עם סטיית תקן של 2000 ₪. נדגמו 100 עובדים מהעיר שמפקידים את משכורותיהם לסניף בנק.

- 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?

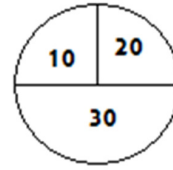
3. בספר כלשהו 500 עמודים. קצב הקריאה הממוצע הוא עמוד אחד ב 4 דקות עם סטיית תקן של 1 דקות.
 - א. מה ההסתברות לסיים את הפרק הראשון (40 עמודים) תוך שעתיים וחצי?
 - ב. מהו האחוזון ה-95 לזמן סיום קריאת הספר?

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

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

- נניח שמספר המכוניות ליחידת דיור בלתי תלויים זה בזה ועם אותה פונקציית 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 $i = 1, 2, \dots, 100$,

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

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

פרק 20 - 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: $|\hat{\theta} - \theta|$ - ההפרש בין האומד לאמת(הפרמטר).

דוגמה: (פתרון בהקלטה)

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

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 אומד חסר הטיות 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 אומד חסר הטיות 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 הטיות 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.

תרגילים:

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 האומדן 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 האומדן?
- 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:
- $$\sum_{i=1}^{20} X_i^2 = 1502.2 \quad \sum_{i=1}^{20} X_i = 162$$
- 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 תהיה:

א. 3.

ב. 2.5.

ג. 1.581.

ד. אין מספיק נתונים לדעת.

7. במדגם מקרי, מתי סכום ריבועי הסטיות מהממוצע, $\sum_{i=1}^n (x_i - \bar{x})^2$, מחולק ב- $n-1$?

א. כאשר n קטן.

ב. כאשר תצפיות המדגם אינן בלתי תלויות.

ג. כאשר האוכלוסייה אינה מתפלגת נורמאלית.

ד. כאשר מעוניינים באומדן חסר הטיה לשונות האוכלוסייה ממנה הוצא המדגם.

ה. כאשר מעוניינים לחשב את שונות התפלגות הדגימה של ממוצע המדגם.

8. $X_1, X_2, \dots, X_{16}$ מדגם מקרי מתוך אוכלוסייה בעלת ממוצע μ לא ידוע ושונות

$\sigma^2 = 64$. טעות התקן של האומדן ל- μ היא:

א. 16

ב. 8

ג. 4

ד. 2

9. מהו אומדן חסר הטיה?

- א. אומדן שערכו שווה לממוצע התפלגות הדגימה שלו.
- ב. אומדן שערכו שווה לערך הפרמטר באוכלוסייה.
- ג. אומדן שממוצע התפלגות הדגימה שלו שווה לערך הפרמטר באוכלוסייה.
- ד. אומדן שהסיכוי שערכו יהיה גבוה מערך הפרמטר באוכלוסייה שווה לסיכוי שיהיה נמוך ממנו.

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 21 - רווח סמך לתוחלת (ממוצע האוכלוסייה)

רווח סמך כששונות האוכלוסייה ידועה

רקע:

ממוצע המדגם הוא אומדן לממוצע האוכלוסייה, אך לא באמת ניתן להבין ממנו על גודלו של ממוצע האוכלוסייה. ההסתברות שממוצע המדגם יהיה בדיוק כמו הממוצע האמיתי הוא אפסי. מה שנהוג לעשות כדי לאמוד את ממוצע האוכלוסייה זה לבנות רווח סמך. נבנה מרווח בטחון שהסיכוי שהפרמטר μ ייכלל בתוכו הוא $1-\alpha$.

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

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

A - גבול התחתון של רווח הסמך

B - הגבול העליון של רווח הסמך

$L = B - A$ - אורך רווח הסמך

דוגמה : (פתרון בהקלטה)

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

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.

שגיאת האמידה המקסימלית:

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

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

דוגמה: (פתרון בהקלטה)

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

קשרים מתמטיים ברווח הסמך:

- אורך רווח הסמך הוא פעמיים שגיאת האמידה המקסימלית: $L = 2\varepsilon$.
- ממוצע המדגם נופל תמיד באמצע רווח הסמך: $\bar{X} = \frac{A+B}{2}$
- ככל שמספר התצפיות (n) גבוה יותר, כך יש יותר אינפורמציה ולכן האומדן 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 נורמאלי 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 ב-50%?
 - 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. $\bar{X} = 6$
- ג. $\bar{X} = 51$
- ד. אורך רווח הסמך הינו 3.
10. איזה מהגורמים הבאים אינו משפיע על גודלו של רווח בר סמך, כאשר שונות האוכלוסייה ידועה? (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:

$$\bar{x} = 74$$

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

פתרונות :

שאלה 2

$$4920.6 < \mu < 4979.4$$

שאלה 3

$$\text{א. } 223.42 < \mu < 236.58$$

$$\text{ב. } 222.16 < \mu < 237.84$$

שאלה 5

$$\text{א. } 102$$

$$\text{ב. } 3$$

$$\text{ג. } 0.9544$$

שאלה 6

$$\text{א. } 3.58 < \mu < 4.42$$

$$\text{ב. יקטן פי } 2$$

$$\text{ג. גדל}$$

שאלה 7

$$\text{א. } 87$$

$$\text{ב. } 5$$

$$\text{ג. } 0.9544$$

שאלה 8

$$\text{א. } 139$$

$$\text{ב. } 21 < \mu < 25$$

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 האמיתי 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 μ ברמת ביטחון של 0.95 כך שהאורך של הרווח יהיה 0.5σ . מהו גודל המדגם הנדרש?

פתרונות :

שאלה 1

780

שאלה 2

א. 139

ב. הדבר יקטין את ε פי 2.

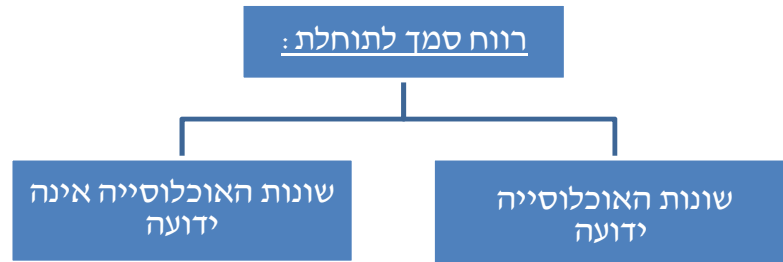
שאלה 3

$n = 62$

רווח סמך לתוחלת (ממוצע האוכלוסייה) כששונות האוכלוסייה אינה ידועה

רקע:

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



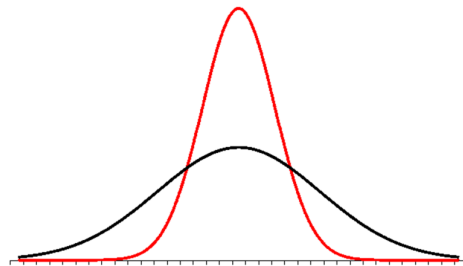
בפרק זה נעסוק במקרה ששונות האוכלוסייה (σ^2) אינה ידועה לנו. מקרה יותר פרקטי.

התנאי: $X \sim N$ או שהמדגם גדול

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

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

התפלגות T:



הינה התפלגות סימטרית פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים. התפלגות T תלויה במושג שנקרא דרגות חופש. דרגות החופש הן $df=n-1$. ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה. כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דיוגמה : (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'.

פתרון :

$$4.39 < \mu < 5.51$$

תרגילים:

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

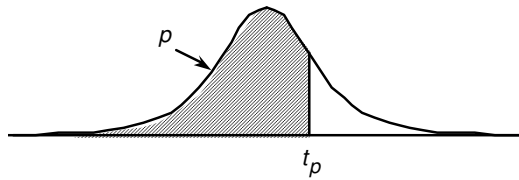
ב. סטטיסטיקאי ב

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?



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 22 - רווח סמך להפרש תוחלות ממדגמים בלתי תלויים

כשונות האוכלוסייה ידועות

רקע:

מטרה: לאמוד את פער התוחלות: $\mu_1 - \mu_2$, כלומר ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$

התנאים לבניית רווח הסמך:

1. σ_1^2, σ_2^2 ידועות.

2. $X_1, X_2 \sim N$ או $n_1, n_2 > 30$

3. שני מדגמים בלתי תלויים.

רווח סמך:

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

אם הערך אפס נופל בגבולות רווח הסמך נגיד שבביטחון של $1 - \alpha$ לא קיים הבדל בין התוחלות.

דוגמה: (פתרון בהקלטה)

נדגמו 100 תושבים מאזור a והמשכורת הממוצעת הייתה שם 9200 ₪.

כמו כן נדגמו 120 תושבים מאזור b וממוצע המשכורות שהתקבל שם 8700 ₪.

לצורך פתרון נניח שסטיית התקן של המשכורות באוכלוסיית שני האזורים היא 1800 ₪.

אמדו ברמת סמך של 90% את הפרש השכר הממוצע בין אזור a לאזור b.

תרגילים:

1. מעוניינים לבדוק האם קיים הבדל בין ממוצע ציוני הפסיכומטרי של חיילים לממוצע ציוני הפסיכומטרי של תלמידי תיכון. ידוע שציוני הפסיכומטרי מתפלגים נורמאלית עם סטיית תקן 100. במדגם של 16 נבחנים חיילים התקבל ממוצע 543. במדגם של 20 תלמידי תיכון התקבל ממוצע 508. בנו רוח סמך לפער תוחלות הציונים בין חיילים לתלמידי תיכון ברמת סמך של 90%. מה ניתן להסיק מרווח סמך זה?
2. ציוני I.Q. מתוכננים כך שיתפלגו נורמאלית עם סטיית תקן של 15. במדגם של 20 נבחנים ישראלים התקבל ממוצע ציונים 104. במדגם של 23 נבחנים אמריקאיים התקבל ממוצע ציונים 99.
א. בנו רוח סמך ברמת סמך של 95% לפער בין ישראל לארה"ב בממוצע הציונים במבחן ה-IQ.
ב. האם קיים הבדל בין ישראלים לאמריקאים מבחינת ממוצע הציונים?
3. חברה להנדסת בניין מעוניינת להשוות ברמת הקשיות של שני סוגי ברגים. ידוע שרמת הקשיות של ברגים מתפלגת נורמלית עם סטיית תקן של 4 יחידות. במדגם של 15 ברגים מסוג א' התקבל רמת קשיות ממוצעת של 28 יחידות ובמדגם של 12 ברגים מסוג ב' התקבל רמת קשיות ממוצעת של 25. עבור אילו רמות בטחון יקבע שאין הבדל בין שני סוגי הברגים מבחינת ממוצע רמת הקשיות שלהם?

פתרונות :

שאלה 1

$(-20, 90)$

שאלה 3:

רמות בטחון הגבוהות מ: 0.9476

כששונויות האוכלוסייה אינן ידועות אך שוות והמדגמים בלתי תלויים

רקע:

מטרה: לאמוד את פער התוחלות: $\mu_1 - \mu_2$, כלומר ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$

התנאים לבניית רווח הסמך:

1. $\sigma_1^2 = \sigma_2^2$.

2. $X_1, X_2 \sim N$

3. מדגמים בלתי תלויים.

השונויות המשוקללת: כיוון שאנו מניחים שבין שתי האוכלוסיות השונויות שוות אנו אומדים את השונויות הזו על ידי שקלול שתי השונויות של שני המדגמים על ידי הנוסחה הבאה:

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

דרגות החופש: $d.f = n_1 + n_2 - 2$

רווח סמך:

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

אם הערך אפס נופל בגבולות רווח הסמך נגיד שבביטחון של $1 - \alpha$ לא קיים הבדל בין התוחלות.

דוגמה: (פתרון בהקלטה)

מחקר מעוניין לבדוק האם קיים הבדל בין תל אביב לבאר שבע מבחינת ההכנסה הממוצעת של אקדמאים. להלן תוצאות המדגם שנעשה:

באר שבע	תל אביב	
10	20	מספר האקדמאים
9500	11,000	ממוצע הכנסות של אקדמאים
250	200	סטיית התקן של הכנסות אקדמאים

בנו רווח סמך ברמת ביטחון של 90% להפרש תוחלות ההכנסה בשני האזורים.
הניחו שהשכר מתפלג נורמלית עם אותה שונות בכל אחד מהאזורים.
פתרון : (1357,1643)

תרגילים:

1. נדגמו 15 ישראלים ו-15 אמריקאים.

כל הנדגמים נגשו למבחן IQ. להלן תוצאות המדגם:

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

מצאו רווח סמך ברמת סמך של 95% לסטייה בין ממוצע הציונים בישראל לממוצע הציונים בארה"ב. רשמו את כל ההנחות הדרושות לצורך פתרון התרגיל.

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

25	21	20	22	X
12	17	25	18	Y

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

פרק 23 - רווח סמך לתוחלת ההפרש במדגם מזווג

רקע:

מדגם מזווג: מדגם אחד שבו יש n צמדים.

כל תצפית במדגם תנפק זוג ערכים: X ו- Y .

ניצור משתנה חדש:

$$D = x - y$$

הפרמטר שנרצה לאמוד: μ_D

התנאים לבניית רווח הסמך:

• $x, y \sim N$

• המדגם מזווג

נוסחת רווח הסמך:

$$\bar{D} \pm t_{1-\frac{\alpha}{2}}^{n-1} \frac{S_D}{\sqrt{n}}$$

כאשר דרגות החופש: $d.f = n - 1$

דוגמה: (פתרון בהקלטה)

מעוניינים לבדוק האם יש הבדל בין מהירות הריצות של שתי תוכנות מחשב.

לקחו 5 קבצים אקראיים והריצו אותם בשתי התוכנות:

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

הניחו כי זמני הריצות מתפלגים נורמלית.

מצאו רווח סמך של 95% להפרש תוחלת הזמן בין שתי התוכנות.

תרגילים:

1. נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'. להלן הציונים בסמסטר א' ו- ב' :

סמסטר א	סמסטר ב
74	80
68	84
90	87
75	76
82	100

נניח שהציונים מתפלגים נורמאלית.

א. בנו רווח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סמסטר א לבין סמסטר ב.

ב. האם על סמך רווח הסמך קיים הבדל בין הסמסטרים מבחינת תוחלת הציונים?

ג. מה צריך לשנות בנתונים כדי שהמדגמים יהיו בלתי תלויים?

2. במטרה לבדוק האם קיים הבדל בין קווי זהב לבזק מבחינת ממוצע המחירים לשיחות בינ"ל.

נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה. להלן התוצאות :

המדינה	בזק-X	קווי זהב-Y
ארה"ב	1.5	1.4
קנדה	2.1	2
הולנד	2.2	1.9
פולין	3	3.1
מצרים	3.5	3.3
סין	3.2	3.2
יפן	4.2	4.2

בהנחה והמחירים מתפלגים נורמלית עבור כל חברה בנו רווח סמך ברמת סמך של 90% לתוחלת הפרש המחירים של שתי החברות.

פרק 24 - רווח סמך לשונות וסטיית תקן

רקע:

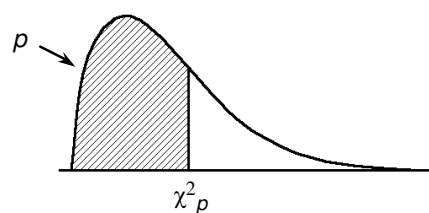
בפרק זה נדון על בניית רווח סמך לשונות האוכלוסייה.

התנאי לבניית רווח הסמך: המשתנה הנחקר מתפלג נורמלית, למרות שנהוג לא לדרוש את התנאי הזה אם המדגם מספיק גדול.

רווח הסמך יתבסס על התפלגות הנקראת חי בריבוע.

התפלגות זו היא התפלגות אסימטרית חיובית המתחילה מהערך אפס ותלויה בדרגות חופש.

דרגות החופש במקרה זה יהיו: $n-1$



$$\text{רווח הסמך לשונות: } \frac{(n-1)S^2}{\chi^2_{1-\alpha/2, n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{\alpha/2, n-1}}$$

$$\text{כאשר } S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n \cdot \bar{X}^2}{n-1} \text{ אומד לשונות הלא-ידועה.}$$

אם נרצה לבנות רווח סמך לסטיית תקן אז נוציא שורש לרווח סמך לשונות.

דוגמה: (פתרון בהקלטה)

זמן התגובה מתפלג נורמאלית. במטרה לאמוד את שונות זמן התגובה נדגמו 4 תצפיות. להלן התוצאות בשניות: 4.7, 5.2, 4.6, 5.3. בנו רווח סמך, ברמת סמך של 95% לשונות זמן התגובה באוכלוסייה.

פתרון:

$$0.039 < \sigma^2 < 1.708$$

תרגילים :

1. חמישה מטופלים קבלו תרופה מסוימת. בדקו לכל מטופל את זמני התגובה שלו. להלן הזמנים שהתקבלו בדקות: 18,17,21,26,28.

בהנחה וזמני התגובה מתפלגים נורמאלית, בנו רווח סמך ברמת סמך של 95% לשונות זמן התגובה.

2. נדגמו 20 ימים אקראיים מחודשי יולי-אוגוסט ונמדדה בהם הטמפ' במעלות צלזיוס בת"א. במדגם התקבל טמפ' ממוצעת 30.8 וסטיית תקן מדגמית 1.1. בהנחה והטמפ' מתפלגת נורמאלית:

- א. בנו רווח סמך לתוחלת הטמפ' בחודשים אלה בת"א ברמת סמך של 95%.
- ב. בנו רווח סמך לסטיית התקן של הטמפ' בחודשים אלה בת"א ברמת סמך של 95%.

3. ציוני IQ בארה"ב מתפלגים נורמאלית עם ממוצע 100 וסטיית תקן 5. נבחנו 20 נבחנים ישראלים במבחן ה-IQ. להלן התוצאות שהתקבלו:

$$\sum_{i=1}^{20} X_i = 2080$$

$$\sum_{i=1}^{20} X_i^2 = 218,220$$

נניח שגם בישראל הציונים מתפלגים נורמאלית.

א. מצאו אומדנים לממוצע הציונים בישראל ולשונות הציונים בישראל באמצעות אומדנים חסרי הטיה.

ב. אמדו ברמת ביטחון של 95% את תוחלת הציונים של נבחנים בישראל.

ג. אמדו ברמת סמך של 90% את סטיית התקן של הציונים של נבחנים ישראלים.

ד. על סמך הסעיפים הקודמים, האם בישראל ממוצע הציונים וסטיית התקן של הציונים שונה מבארה"ב? הסבירו.

4. באוכלוסייה מסוימת נדגמו 10 תצפיות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{10} X_i = 750$$

$$\sum_{i=1}^{10} (X_i - \bar{X})^2 = 900$$

נתון ש $X_i \sim N(\mu, \sigma^2)$

א. בנו רווח סמך ל- μ ברמת סמך של 95%.

ב. בנו רווח סמך ל- σ^2 ברמת סמך של 95%.

פתרונות :

שאלה 2

א. $30.285 < \mu < 31.315$

ב. $0.837 < \sigma < 1.607$

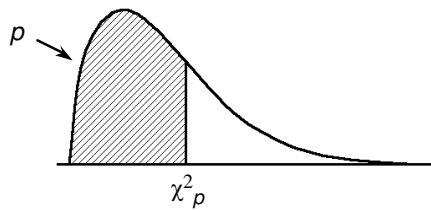
תשובה 3

א. לממוצע 104, לשונות 100.

ב. $99.32 < \mu < 108.68$

ג. $7.94 < \sigma < 13.7$

טבלת התפלגות חי-בריבוע – ערכי החלוקה χ^2_p



p

df	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995
1	0.0 ⁴ 393	0.0 ³ 157	0.0 ³ 982	0.0 ² 393	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

פרק 25 - רווח סמך ליחס שונות

רקע:

נרצה לאמוד את ההבדל בין שתי שונות משתי אוכלוסיות שונות.

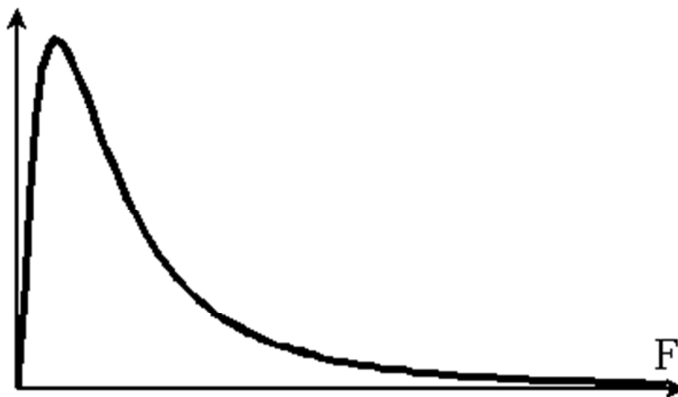
הפרמטר יהיה : $\frac{\sigma_1^2}{\sigma_2^2}$: כלומר היחס בין השונות.

התנאים:

• $X_1, X_2 \sim N$ או מדגמים גדולים.

• מדגמים בלתי תלויים.

רווח הסמך יבנה על סמך התפלגות הנקראת התפלגות F. התפלגות זו היא אסימטרית חיובית ומושפעת משתי דרגות החופש זו של המונה וזו של המכנה.



$$df_1 = n_1 - 1$$

$$df_2 = n_2 - 1$$

רווח הסמך יהיה :

$$\frac{s_1^2}{s_2^2} \frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} F_{1-\frac{\alpha}{2}}(n_2-1, n_1-1)$$

דוגמה: (פתרון בהקלטה)

מחקר סוציולוגי מעוניין לחקור את הרגלי הבילויים בקבוצות גיל שונות :
במדגם שנעשה על סטודנטים בגילאי 21-26 התקבל אומד חוסר הטיה לשונות ההוצאה החודשית
על בילויים 10,000. כמות הסטודנטים שנדגמו 16.
במדגם שנעשה על 11 מבוגרים בשנות השלושים התקבל אומד חסר הטיה לשונות ההוצאה
החודשית על בילויים 490,000.
נניח שההוצאה החודשית לבילוי בכל קבוצת גיל מתפלג נורמאלית.
בנו רווח סמך ברמת סמך של 95% ליחס בין השונות.

תרגילים:

1. בתחום הבינוי משתמשים בשני סוגי מתכות: מתכת A ומתכת B. מחקר מעוניין לבדוק האם קיים הבדל בין שני סוגי המתכות מבחינת שונות החוזק שלהן. דגמו מספר יחידות. מתכת מכל סוג והתקבלו התוצאות הבאות:

סוג המתכת	A	B
N	8	10
$\sum X_i$	16	30
$\sum X_i^2$	60	198

- יש להניח שרמת החוזק של המתכות מתפלגת נורמאלית.
- א. בנו רווח סמך ליחס השונויות של רמות החוזק בין שני סוגי המתכות ברמת סמך של 90%.
- ב. בנו רווח סמך ליחס סטיות התקן של רמות החוזק בין שני סוגי המתכות ברמת סמך של 90%.
- ג. האם בבטחון של 90% ניתן לומר שסטיות התקן של שני סוגי המתכות שונות?

2. מעוניינים להשוות בין נשים וגברים מבחינת השונויות בזמנים שלהם לבצע משימה מסוימת. במדגם של 10 גברים התקבלו התוצאות הבאות לגבי זמני ביצוע המשימה:

$$\sum (y_i - \bar{y})^2 = 204$$

במדגם של 13 נשים התקבלו התוצאות הבאות:

$$\sum (x_i - \bar{x})^2 = 200$$

- אמוד ברמת בטחון של 95% פי כמה גדולה השונויות של הגברים באוכלוסייה מהשונויות של הנשים. מה יש להניח לצורך פתרון?

טבלת התפלגות F ערכי החלוקה F_p של התפלגות $F(m, n)$
 m — דרגות חופש המונה n ; דרגות חופש המכנה

		m										
p	n	1	2	3	4	5	6	7	8	9	10	12
.95	1	161	200	216	225	230	234	237	239	241	242	244
.975		648	800	864	900	922	937	948	957	963	969	977
.99		4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6106
.995		16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426
.95	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41
.975		38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41
.99		98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.42
.995		198.50	199.01	199.16	199.24	199.30	199.33	199.36	199.38	199.39	199.39	199.42
.95	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74
.975		17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34
.99		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.05
.995		55.55	49.80	47.47	46.20	45.39	44.84	44.43	44.13	43.88	43.68	43.39
.95	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91
.975		12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75
.99		21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37
.995		31.33	26.28	24.26	23.15	22.46	21.98	21.62	21.35	21.14	20.97	20.70
.95	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68
.975		10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52
.99		16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89
.995		22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.38
.95	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00
.975		8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37
.99		13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72
.995		18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.03
.95	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57
.975		8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67
.99		12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47
.995		16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.18
.95	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28
.975		7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20
.990		11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67
.995		14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01

		m										
	p	n	1	2	3	4	5	6	7	8	9	10
.95	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07
.975		7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87
.99		10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11
.995		13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23
.95	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91
.975		6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62
.99		10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71
.995		12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66
.95	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69
.975		6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28
.99		9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16
.995		11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91
.95	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48
.975		6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96
.99		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67
.995		10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25
.95	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28
.975		5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68
.99		8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23
.995		9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68
.95	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09
.975		5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41
.99		7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84
.995		9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18
.95	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92
.975		5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17
.99		7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50
.995		8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74
.95	120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83
.975		5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05
.99		6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34
.995		8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54
.95	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75
.975		5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94
.99		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18
.995		7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36

פרק 26 - תרגול מסכם ברווחי סמך

1. מהירות הגלישה באינטרנט במקום מסוים מתפלגת נורמאלית. בדקו את מהירות הגלישה ב-30 זמנים אקראיים. מהירות הגלישה נמדדה ב-Mbps. מהירות מתחת ל-10 Mbps מוגדרת על ידי החברה כנמוכה.
- התוצאות שהתקבלו במדגם: ממוצע היה 87 עם סטיית תקן 17 ו-12 פעמים המהירות הייתה נמוכה. בנו רווחי סמך ברמת סמך של 95% לפרמטרים הבאים:
- א. תוחלת מהירות הגלישה.
- ב. סטיית תקן של מהירות הגלישה.
- ג. הסיכוי שמהירות הגלישה תהיה נמוכה.

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

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

- א. תנו רווח סמך לממוצע מספר כוסות הקפה שאנשים נוהגים לשתות ביום. $\alpha = 0.05$
- ב. אדם השותה לפחות 4 כוסות קפה ביום נקרא "מכור לקפה". בנו רווח סמך לאחוז "המכורים לקפה" $\alpha = 0.1$
3. חוקר בנה רווח סמך לאחוז האנשים שהתקררו לפחות פעם אחת בשנה. רווח הסמך שהתקבל הוא $81 < p < 91$ רווח הסמך הני"ל התבסס על מדגם של 500 איש.
- א. כמה אנשים במדגם טענו שכלל לא התקררו השנה?
- ב. באיזו רמת סמך נבנה רווח הסמך?
- ג. בנו רווח סמך לאחוז האנשים שהתקררו לפחות פעם אחת השנה ברמת סמך של 95% על סמך תוצאות המדגם

4. ציוני IQ בארה"ב מתפלגים נורמאלית עם תוחלת 100. במדגם של 20 ישראלים שנבחנו במבחן

$$\sum_{i=1}^{20} x_i = 2040$$

ה-IQ התקבלו התוצאות הבאות:

$$\sum_{i=1}^{20} x_i^2 = 210740$$

א. אמדו ברמת ביטחון של 90% את ממוצע ציוני בחינת ה-IQ בישראל – מהי ההנחה

הדרושה לפתרון?

ב. על סמך רווח הסמך של סעיף א האם תקבלו את הטענה שבישראל ממוצע הציונים שונה

מארה"ב?

ג. מה היה קורה לרווח הסמך אם הינו מגדילים את רמת הסמך שלו?

5. להלן תוצאות מדגם שבדק עבור כל משפחה האם יש לה בבית מכשיר טאבלט.

אזור מגורים	גוש דן	שאר הארץ
גודל המדגם	200	240
מספר משפחות בעלי טאבלט	160	168

א. בנו רווח סמך להבדל בין אחוז המשפחות עם טאבלט בגוש דן ואחוז המשפחות בעלי

טאבלט בשאר חלקי הארץ. ברמת סמך של 98%.

ב. בנו רווח סמך לפרופורצית משפחות בעלות טאבלט בכלל הארץ ברמת סמך של 95%.

6. הגובה של מתגייסים לצה"ל מתפלג נורמלית במדגם של 25 מתגייסים התקבלו התוצאות

הבאות:

$$\bar{x} = 176.2cm$$

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

א. אמדו את הגובה הממוצע של המתגייסים ברמת סמך של 98%.

ב. אמדו ברמת סמך של 90% את סטיית התקן של הגובה של מתגייסים של צה"ל.

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

8. להלן מדגם של שכר הדירה בשייח של 5 דירות שלושה חדרים בשכונת בבלי בתל אביב :

שנת 2012	8000	7500	7000	6500	7500
שנת 2013	8000	8200	7800	6800	7700

- בנו רווח סמך ברמת סמך של 95% לתוחלת עליית שכר הדירה משנת 2012 לשנת 2013 בשכונת בבלי. ניתן להניח ששכר הדירה בשכונה מתפלג נורמלית.

פתרונות:

שאלה 1

א. $80.65 \leq \mu \leq 93.35$

ב. $13.5 \leq \sigma \leq 22.9$

ג. $0.225 \leq p \leq 0.575$

שאלה 2

א. $1.21 \leq \mu \leq 1.65$

ב. $10.85\% \leq p \leq 19.15\%$

שאלה 3

א. 70

ב. 99.88%

ג. $83\% \leq p \leq 89\%$

שאלה 4

א. $97.4 \leq \mu \leq 106.6$

ב. לא

ג. יגדל

שאלה 5

א. $0.5\% < p_1 - p_2 < 19.5\%$

ב. $0.704 \leq p \leq 0.768$

שאלה 6

א. $170.8 \leq \mu \leq 181.6$

ב. $8.8 \leq \sigma \leq 14.3$

שאלה 7

א. $-372 \leq \mu_A - \mu_B \leq 772$

ב. $6467 \leq \mu \leq 7133$

שאלה 8

$-21 \leq \mu_{2013} - \mu_{2012} \leq 821$

פרק 27 - 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 הדחייה מוכתב על ידי סיכון שלוקח החוקר מראש שניקרא 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:

$$n = 20$$

$$\bar{X} = 3120$$

$$S = 280$$

- 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) 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 :

מציאות	הכרעה		
		H0	H1
	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?

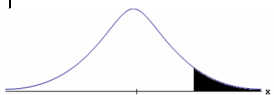
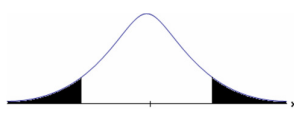
ה. מה מסקנת המחקר?

ו. מהי סוג הטעות האפשרית במחקר?

פרק 28 - בדיקת השערות על תוחלת (ממוצע)

כאשר שונות האוכלוסיה ידועה

רקע:

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

סטטיסטי המבחן :

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

חלופה אחרת לכלל הכרעה :

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

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 :

שאלה 1:

נקבל H_0

שאלה 2:

נדחה H_0

שאלה 3:

נדחה H_0

שאלה 4:

נדחה H_0

שאלה 5:

נקבל H_0

שאלה 6:

ב

שאלה 7:

ג

שאלה 8:

א. אותה מסקנה

ב. לא ניתן לדעת.

סיכוי לטעויות ועוצמה כאשר שונות האוכלוסייה ידועה

רקע:

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

נגדיר את ההסתברויות הבאות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות)

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

הסיכוי לבצע טעות מסוג 2:

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

רמת בטחון:

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

עוצמה:

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

התהליך לחישוב סיכוי לטעות מסוג שני:

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

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

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

דוגמה : (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.

השאלות שלהלן הן שאלות רב בררתיות. בחר בכל שאלה את התשובה הנכונה ביותר:

6. אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו אזי:

א. הסיכוי לטעות מסוג ראשון גדל.

ב. העוצמה של המבחן גדלה.

ג. הסיכוי לטעות מסוג שני גדל.

ד. תשובות א ו-ב נכונות.

7. חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן:

א. השערת האפס נכונה.

ב. השערת האפס נדחתה.

ג. השערת האפס לא נדחתה.

ד. אף אחת מהתשובות לא נכונה בהכרח.

8. מה המצב הרצוי לחוקר המבצע בדיקת השערה :

$1 - \beta$	α
גדולה	א. גדולה
קטנה	ב. גדולה
גדולה	ג. קטנה
קטנה	ד. קטנה

9. נערך שינוי בכלל ההחלטה של בדיקת השערה מסוימת ובעקבותיו אזור דחיית H_0 קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה מכך :

- א. הן α , והן $(1 - \beta)$, יקטנו.
- ב. α יישאר ללא שינוי ואילו $(1 - \beta)$ יגדל.
- ג. α יגדל ואילו $(1 - \beta)$ יקטן.
- ד. הן α והן $(1 - \beta)$ יגדלו.

10. ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניח שלחץ הדם בקרב עיתונאים גבוה יותר מהממוצע באוכלוסייה. הוא לקח מדגם של 60 עיתונאים וקיבל ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק שלחץ הדם בקרב העיתונאים אינו גבוה יותר. מה הטעות האפשרית שהרופא עושה ?

- 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

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 :

$$\text{אם } p_v \leq \alpha \text{ דוחים את } H_0$$

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

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

אם ההשערה היא דו צדדית :

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

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

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

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

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

דוגמה: (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 נורמלית עם תוחלת אליה כוונה המכונה וסטיית תקן 2 ס"מ. ביום מסוים כוונה המכונה לחתוך מוטות באורך 80 ס"מ. אחראי האיכות מעוניין לבדוק האם המכונה מכוילת. לצורך כך נדגמו מקו הייצור 16 מוטות שנחתכו אורכן הממוצע היה 81.7 ס"מ.

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

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

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

5. אם מקבלים בחישובים אלפא מינימלית (P value) קטנה 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 $p\text{ value}=0.0254$ 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:

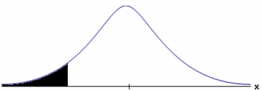
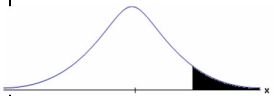
תשובה : א

שאלה 8:

תשובה : ג

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

רקע:

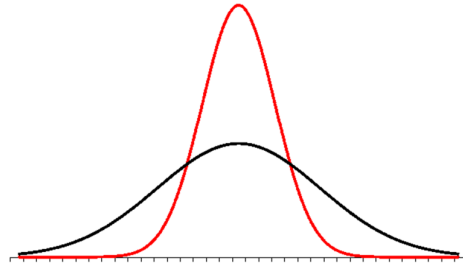
השערת האפס : השערה אלטרנטיבה :	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$
תנאים :	7. σ אינה ידועה 8. $X \sim N$ או מדגם מספיק גדול		
כלל ההכרעה : אזור הדחייה של H_0 :	$t_{\bar{x}} < -t_{1-\frac{\alpha}{2}}^{(n-1)}$ או $t_{\bar{x}}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

סטטיסטי המבחן :

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

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

התפלגות T:



הינה התפלגות סימטרית פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים. התפלגות T תלויה במושג שנקרא דרגות חופש. דרגות החופש הן $df=n-1$. ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה. כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה: (פתרון בהקלטה)

- מפעל קיבל הזמנה לייצור משטחים בעובי של 0.1 ס"מ.
- כדי לבדוק האם המפעל עומד בדרישה נדגמו 10 משטחים ונמצא שהעובי הממוצע הוא 0.104 עם אומדן לסטיית תקן 0.002 ס"מ.
- א. מהן השערות המחקר?
- ב. מה ההנחה הדרושה לצורך פתרון?
- ג. בדוק ברמת מובהקות של 5%.

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.

פתרונות:

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 $p_v \leq \alpha$ 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.

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

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

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

$$d.f = n - 1$$

דוגמה : (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% ?

תרגילים :

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 גבוהה מ-174 ס"מ באופן מובהק. מהי בקרוב מובהקות התוצאה ועל פיה מה תהיה המסקנה ברמת מובהקות של 6%?

פתרונות :

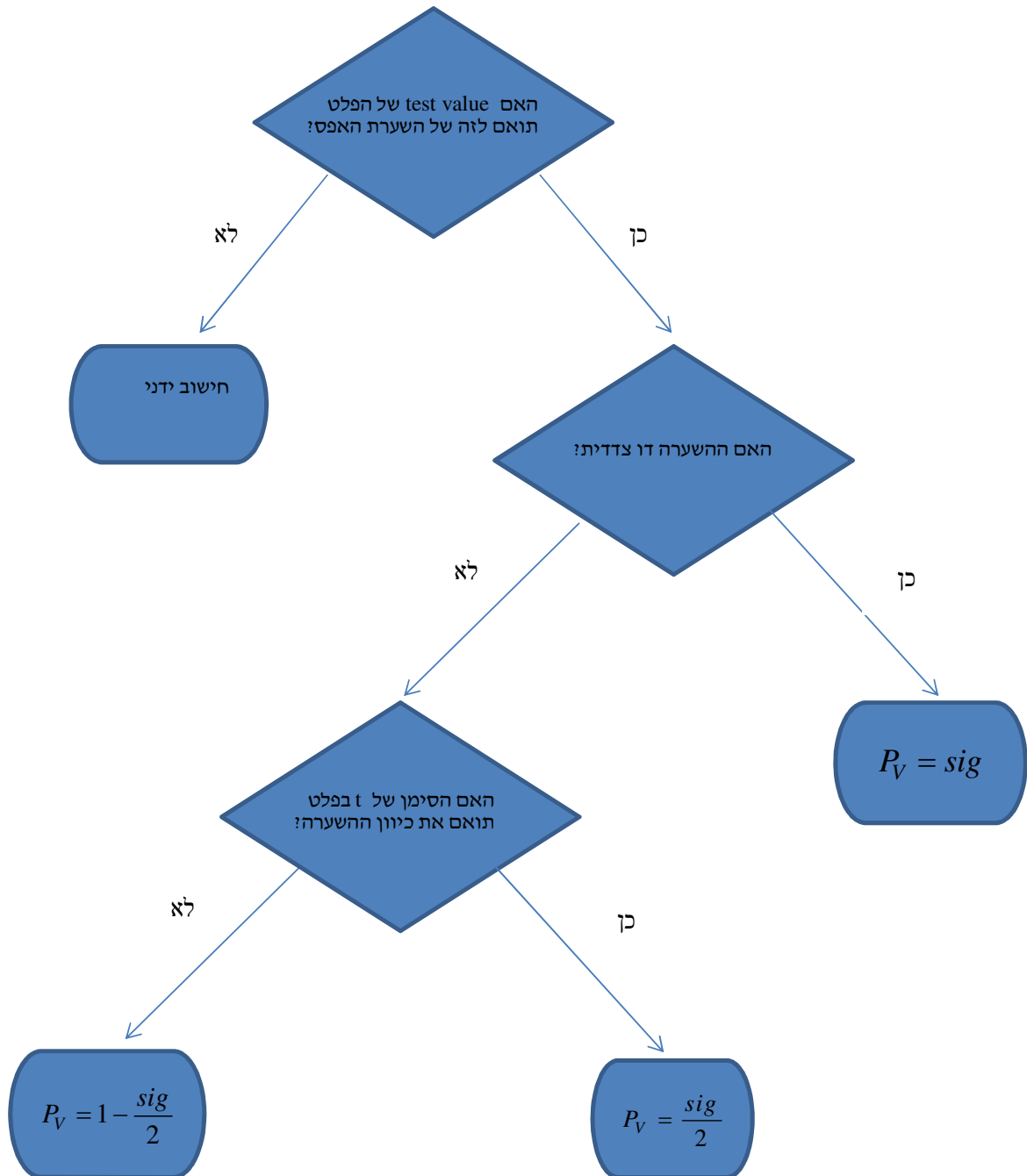
שאלה 3:

נקבל H_0

ניתוח פלטי SPSS

רקע:

חישוב מובהקות התוצאה באמצעות פלט תוכנת SPSS



דוגמה: (פתרון בהקלטה)

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
×	25	87.6400	64.90434	12.98087

One-Sample Test

	Test Value = 60					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
×	2.129	24	.044	27.64000	.8488	54.4312

ממוצע הציונים במבחן המיצב בחשבון הוא 60. הוחלט לדגום כיתה אקראית של 25 תלמידים וללמד אותם בשיטת לימוד חדשה.

א. מהו רווח הסמך לממוצע הציונים בחשבון אם יוחלט ליישם את שיטת הלימוד החדשה?

ב. מהו P_V לבדיקת יעילותה של שיטת הלימוד החדשה?

ג. מה יוכרע ברמת מובהקות של 5% לגבי יעילותה של שיטת הלימוד החדשה?

תרגילים :

1. באוניברסיטה גדולה גיל הסטודנטים לתואר ראשון מתפלג נורמאלי. בעבר פורסם שהגיל הממוצע של הסטודנטים הינו 23. להלן פלט תוכנת SPSS על מדגם של 16 סטודנטים אקראיים מתואר ראשון:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
age	16	23.4375	2.50250	.62562

One-Sample Test

	Test Value = 23					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
age	.699	15	.495	.43750	-.8960	1.7710

- א. מהו המבחן הסטטיסטי שנעשה כאן?
- ב. מה ערכו של הפרמטר לפי השערת האפס?
- ג. רשום רווח סמך ברמת סמך של 95% לתוחלת גיל הסטודנטים באוניברסיטה לתואר ראשון.
- ד. בדוק ברמת מובהקות של 5% האם הגיל הממוצע כיום שונה מבעבר?

2. קבוצת ילדים בגיל 6 קיבלה משימה לביצוע. עבור כל ילד בדקו כמה זמן לקח לו לסיים את המשימה בדקות. להלן תוצאות הניתוח הסטטיסטי:

One-Sample Test

	Test Value = 4.5					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
time	-1.853	24	.076	-.09200	-.1944	.0104

- א. כמה ילדים השתתפו במחקר?
- ב. מצא רווח סמך ברמת סמך של 95% לתוחלת זמן ביצוע המשימה עבור ילדים בני 6.
- ג. מה יש להניח כדי שרווח הסמך מסעיף א' יהיה תקף?
- ד. בדוק ברמת מובהקות של 5% שזמן ביצוע המשימה הממוצע נמוך מ-4.5 דקות.

3. לפניך פלט מחשב עבור ניתוח סטטיסטי שנעשה בתוכנת SPSS. הניתוח הוא עבור מדגם אקראי של קבוצת נבחנים בבגרות באנגלית.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
grade	???	???	19.62787	2.95901

One-Sample Test

	Test Value = 75					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
grade	???	43	.017	-7.34091	-13.3083	???

- השלם את הגדלים החסרים המסומנים בסמני שאלה בפלט.
- מהי מובהקות התוצאה לבדיקת ההשערה שהתוחלת של הציונים שונה מ-75?
- מהי מובהקות התוצאה לבדיקת ההשערה שהתוחלת של הציונים קטנה מ-75?
- מהי מובהקות התוצאה לבדיקת ההשערה שהתוחלת של הציונים גדולה מ-75?

4. יצרן סיגריות מפרסם כי תוחלת הניקוטין בסיגריות שהוא מיצר קטנה מ- 27 מ"ג. בבדיקה מקרית של 5 סיגריות מתוצרתו נמצאו כמויות הניקוטין הבאות:
21, 20, 24, 22 מ"ג. הנח כי כמות הניקוטין בסיגריות מפולג נורמאלי.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
nicotine	5	21.6000	1.51658	.67823

One-Sample Test

	Test Value = 27					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
nicotine	-7.962	4	.001	-5.40000	-7.2831	-3.5169

- האם ברמת מובהקות של 5% ניתן להסיק שיש אמת בפרסום?
- אם הינו מוסיפים עוד תצפית שערכה 20. כיצד הדבר היה משפיע על הערך Sig ועל המסקנה?
- בדוק האם ניתן להגיד שתוחלת רמת הניקוטין שונה מ – 26 ברמת מובהקות של 5%.

פתרונות :

שאלה 1:

ב. 23

ג. (22.104, 24.771)

ד. נקבל H_0

שאלה 2:

א. 25

ב. (4.3056, 4.5104)

ד. נדחה H_0

שאלה 3

א. $n = 44$

$$\bar{X} = 67.66$$

$$t = -2.48$$

$$upper = -1.3735$$

ב. 0.017

ג. 0.0085

ד. 0.9915

שאלה 4:

א. נכריע שיש אמת בפרסום.

ב. מסקנה לא תשתנה.

ג. נכריע שהתוחלת שונה מ-26.

הקשר בין רווח סמך לבדיקת השערות על תוחלת

רקע:

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 $\leftarrow$ נקבל את H_0

אם μ_0 לא נופל ברווח $\leftarrow$ נדחה את H_0

דוגמה: (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 10.4 מ"ג. נלקח מדגם של 60 נבדקים שניזונו מדיאטה זו. נמצא כי ממוצע מספר המיליגרם סוכר היה 115.5 מ"ג לסמ"ק.

א. בנה רווח סמך ברמת סמך 95% לתוחלת רמת הסוכר בדם אצל הניזונים מדיאטה זו.
ב. ידוע שתוחלת רמת הסוכר בדם באוכלוסיה היא 90 מ"ג לסמ"ק. האם לדעתך ניתן להסיק על סמך תוצאת סעיף א שהדיאטה משפיעה על רמת הסוכר בדם? 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 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. $112.87 \leq \mu \leq 118.13$

ב. נכריע שהדיאטה משפיעה על תוחלת רמת הסוכר בדם.

שאלה 3:

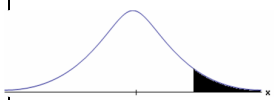
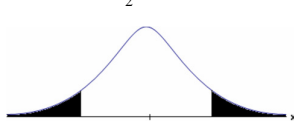
א. $191.8 \leq \mu \leq 200.2$

ב. נכריע שיש אמת בפרסום.

פרק 29 - בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

כשהשונות של האוכלוסייה ידועות

רקע:

$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 > c$	$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 < c$	$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 \neq c$	השערת האפס: השערה אלטרנטיבית:
תנאים: 1. מדגמים בלתי תלויים 2. σ_1, σ_2 ידועות 3. $X_1, X_2 \sim N$ או מדגמים מספיק גדולים			
$Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\alpha}$  $Z_{1-\alpha}$ - דוחים את H_0	$Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ - דוחים את H_0	$Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\frac{\alpha}{2}}$ או $Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}} \quad Z_{1-\frac{\alpha}{2}}$ - דוחים את H_0	כלל ההכרעה: אזור הדחייה של H_0 :

סטטיסטי המבחן:

$$Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

חלופה אחרת לכלל הכרעה:

$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ או $\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	נדחה H_0 אם מתקיים:
---	---	--	---

התפלגות הפרש הממוצעים:

$$\bar{x}_1 - \bar{x}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

התקנון:

$$Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

דוגמה : (פתרון בהקלטה)

בשנת 2004 הפער בין השכר הממוצע של הגברים לנשים היה 3000 ₪ לטובת הגברים.
מעוניינים לבדוק האם כיום הצטמצם הפער בין הגברים לנשים מבחינת השכר הממוצע.
נדגמו 100 עובדים גברים. שכרם הממוצע היה 9,072 ₪. נדגמו 80 עובדות, שכרן הממוצע היה
7809 ₪. לצורך פתרון נניח שסטיות התקן של השכר ידועות ושוות ל-2000 ₪ באוכלוסיית הנשים
ו-3000 ₪ באוכלוסיית הגברים. מה המסקנה ברמת מבוהקות של 5%?

תרגילים :

1. מחקר טוען שאנשים החיים במרכז הארץ צופים בממוצע בטלוויזיה יותר מאנשים שלא חיים במרכז. נדגמו 100 אנשים מהמרכז ו-107 אנשים לא מהמרכז. אנשים אלו נשאלו כמה שעות ביום הם נוהגים לצפות בטלוויזיה.
במדגם של מרכז הארץ התקבל ממוצע 2.7 שעות.
במדגם של מחוץ למרכז הארץ התקבל ממוצע 1.8 שעות.
לצורך פתרון הניחו שבכל אזור, סטיית התקן היא שעה 1 ביום. בדקו את טענת המחקר ברמת מובהקות של 1%.
2. ציוני פסיכומטרי מתפלגים נורמלית עם סטיית תקן 100. מכון ללימוד פסיכומטרי טוען שהוא יכול לשפר את ממוצע הציונים ביותר מ-30 נקודות. במדגם של 20 נבחנים שניגשו למבחן ללא הכנה במכון התקבל ממוצע 508. במדגם של 25 נבחנים שעברו הכנה במכון התקבל ממוצע ציונים 561. מה מסקנתכם ברמת מובהקות של 5%.
3. במדגם אקראי של 20 ימים נבדקה התפוקה של מפעל ביום. התפוקה הממוצעת הייתה של 340 מוצרים ליום. במדגם אקראי של 20 ימים אחרים נבדקה התפוקה של המפעל בלילה והתפוקה הממוצעת הייתה 295. לצורך פתרון נניח שסטיית התקן של התפוקה ביום היא 40 מוצרים ובלילה 30 מוצרים.
א. מהי מובהקות התוצאה לבדיקה האם התפוקה הממוצעת היומית גבוהה מהתפוקה הממוצעת הלילית.
ב. מה תהיה המסקנה ברמת מובהקות של 8%?
4. במחקר מקיף שנעשה באירופה נקבע שגברים גבוהים מנשים ב-8 ס"מ בממוצע.
מחקר ישראלי מתעניין לבדוק האם בישראל הפער גדול יותר. לצורך המחקר נדגמו 40 גברים ו 40 נשים באקראי. כמו כן, נניח שסטיות התקן של הגברים והנשים ידועות ושוות ל-6 ס"מ אצל הנשים. ו-12 ס"מ אצל הגברים.
א. מהן השערות המחקר ומהו כלל ההכרעה ברמת מובהקות של 10%?
ב. אם בישראל הפער בין גברים לנשים מבחינת הגובה הממוצע הוא 11 ס"מ, מה ההסתברות שהמחקר לא יגלה זאת? איך קוראים להסתברות הזאת?

פתרונות:

שאלה 1:

נדחה H_0

שאלה 2:

לא נדחה את H_0

שאלה 3:

א. 0

ב. נדחה H_0

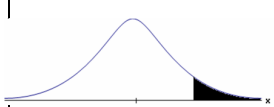
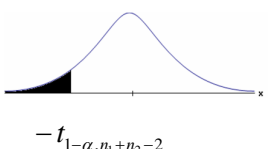
שאלה 4:

א. נדחה H_0 אם במדגם הגברים יהיו גבוהים בממוצע מהנשים ביותר מ-10.72 ס"מ.

ב. 0.6331

כששונויות האוכלוסיה לא ידועות ומניחים שהן שוות

רקע:

השערת האפס : השערה אלטרנטיבית:			תנאים:
$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 > c$	$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 < c$	$H_0 \quad \mu_1 - \mu_2 = c$ $H_1 \quad \mu_1 - \mu_2 \neq c$	4. מדגמים בלתי תלויים 5. σ_1, σ_2 לא ידועות אך שוות 6. המשתנים בכל אוכלוסייה מתפלגים נורמלית
$t_{\bar{x}_1 - \bar{x}_2} > t_{1-\alpha}^{(n_1+n_2-2)}$  $t_{1-\alpha, n_1+n_2-2}$ דוחים את H_0	$t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\alpha}^{(n_1+n_2-2)}$  $-t_{1-\alpha, n_1+n_2-2}$ דוחים את H_0	או $t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$ או $t_{\bar{x}_1 - \bar{x}_2} > t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$  $-t_{1-\frac{\alpha}{2}, n_1+n_2-2}$ $t_{1-\frac{\alpha}{2}, n_1+n_2-2}$ דוחים את H_0	אזור הדחייה של H_0 :

סטטיסטי המבחן :

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \quad \text{השונוות המשוקללת} \quad t_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

חלופה אחרת לכלל הכרעה:

$\bar{x}_1 - \bar{x}_2 > c + t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 < c - t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$	או $\bar{x}_1 - \bar{x}_2 > c + t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$ או $\bar{x}_1 - \bar{x}_2 < c - t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$	נדחה H_0 אם מתקיים:
---	---	--	-----------------------

דוגמה : (פתרון בהקלטה)

חברה המייצרת מוצרי בנייה טוענת שפיתחה סגסוגת (תערובת מתכות) שטמפרטורת ההתכה שלה גבוהה משמעותית מטמפרטורת ההתכה של הסגסוגת לבנייה שמשמשים בה כיום לבניית בניינים.

לצורך בדיקת טענת המחקר נדגמו 10 יחידות של מתכות מהסוג הישן ו-12 יחידות של מתכות מהסוג החדש.

להלן תוצאות המדגם:

טמפרטורת ההתכה הממוצעת במתכת הישנה 1170 מעלות עם אומד חסר הטיה לשונות

$$S^2 = 200$$

טמפרטורת ההתכה הממוצעת במתכת החדשה 1317 מעלות עם אומד חסר הטיה לשונות

$$S^2 = 260$$

נניח לצורך פתרון שטמפרטורת ההתכה מתפלגת נורמאלית עם אותה שונות במתכות השונות. בדקו ברמת מובהקות של 5%.

תרגילים:

1. להלן נתונים של שטחי דירות מתוך דירות שנבנו בשנת 2012 ובשנת 2013 (מטרים רבועים):

120	94	90	130	95	112	120	2012
	69	74	105	91	82	100	2013

בדקו שבשנת 2013 הייתה ירידה משמעותית בשטחי הדירות לעומת שנת 2012 עבור רמת מובהקות של 5%. הניחו ששטחי הדירות בכל שנה מתפלגים נורמלית עם אותה שונות.

2. נדגמו 15 ישראלים ו-15 אמריקאים. כל
הנדגמים נגשו למבחן IQ. להלן תוצאות המדגם:

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

בדקו ברמת מובהקות של 5% האם קיים הבדל של נקודה בין ישראלים לאמריקאים מבחינת ממוצע הציונים במבחן ה-IQ לטובת ישראל. רשמו את כל ההנחות הדרושות לצורך פתרון התרגיל.

3. להלן תוצאות מדגם הבדק אורך חיים של נורות מסוג W60 ומסוג W100.
אורך החיים נמדד בשעות.

הקבוצה	2-60W	1-100W
$\bar{x}$	1007	956
S	80	72
n	13	15

- א. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקות בממוצע יותר מאשר נורות מסוג W100. רשמו את כל ההנחות הדרושות לפתרון.
- ב. עבור איזו רמת מובהקות ניתן לקבוע שנורות מסוג W60 דולקות בממוצע יותר מאשר נורות מסוג W100?
- ג. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקות יותר מ-1000 שעות. רשמו את כל ההנחות הדרושות.

פתרונות :

שאלה 1:

לא נדחה H_0

שאלה 2:

שאלה 3:

א. נדחה H_0

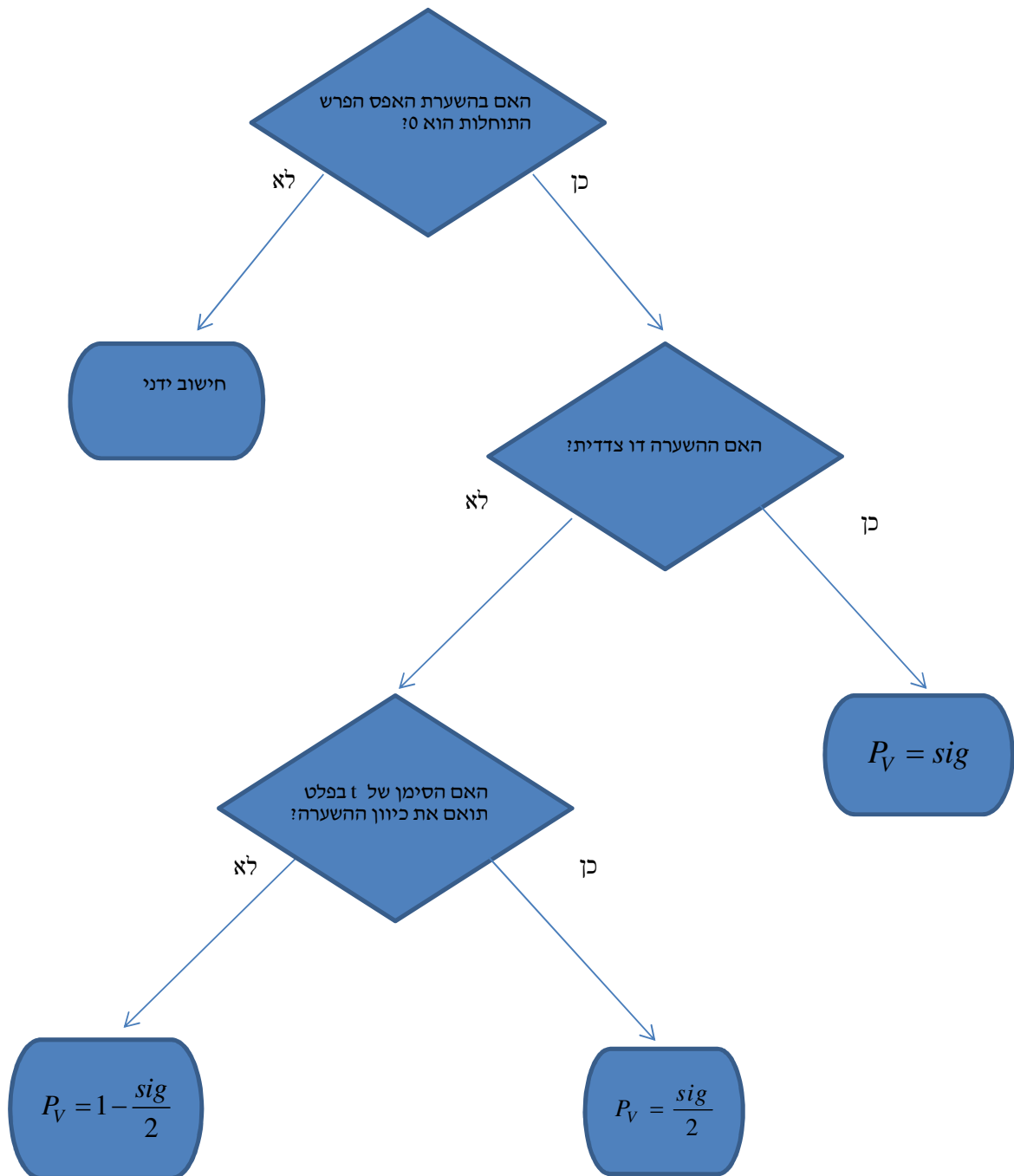
ב. רמות מובהקות של לפחות 5%

ג. לא נדחה H_0

ניתוח פלטים

רקע :

מובהקות התוצאה על סמך הפלט:



דוגמה: (פתרון בהקלטה)

בסקר שנערך בארה"ב בשנת 1993 נשאלו נסקרים משני אזורים שונים במדינה על מס' האחים והאחיות שלהם. להלן הפלט שהתקבל:

Group Statistics

	Region of the United States	N	Mean	Std. Deviation	Std. Error Mean
Number of Brothers and Sisters	North East	676	3.76	2.939	.113
	South East	410	4.05	2.993	.148

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Number of Brothers and Sisters	Equal variances assumed	.173	.677	-1.583	1084	.114	-.293	.185	-.657	.070
	Equal variances not assumed			-1.576	850.945	.115	-.293	.186	-.658	.072

א. מהו המבחן הסטטיסטי שנעשה כאן?

ב. בדוק ברמת מובהקות של 5% האם קיים שוויון שונויות בין שני האזורים?

ג. בדוק האם קיים הבדל בין "South East" ל-"North East" ברמת מובהקות של 5% מבחינת מספר האחים והאחיות הממוצע.

ד. מהי מובהקות התוצאה לבדיקת הטענה שההפרש הממוצע בין-"South East"

לבין-"North East" חיובי ?

תרגילים :

1. להלן פלט מתוכנת SPSS מתוך מחקר שבחן את רמת האופטימיות של גברים ונשים.
רמת האופטימיות נמדדה בסקאלת ציונים של 1 עד 5.

Group Statistics

	GENDER	N	Mean	Std. Deviation	Std. Error Mean
optimizm	MALE	633	2.6053	.49781	.01979
	FEMALE	568	2.5503	.48483	.02034

Independent Samples Test

		Levene's Test for quality of Variance:		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
optimizm	Equal varianc assumed	.383	.536	1.935	1199	.053	.05500	.02842	-.00076	?
	Equal varianc not assumed			1.938	190.977	.053	.05500	.02838	-.00068	.11067

- א. האם ניתן להניח ששונות האופטימיות של נשים וגברים שווה ברמת מובהקות של 5%?
- ב. ברמת מובהקות של 5% האם קיים הבדל בין הנשים לגברים ברמת האופטימיות הממוצעת שלהם?
- ג. מצא את הגבול העליון של רווח הסמך המסומן בסימן שאלה בפלט. דייק עד 5 ספרות אחרי הנקודה.
- ד. בנה רווח סמך לתוחלת רמת האופטימיות של הגברים ברמת סמך של 95%.

2. פסיכולוגים טוענים שאנשים שניגשים למבחן אינטליגנציה יותר מפעם אחת נוטים לקבל ציונים גבוהים יותר.

להלן הפלט שהתקבל:

T-Test

Group Statistic:

		N	Mean	Std. Deviation	Std. Error Mean
grade	A	9	96.8889	9.40006	3.13335
	B	11	108.4545	11.46616	3.45718

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
grade	Equal variances assumed	.206	.656	-2.428	18	.026	-11.56566	4.76333	-21.57304	-1.55828
	Equal variances not assumed			-2.479	17.997	.023	-11.56566	4.66583	-21.36832	-1.76299

מקרא:

A = נגשו פעם אחת.

B = נגשו יותר מפעם אחת.

- רשמו את השערות המחקר והסבירו מהו המבחן המתאים כאן.
 - כיצד הייתה משתנה התשובה לסעיף הקודם אם היה מדובר על אותם אנשים שציונם נבדק פעם אחרי המבחן הראשון שעשו ופעם אחרי המבחן השני?
 - האם ניתן לומר כי מידת הפיזור של ציוני אנשים הנבחנים בפעם הראשונה שונה ממידת הפיזור של ציוני האנשים אשר נבחנים בפעם השנייה. בדוק ברמת מובהקות של $\alpha=0.05$.
 - האם נכונה טענת הפסיכולוגים ברמת מובהקות של $\alpha=0.01$.
3. כחלק ממחקר בנושא הנישואין בישראל, אחד החוקרים העלה השערה שיש הבדל במוצא גיל הנישואין (הראשונים), בין נשים הגרות בערים מרכזיות לבין נשים הגרות בערים מרוחקות מהמרכז.
- לשם כך נדגמו 50 כלות מכל אחת משתי ערים עיר א'-מרכזית ועיר ב'-מרוחקת ונרשם גילן. תוצאות עיבוד הנתונים מופיעות בטבלאות שלהלן:

T-Test

Group Statistics

מקום המגורים		N	Mean	Std. Deviation	Std. Error Mean
גיל הנישואין	עיר א	50	24.8072	1.38978	.19654
	עיר ב	50	23.0131	1.62070	.22920

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
גיל הנישואין	Equal variances assumed	.330	.567	5.942	98	.000	1.79415	.30193	1.19497	2.39332
	Equal variances not assumed			5.942	95.772	.000	1.79415	.30193	1.19480	2.39350

א. מהו המבחן הסטטיסטי שנעשה כאן?

ב. מצא רווח סמך ברמת סמך של 95% להפרש בין עיר א לעיר ב מבחינת גיל הנשים הממוצע בנישואין הראשונים.

ג. האם ניתן לומר ברמת מובהקות של 1% שנשים בערים מרכזיות מתחתנות בגיל מאוחר יותר מאשר נשים הגרות בערים מרוחקות?

4. להלן פלט של תוכנת SPSS:

T-Test

	N	Mean	Std. Deviation	Std. Error Mean
X	26	36.3077	13.23259	2.59513
Y	24	46.4583	20.96369	4.27920

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2- ailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	4.446	.040	-2.164	???	.044	-10.15064	???	-20.03781	-.26347
Equal variances not assumed			-2.038	38.267	.048	???	5.00462	-20.27964	-.02164

א. השלם את סימני השאלה בטבלה.

ב. מהי מובהקות התוצאה לבדיקת הטענה שקיים הבדל בין השונות של X לזה של Y?

ג. מהי מובהקות התוצאה לבדיקת הטענה שהתוחלת של X גדולה מהתוחלת של Y?

ד. מהי מובהקות התוצאה לבדיקת הטענה שהתוחלת של X קטנה מהתוחלת של Y?

פתרונות:

שאלה 1:

- א. נכריע שוויון שוניות.
ב. נקבע שלא קיים הבדל בין נשים לגברים מבחינת האופטימיות הממוצעת.
ג. 0.11076
ד. $2.5665 \leq \mu \leq 2.6441$

שאלה 2:

- א. מבחן t למדגמים בלתי תלויים.
ב. מבחן t למדגמים מזווגים.
ג. נקבע שקיים שוויון שוניות.
ד. לא נקבל את טענת הפסיכולוגים.

שאלה 3:

- א. מבחן t להשוואת תוחלת במדגמים בלתי תלויים.
ב. $1.19497 \leq \mu_1 - \mu_2 \leq 2.39332$
ג. כן.

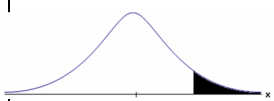
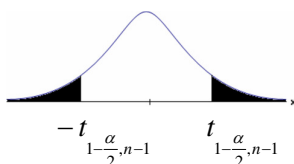
שאלה 4:

- א. -10.15, 4.69, 48
ב. 0.04
ג. 0.978
ד. 0.022

**פרק 30 - בדיקת השערות על תוחלת ההפרשים במדגמים מזווגים
(תלויים)**

בדיקת השערות למדגמים מזווגים

רקע:

$H_0: \mu_D = C$ $H_1: \mu_D > C$	$H_0: \mu_D = C$ $H_1: \mu_D < C$	$H_0: \mu_D = C$ $H_1: \mu_D \neq C$	השערת האפס : השערה אלטרנטיבית:
7. σ_D אינה ידועה 8. $D \sim N$ או מדגם מספיק גדול			תנאים:
$t_{\bar{D}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha, n-1}$ דוחים את H_0	$t_{\bar{D}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha, n-1}$ דוחים את H_0	$t_{\bar{D}} < -t_{1-\frac{\alpha}{2}}^{(n-1)}$ או $t_{\bar{D}} > t_{1-\frac{\alpha}{2}}^{(n-1)}$  $-t_{1-\frac{\alpha}{2}, n-1}$ $t_{1-\frac{\alpha}{2}, n-1}$ דוחים את H_0	כלל ההכרעה: אזור הדחייה של H_0:
$\bar{D} > C + t_{1-\alpha}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} < C - t_{1-\alpha}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} > C + t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$ או $\bar{D} < C - t_{1-\frac{\alpha}{2}}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	חלופה לכלל הכרעה : נדחה H_0 אם מתקיים:

סטטיסטי המבחן :

$$t_{\bar{D}} = \frac{\bar{D} - \mu_D}{S_D / \sqrt{n}}$$

$$S_D^2 = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n-1} = \frac{\sum_{i=1}^n D_i^2 - n\bar{D}^2}{n-1}$$

דוגמה : (פתרון בהקלטה)

חברה שיווקית מעוניינת לבדוק את טענת רשת השיווק "מגה בעיר" הטוענת שמחירה נמוכים מהמחירים מרשת השיווק "שופרסל".

לצורך בדיקה נבחרו באקראי 4 מוצרים שונים. המחירים נבדקו בשתי הרשתות. להלן המחירים :

המוצר	מגה בעיר	שופרסל
שמפו	17	18
גיל כביסה	48	57
עוגת גבינה	35	35
לחם	12	10
קפה נמס	49	47
בקבוק יין	113	142
גבינה בולגרית	20	26

בהנחה והמחירים מתפלגים נורמאלית בדקו ברמת מובהקות של 5% את טענת רשת "מגה בעיר".

תרגילים:

1. במטרה לבדוק האם קיים הבדל בין חברת X לחברת Y מבחינת המחירים לשיחות בינ"ל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה. להלן התוצאות:

המדינה	X	Y
ארה"ב	1.5	1.4
קנדה	2.1	2
הולנד	2.2	1.9
פולין	3	3.1
מצרים	3.5	3.2
סין	3.2	3.2
יפן	4.2	4.2

בהנחה והמחירים מתפלגים נורמלית בכל חברה, בדקו ברמת מובהקות של 5% האם קיים הבדל בין החברות מבחינת המחירים בממוצע?

2. מכון המכין לפסיכומטרי טוען שהוא מעלה את ממוצע הציונים ביותר מ-30 נקודות. 8 נבחנים נבדקו לפני ואחרי שהם למדו במכון. להלן התוצאות שהתקבלו:

לפני	506	470	420	640	670	390	500	590
אחרי	570	540	430	610	680	510	520	580

מה מסקנתכם ברמת מובהקות 5%? הניחו שציוני פסיכומטרי מתפלג נורמלית.

3. נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'. להלן הציונים שלהם בסמסטר א' ו- ב':

סטטיסטיקה א	סטטיסטיקה ב
74	80
68	84
90	87
75	76
82	100

- פורסם שתלמידים שמסיימים את סמסטר ב משפרים בממוצע את הציונים ב-5 נקודות לעומת סמסטר א' . הנח שהציונים מתפלגים נורמלית.
- א. מהי מובהקות התוצאה לבדיקת הטענה שהשיפור הוא יותר מ 5 נקודות?
- ב. על סמך הסעיף הקודם , מהי רמת המובהקות המינימלית להכרעה שהשיפור הוא יותר מ- 5 נקודות?
- ג. לאור זאת, מה המסקנה ברמת מובהקות של 10% ?

4. לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים.
- אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני ללא היפנוזה. לאחר מכן נערך לשניהם מבחן באנגלית. נניח שציוני המבחן מתפלגים נורמאלית ללא ידיעת השונות האמתית.
- המבחן שיש לבצע כאן הוא:
- א. מבחן Z למדגם יחיד.
- ב. מבחן T למדגם יחיד.
- ג. מבחן T למדגמים בלתי תלויים.
- ד. מבחן T למדגמים מזווגים.

5. בתחנת טיפת חלב מסוימת יש שני מכשירי שקילה. על מנת להשוות בין שני המשקלים נדגמו 4 תינוקות. כל תינוק בן חודשיים נשקל בכל אחד מהמשקלים. להלן תוצאות השקילה (בק"ג):

משקל במכשיר 1	5.4	6.9	7.0	5.2
משקל במכשיר 2	5.3	6.9	7.1	5.0

נניח שהמשקלים מתפלגים נורמלית.

המבחן שיש לבצע כאן הוא :

א. מבחן Z למדגם יחיד.

ב. מבחן T למדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

6. כדי להשוות בין שני אצים נדגמו 5 תוצאות מריצת 100 מטר של כל אצן. זמני הריצה נרשמו ויש להניח שמתפלגים נורמלית. המטרה להשוות בין האצנים.

המבחן שיש לבצע כאן הוא :

א. מבחן Z למדגם יחיד.

ב. מבחן T למדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

פתרונות:

שאלה 1:

לא נדחה H_0

שאלה 2:

לא נדחה H_0

שאלה 3:

א. $0.1 \leq p \leq 0.25$

ב. 0.25

ג. לא נדחה H_0

שאלה 4:

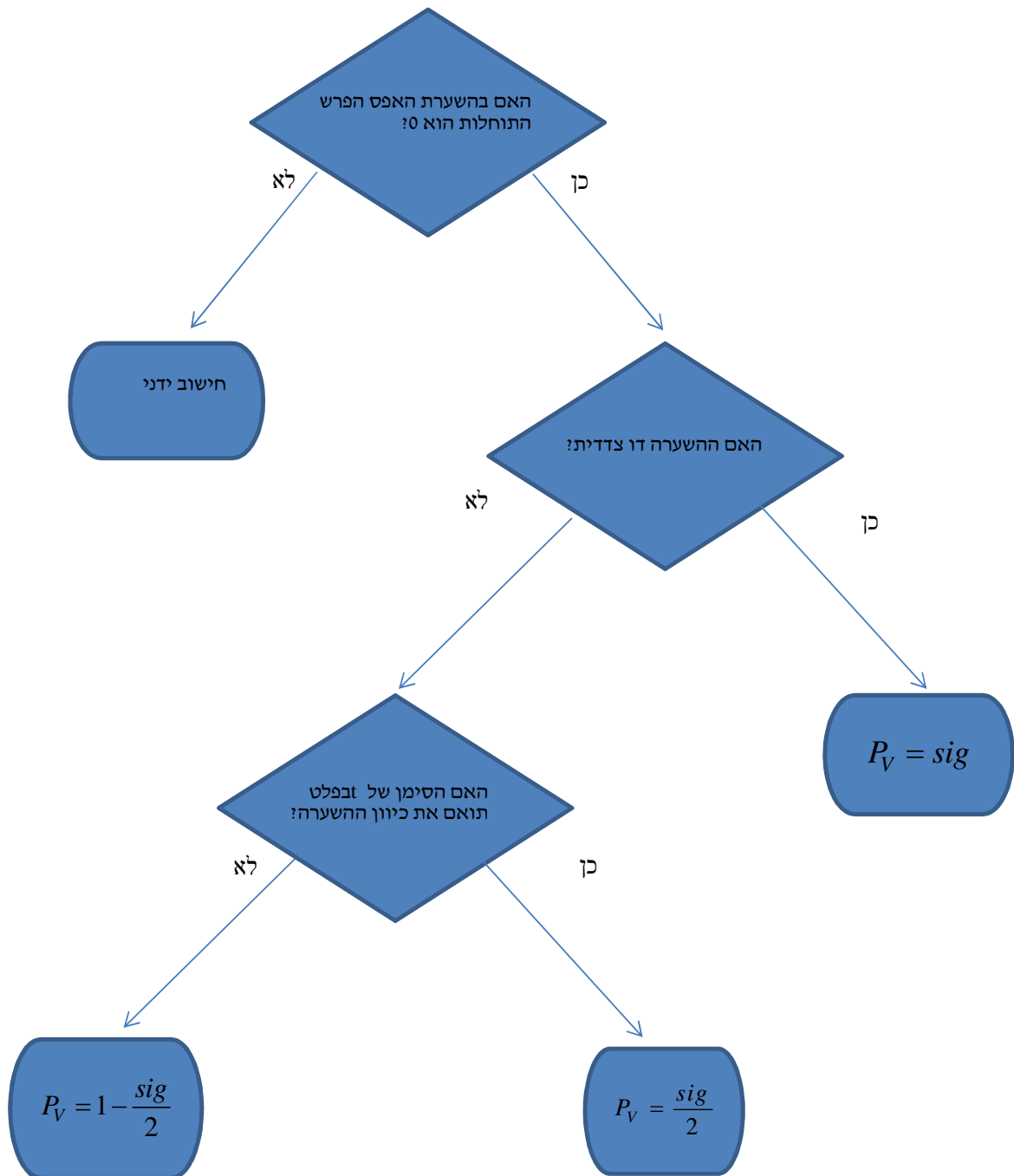
התשובה היא ד.

שאלה 5:

התשובה היא ד.

שאלה 6:

התשובה היא ג.



כדי לבדוק את ההשפעה של קורס לגמילה מעישון נלקח מדגם מקרי של 5 נבדקים. עבור כל אחד מהם נמדדה צריכת הסיגריות היומית לפני הקורס וחודשיים אחריו. הניחו שצריכת הסיגריות מתפלגת נורמלית. להלן התוצאות:

נבדק	1	2	3	4	5
לפני	40	22	25	28	30
אחרי	30	24	13	10	12

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 BEFORE	29.0000	5	6.85565	3.06894
AFTER	17.8000	5	8.72926	3.90364

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	90% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	BEFORE - AFTER	11.20000	8.19756	3.66606	3.38452	19.01548	3.055	4	.038

בדוק ברמת מובהקות של 5% האם הקורס יעיל.

תרגילים :

1. בסקר שנערך בארה"ב בשנת 1993 נשאלו נסקרים על השכלת הוריהם, להלן הפלט שהתקבל :

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Highest Year School Completed, Father - Highest Year School Completed, Mother	-.007	3.115	.100	-.203	.189	-.072	973	.943

- א. תן אומדן להפרש הממוצעים.
- ב. תן אומדן לטעות התקן של הפרש הממוצעים.
- ג. האם קיים הבדל מובהק בין השכלת האבות להשכלת האימהות ברמת מובהקות של 5%?

2. בתחרות קפיצה למים שופטים באופן קבוע שופט איטלקי ושופט דרום קוריאני. להלן פלט המנתח את הציונים ששופטים אלה נתנו בתחרויות השונות :

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Italy	???	300	.86742	.05008
	South Korea	8.9183	???	.81992	.04734

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
				Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Italy - South Korea	-.42233	.36153	.02087	-.46341	-.38126	-20.234	???	???

- א. השלימו את החלקים החסרים בפלט. (מסומנים בסימני שאלה).
- ב. בדוק את הטענה שהשופט הדרום קוריאני נותן בממוצע 0.2 נקודות יותר מאשר השופט האיטלקי ברמת מובהקות של 5%.
- ג. מהו רווח הסמך ברמת סמך של 95% לתוחלת פער הציונים בין השופטים.
- ד. בנה את הרווח כעת ברמת סמך של 98% לתוחלת פער בציונים בין השופטים.

3. בדקו את ציוניהם של 44 נבדקים אקראיים במבחן הפסיכומטרי. פעם אחת לפני הכנה (Before) ופעם אחת אחרי הכנה (After).

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Before - After	-7.45455	19.28303	2.90703	-13.31712	-1.59197	-2.564	43	.014

- א. רשום מהו המבחן הסטטיסטי ונסח את ההשערות אליהם מתייחס הפלט.
- ב. בדוק את ההשערה שממוצע ציונים משתפרים לאחר ההכנה ברמת מובהקות של 5%.
- ג. בדוק את ההשערה שממוצע ציונים משתפרים לאחר ההכנה ביותר מ-5 נקודות ברמת מובהקות של 5%.
- ד. מצא רווח סמך לתוחלת שיפור ממוצע הציונים לאחר ההכנה ברמת ביטחון של 95%.

4. להלן פלט של תכנת SPSS :

T-Test

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 x	54.0000	6	5.86515	2.39444
y	46.5000	6	10.72847	4.37988

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 $\bar{x} - \bar{y}$	7.50000	??	4.72405	-4.64356	19.64356	??	5	.173

א. מלא את החלקים החסרים בטבלה.

ב. מהי רמת המובהקות המינימלית לקבלת הטענה שיש הבדל בין X ל- Y בממוצע ?

ג. האם התשובה לסעיף הקודם הייתה משתנה , ואם כן גדלה או קטנה , אם הינו מוסיפים

עוד תצפית שההפרש בין X ל- Y הוא 0.

ד. מהי מובהקות התוצאה לבדיקת הטענה ש X גדול מ- Y בממוצע?

ה. מהי מובהקות התוצאה לבדיקת הטענה ש X קטן מ- Y בממוצע?

ו. בנו רווח סמך לתוחלת של X ברמת סמך של 90% .

פתרונות:

שאלה 1:

א. -0.007

ב. 0.1

ג. אין הבדל מובהק

שאלה 2:

א. $d.f = 299$

$n = 300$

$\bar{X} = 8.496$

Sig=0

ב. נדחה H_0

ג. (0.46341, 0.38126)

ד. (0.4708, 0.3738)

שאלה 3:

ב. נדחה H_0

ג. לא נדחה את H_0

ד. (1.592, 13.317)

שאלה 4:

א. 1.5876, 11.5715

ב. 0.173

ג. יגדל

ד. 0.0865

ה. 0.9135

ו. $49.18 < \mu < 58.82$

פרק 31 - הקשר בין רווח סמך לבדיקת השערות על הפרש תוחלות

רקע:

ניתן לבצע בדיקת השערות דו צדדית ברמת מובהקות α על $\mu_1 - \mu_2$:

$$H_0 : \mu_1 - \mu_2 = C$$

$$H_1 : \mu_1 - \mu_2 \neq C$$

על ידי בניית רווח סמך ברמת סמך של $1 - \alpha$ ל $\mu_1 - \mu_2$:

אם C נופל ברווח $\leftarrow$ נקבל את H_0

אם C לא נופל ברווח $\leftarrow$ נדחה את H_0

דוגמה: (פתרון בהקלטה)

חוקר ביצע בדיקת השערות לתוחלת ההפרש במדגם מזווג. להלן השערותיו :

$$H_0 : \mu_D = 80$$

$$H_1 : \mu_D \neq 80$$

$$\alpha = 5\%$$

החוקר בנה רווח סמך ברמה של 90%

$$78 < \mu_D < 83$$

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

תרגילים:

1. נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'. להלן ציוניהם בסמסטר א' ו- ב' :

סטטיסטיקה א	סטטיסטיקה ב
74	80
68	84
90	87
75	76
82	100

- א. בנו רווח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סמסטר א לבין סמסטר ב.
ב. פורסם שתלמידים שמסיימים את סמסטר ב משפרים בממוצע את הציונים ב-5 נק' לעומת סמסטר א' האם יש אמת בפרסום?

2. הוחלט להשוות הציונים אצל מרצה X ואצל מרצה Y. נבחרו באקראי 6 סטודנטים, 3 סטודנטים של מרצה X ו- 3 סטודנטים של מרצה Y, עבורם התקבלו הציונים הבאים :

68	90	82	מרצה X
64	81	68	מרצה Y

- א. חשבו רווח סמך ברמת סמך 90% להפרש בין התוחלות של הציונים אצל שני המרצים.
ב. האם ברמת מובהקות של 10% נכריע שיש הבדל בין תוחלות הציונים אצל שני המרצים?

שאלות אמריקאיות:

3. סטטיסטיקאי נתבקש לאמוד את הפרש הממוצעים של שני טיפולים לפי שני מדגמים מקריים בלתי תלויים.
הוא חישב רווח סמך להפרש ברמת סמך 0.98, וקיבל את הרווח $-2 < \mu_1 - \mu_2 < 4.5$.
אילו יתבקש החוקר לבדוק לפי אותם נתונים את ההשערות :
 $H_0 : \mu_1 - \mu_2 = 0$; $H_1 : \mu_1 - \mu_2 \neq 0$, ברמת מובהקות 0.05 מסקנתו תהיה :
א. לדחות את השערת האפס.
ב. לא לדחות את השערת האפס.
ג. שלא ניתן לדעת את המסקנה עבור רמת מובהקות 0.05 .
ד. שלא נתונות בשאלה סטיות התקן של האוכלוסיות, ולכן לא ניתן להסיק דבר.

4. במטרה לבדוק האם קיים הבדל בין קווי זהב לבזק מבחינת ממוצע המחירים לשיחות בינ"ל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה. בהנחה והמחירים מתפלים נורמלית בנו ריוח סמך לממוצע ההפרשים וקיבלו :
- $$-0.0293 < \mu_d < 0.2145$$
- רווח הסמך הוא ברמת סמך של 95%.
- לכן מסקנת המחקר היא :
- א. ברמת מובהקות של 5% לא נוכל לקבוע שקיים הבדל בין החברות.
- ב. ברמת מובהקות של 5% נקבע שקיים הבדל מובהק בין החברות.
- ג. לא ניתן לדעת מה המסקנה ברמת מובהקות של 5% כיוון שלא נאמר מה ההגדרה של D.

פתרונות:

שאלה 1:

א. $-3.8 \leq \mu_D \leq 19$

ב. נכריע שיש אמת בפרסום.

שאלה 2:

א. $-8.5 \leq \mu_X - \mu_Y \leq 26.5$

ב. נכריע שאין הבדל.

שאלה 3:

התשובה היא ג.

שאלה 4

התשובה היא א.

פרק 32 - ניתוח שונות חד כיוונית

רקע תיאורטי:

ניתוח שונות (חד כיווני) הוא מבחן להשוואת תוחלות ($\mu_1, \dots, \mu_k$) של k אוכלוסיות שונות.

ולכן בניתוח שונות השערות המחקר הן:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k \quad (\text{התוחלות של כל האוכלוסיות שוות})$$

$$H_1: \quad \text{אחרת} \quad (\text{לפחות שתיים מהתוחלות שונות})$$

ההנחות הדרושות לביצוע התהליך הן:

1. בכל אוכלוסייה מתוך k האוכלוסיות ההתפלגות נורמלית.

2. כל האוכלוסיות הן עם אותה שונות σ^2 .

3. המדגמים בלתי תלויים זה בזה.

ישנו משתנה המבדיל בין הקבוצות השונות, הוא המשתנה הבלתי תלוי הנקרא גורם (factor)

משתנה זה הוא קטגוריאלי עם k רמות (levels).

כדי לבצע את התהליך יש לבצע מדגם מכל אוכלוסייה:

נסמן ב- n_i את גודל המדגם בקבוצה i.

$$n = \sum_{i=1}^k n_i \quad - \text{ מספר התצפיות סך הכול (בכל המדגמים)}$$

$$\bar{X}_1 - \text{ ממוצע המדגם הראשון, } \dots, \bar{X}_k - \text{ ממוצע המדגם ה-}k\text{-י.}$$

$$\bar{X} - \text{ ממוצע כללי (של כל המדגמים).}$$

$$SS_B = \sum_{i=1}^k n_i [\bar{X}_i - \bar{X}]^2 \quad \text{סכום ריבועים בין הקבוצות}$$

$$SS_W = \sum_{i=1}^k [n_i - 1] \cdot \hat{S}_i^2 \quad \text{סכום ריבועים בתוך הקבוצות}$$

$$SS_T = \sum_{i=1}^k \sum_{j=1}^{n_j} [X_{ij} - \bar{X}]^2 : \text{ סכום ריבועים כללי}$$

$$SST = SSB + SSW$$

יש למלא את טבלת ניתוח השונות הבאה :

טבלת ניתוח שונות

מקור השונות	סכום הריבועים SS	דרגות חופש df	ממוצע הריבועים MS	F
B-בין הקבוצות	SSB	k - 1	$\frac{SSB}{k-1}$	$\frac{MSB}{MSW}$
W-בתוך הקבוצות	SSW	n - k	$\frac{SSW}{n-k}$	
T-סה"כ	SST	n - 1		

$$F = \frac{SS_B / (k-1)}{SS_W / (n-k)} \sim F(k-1, n-k)$$

$$F > F_{(k-1), (n-k); 1-\alpha} : H_0 \text{ איזור דחיית}$$

תרגילים:

1. מחקר מעוניין להשוות בין שלוש תרופות לשיכוך כאבים במטרה לבדוק האם קיים הבדל בין התרופות מבחינת הזמן בדקות שלוקח עד שהתרופה משפיעה. לצורך הבדיקה נלקחו 15 אנשים שסובלים מכאבי ראש. אנשים אלה חולקו באקראי לשלוש : קבוצה 1 קיבלה "אקמול" קבוצה 2 קיבלה "אופטלגין" קבוצה 3 קיבלה "נורופן". כל אדם במחקר מסר את מספר הדקות עד שהתרופה השפיעה עליו.
 - א. מהו המשתנה התלוי ומהו המשתנה הבלתי תלוי במחקר? מהו ה"גורם" וכמה רמות יש לו?
 - ב. מהו המבחן הסטטיסטי המתאים כאן? רשמו את ההשערות.
 - ג. מה הן ההנחות הדרושות כדי לבצע את המבחן הסטטיסטי שהצעת בסעיף הקודם?

2. בעיר מסוימת שלושה בתי ספר תיכון. ראש העיר התעניין לבדוק האם קיים הבדל בהצלחה של בתי הספר במקצוע מתמטיקה. לצורך כך הוא דגם מספר תלמידים שנבחנו במבחן הבגרות במתמטיקה ברמה של 3 יחידות בעירו ובדק עבור כל תלמיד מה ציון הבגרות שלו במתמטיקה. להלן הציונים שהתקבלו:

בית הספר	"המתמיד"	"רבין"	"הס"
	78	98	85
	65	62	83
	70	55	74
	90	80	85
	56		75

- א. מהו המבחן הסטטיסטי המתאים? רשמו את ההשערות ואת ההנחות של המבחן.
- ב. מהו גודל המדגם? מהו המשתנה הבלתי תלוי (FACTOR) כמה רמות יש לו?
- ג. חשבו את הממוצע ואת סטיית התקן של הציונים בכל אחד מהמדגמים.
- ד. מלאו את טבלת ANOVA.
- ה. רשמו את כלל ההכרעה למבחן שהוצע בסעיף א ברמת מובהקות של 5%.
- ו. האם קיים הבדל בין בתי הספר בעיר מבחינת רמת הצלחת התלמידים במקצוע המתמטיקה? ענה על סמך הסעיפים הקודמים.

3. מעוניינים לבדוק האם יש הבדל בהשפעה של שיטות טפול שונות על לחץ הדם הסיסטולי (SBP) באוכלוסייה של קשישים. נבדקו 4 שיטות שונות. בטבלה המצורפת מרוכזים ממצאי המחקר.

השיטה	A	B	C	D
גודל המדגם	12	14	8	12
הממוצע	178	172	180	182
סטיית התקן	4	8	5	3

א. רשמו את השערות המחקר וההנחות הדרושות כדי לבצע את המבחן המתאים.

ב. מה מסקנת המחקר ברמת מובהקות של 5%?

ג. האם יש צורך לבצע השוואות מרובות?

4. שלושה אופים נתבקשו להכין עוגת שוקולד. לכל אופה בדקו את משך זמן ההכנה בדקות.

כל אופה נדרש לאפות בכל יום 4 עוגות.

האופה	ניר	מוזס	שלום
סכום הזמנים	206	212	182
סכום ריבועי הזמנים	10644	11250	8982

האם קיים הבדל בין האופים מבחינת תוחלת זמני ההכנה של העוגות? בדקו ברמת מובהקות של 5%.

5. להלן טבלת ניתוח שונות חד כיוונית. במחקר בחנו 4 סוגי סוללות. רצו לבדוק האם לסוג הסוללה השפעה על תוחלת אורך החיים שלה. הפעילו את כל הסוללות על אותו מכשיר ובדקו את אורך החיים של כל סוללה בשעות.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10.317	3	3.439	1.361	.279
Within Groups	60.648	24	2.527		
Total	70.964	27			

מה המסקנה ברמת מובהקות של 10%? רשמו את ההשערות וההנחות הדרושות.

6. להלן טבלת ANOVA בטבלה הושמטו חלקים. השלם את החלקים בטבלה שהושמטו ומסומנים באותיות.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	357.450	ב	ג	ה	.000
Within Groups	א	17	ד		
Total	522.950	19			

7. חברת תרופות לקחה 15 אנשים ברמת בריאות דומה. החברה חילקה את האנשים ל שלוש קבוצות שוות בגודלן. לכל קבוצה ניתנה אותה תרופה במינון שונה (dosage). המינונים שניתנו הם: 10 מ"ג, 20 מ"ג ו- 30 מ"ג. לאחר שעה מזמן לקיחת התרופה ניבדק קצב פעימות הלב של כל אדם (pulse).
הנתונים הוזנו לתוכנה סטטיסטית והתקבלו התוצאות הבאות:

ANOVA

pulse

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	414.400	2	207.200	19.733	.000
Within Groups	126.000	12	10.500		
Total	540.400	14			

Post Hoc Tests

Multiple Comparisons

pulse
Tukey HSD

(I) dosage	(J) dosage	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
10.00	20.00	3.20000	2.04939	.299	-2.2675	8.6675
	30.00	12.40000*	2.04939	.000	6.9325	17.8675
20.00	10.00	-3.20000	2.04939	.299	-8.6675	2.2675
	30.00	9.20000*	2.04939	.002	3.7325	14.6675
30.00	10.00	-12.40000*	2.04939	.000	-17.8675	-6.9325
	20.00	-9.20000*	2.04939	.002	-14.6675	-3.7325

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

pulse

Tukey HSD^a

dosage	N	Subset for alpha = 0.05	
		1	2
30.00	5	71.0000	
20.00	5		80.2000
10.00	5		83.4000
Sig.		1.000	.299

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

- א. בדוק ברמת מובהקות של 5% האם קיים הבדל בין המינונים השונים מבחינת תוחלת הדופק של האנשים? רשמו את ההשערות וההנחות הדרושות לצורך פתרון.
- ב. הסבירו ללא חישוב כיצד הייתה משתנה התשובה לסעיף הקודם אם הינו מעלים את הדופק של כל התצפיות במחקר ב- 2.
- ג. האם יש צורך במחקר בהשוואת מרובות. נמק!
- ד. לטבלת ה ANOVA צורפו טבלאות של השוואות מרובות בשיטה הנקראת "טוקי". ברמת בטחון של 95% מה הם הממצאים לפי שיטה זו?

8. בעיר מסוימת רצו לבדוק האם קיים הבדל ברמה של התלמידים בין בתי הספר השונים בעיר. ביצעו מדגם מכל בית ספר ונתנו מבחן זהה לכל הנדגמים. לאחר מכן ריכזו את הנתונים בתוכנה סטטיסטית והפעילו ניתוח שונות. מצורפים הפלטים שהתקבלו.
- ענו על הסעיפים הבאים:
- א. כמה בתי ספר יש בעיר?
- ב. כמה תלמידים השתתפו בסך הכול במחקר?
- ג. האם קיים הבדל בין בתי הספר בעיר מבחינה רמת הציונים? בדקו ברמת מובהקות של 1%
- ד. בביטחון של 95% אילו בתי ספר שונים זה מזה ברמת התלמידים? נמקו והסבירו.

Oneway

ANOVA

grade					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7799.600	4	1949.900	13.586	.000
Within Groups	2870.400	20	143.520		
Total	10670.000	24			

Post Hoc Tests

Multiple Comparisons

grade

Scheffe

(I) school	(J) school	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	5.40000	7.57681	.971	-20.2543	31.0543
	3.00	36.80000*	7.57681	.003	11.1457	62.4543
	4.00	36.40000*	7.57681	.003	10.7457	62.0543
	5.00	-2.60000	7.57681	.998	-28.2543	23.0543
2.00	1.00	-5.40000	7.57681	.971	-31.0543	20.2543
	3.00	31.40000*	7.57681	.011	5.7457	57.0543
	4.00	31.00000*	7.57681	.013	5.3457	56.6543
	5.00	-8.00000	7.57681	.888	-33.6543	17.6543
3.00	1.00	-36.80000*	7.57681	.003	-62.4543	-11.1457
	2.00	-31.40000*	7.57681	.011	-57.0543	-5.7457
	4.00	-.40000	7.57681	1.000	-26.0543	25.2543
	5.00	-39.40000*	7.57681	.001	-65.0543	-13.7457
4.00	1.00	-36.40000*	7.57681	.003	-62.0543	-10.7457
	2.00	-31.00000*	7.57681	.013	-56.6543	-5.3457
	3.00	.40000	7.57681	1.000	-25.2543	26.0543
	5.00	-39.00000*	7.57681	.001	-64.6543	-13.3457
5.00	1.00	2.60000	7.57681	.998	-23.0543	28.2543
	2.00	8.00000	7.57681	.888	-17.6543	33.6543
	3.00	39.40000*	7.57681	.001	13.7457	65.0543
	4.00	39.00000*	7.57681	.001	13.3457	64.6543

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

grade

Scheffe^a

school	N	Subset for alpha = 0.05	
		1	2
3.00	5	45.0000	
4.00	5	45.4000	
2.00	5		76.4000
1.00	5		81.8000
5.00	5		84.4000
Sig.		1.000	.888

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

פתרונות סופיים חלקיים - ניתוח שונות חד כיוונית

2. אם חישובת נכון ה F הסטטיסטי יוצא: 0.58

3. נדחה את השערת האפס.

4. להלן טבלת הניתוח השונות המתקבלת:

	Sum of Squares	df	Mean Square	F
Between Groups	126.000	2	63.000	.756
Within Groups	750.000	9	83.333	
Total	876.000	11		

5. נקבל את השערת האפס.

6. א. 165.5 ב. 2 ג. 178.725 ד. 9.375 ה. 18.36

7. א. נדחה את השערת האפס. ב. לא משתנה. ג. כן

8. א. 5 ב. 25 ג. כן