

סטודנטים יקרים

לפניכם ספר תרגילים בקורס סטטיסטיקה מושגי יסוד. 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 לחוויית לימודים יוצאת דופן וחדשה לחלוטין, היכנסו עכשיו לאתר www.gool.co.il.



אנו מאחלים לכם הצלחה מלאה בבחינות

צוות האתר GooL

גול זה בול. בשבילך!

לפתרון מלא בסרטון וידאו היכנסו ל- www.Gool.co.il

© כתב ופתר - ברק קנדל

תוכן

| | |
|-----|---|
| 4 | פרק 1 - סטטיסטיקה תיאורית - סיווג משתנים וסולמות מדידה |
| 7 | פרק 2 - סטטיסטיקה תיאורית - הצגה של נתונים |
| 15 | פרק 3 - סטטיסטיקה תיאורית - סכימה |
| 18 | פרק 4 - סטטיסטיקה תיאורית - מדדי מיקום מרכזי |
| 28 | פרק 5 - סטטיסטיקה תיאורית - מדדי פיזור: הטווח, השונות וסטיית התקן |
| 33 | פרק 6 - סטטיסטיקה תיאורית - מדדי פיזור- טווח בין- רבעוני |
| 36 | פרק 7 - סטטיסטיקה תיאורית - ממוצע משוקלל ושונות מצורפת |
| 39 | פרק 8 - סטטיסטיקה תיאורית - מדדי מיקום יחסי - ציון תקן |
| 42 | פרק 9 - סטטיסטיקה תיאורית - מדדי מיקום יחסי - אחוזונים במחלקות |
| 47 | פרק 10 - סטטיסטיקה תיאורית - טרנספורמציה לינארית |
| 50 | פרק 11 - סטטיסטיקה תיאורית - שאלות מסכמות |
| 59 | פרק 12 - בעיות בסיסיות בהסתברות |
| 64 | פרק 13 - המשתנה המקרי הבדיד - פונקציית ההסתברות |
| 66 | פרק 14 - המשתנה המקרי הבדיד - תוחלת, שונות וסטיית תקן |
| 69 | פרק 15 - המשתנה המקרי הבדיד - טרנספורמציה לינארית |
| 72 | פרק 16 - התפלגויות בדידות מיוחדות - התפלגות בינומית |
| 76 | פרק 17 - התפלגויות בדידות מיוחדות - התפלגות גיאומטרית |
| 79 | פרק 18 - התפלגויות רציפות מיוחדות - התפלגות אחידה |
| 82 | פרק 19 - התפלגויות רציפות מיוחדות - התפלגות נורמלית |
| 90 | פרק 20 - התפלגות הדגימה |
| 90 | ממוצע המדגם ומשפט הגבול המרכזי |
| 98 | התפלגות סכום תצפיות המדגם ומשפט הגבול המרכזי |
| 101 | התפלגות מספר ההצלחות במדגם - הקרוב הנורמלי להתפלגות הבינומית |
| 106 | התפלגות פרופורציית ההצלחות במדגם |
| 111 | פרק 21 - הסקה סטטיסטית - הקדמה |
| 114 | פרק 22 - מושגים בסיסיים באמידה |
| 121 | פרק 23 - רווח סמך לתוחלת (ממוצע האוכלוסייה) |
| 121 | רווח סמך כששונות האוכלוסייה ידועה |
| 127 | קביעת גודל מדגם באמידת תוחלת עם שונות אוכלוסייה ידועה |
| 130 | רווח סמך לתוחלת (ממוצע האוכלוסייה) כששונות האוכלוסייה אינה ידועה |
| 136 | פרק 24 - רווח סמך לפרופורציה |
| 139 | קביעת גודל מדגם באמידת פרופורציה |

| | |
|-----|--|
| 143 | פרק 25 - בדיקת השערות על פרמטרים |
| 143 | הקדמה |
| 146 | טעויות בבדיקת השערות |
| 148 | פרק 26 - בדיקת השערות על תוחלת (ממוצע) |
| 148 | כאשר שונות האוכלוסיה ידועה |
| 153 | סיכוי לטעויות ועוצמה כאשר שונות האוכלוסייה ידועה |
| 160 | מובהקות התוצאה (P-VALUE) בבדיקת השערות על תוחלת עם שונות ידועה |
| 165 | הקשר בין רווח סמך לבדיקת השערות על תוחלת |
| 168 | פרק 27 - בדיקת השערות על פרופורציה |
| 168 | התהליך |
| 172 | סיכוי לטעויות ועוצמה |
| 177 | מובהקות התוצאה |

פרק 1 - סטטיסטיקה תיאורית - סיווג משתנים וסולמות מדידה

רקע:

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

מיון משתנים לפי סולמות המדידה:

1. **סולם שמי** (נומינאלי) – משתנה שלערכיו יש משמעות רק מבחינת הזהות ואין עניין של יותר או פחות לדוגמה: מצב משפחתי רווק/נשוי/אלמן/גרש; אזור מגורים. משתנה דיכוטומי (הינו מסולם שמי) אותם משתנים שיש להם רק שני ערכים אפשריות זכר/נקבה. מעשך/לא מעשך.
2. **סולם סדר** (אורדינאלי) – כאשר לערכים של המשתנה בנוסף לשם ישנה גם משמעות לסדר אבל אין משמעות לגודל ההפרש. למשל, דרגה בצבא.
3. **סולם רווחים** (אינטרוואלי) – משתנה שלערכים שלו בנוסף לשם ולסדר בניהם יש משמעות לרווחים בין הערכים אבל אין משמעות ליחס בין הערכים. למשל, קומה בבניין. סולם לא כל כך פופולרי.
4. **סולם מנה/יחס** – משתנה שלערכיו בנוסף לשם, לסדר ולרווח יש משמעות גם ליחס בין הערכים. למשל, מספר מכוניות למשפחה, משקל אדם בק"ג. הדרך הקלה ביותר כדי לזהות עם הסולם הוא סולם מנה היא על ידי מבחן האפס. בסולם מנה האפס הוא מוחלט, אבסולוטי, ומייצג אין.

סוגי משתנים:

נבצע סיווג של המשתנים :

משתנה איכותי הוא משתנה שלערכיו אין משמעות של יותר או פחות , אין עניין כמותי לערכים המתקבלים.

כמו : מקום מגורים של אדם (רעננה, תל אביב, אשדוד..)

מין האדם (זכר, נקבה)

מצב משפחתי (רווק, נשוי, גרוש, אלמן)

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

את המשתנה הכמותי נסווג לשני סוגים :

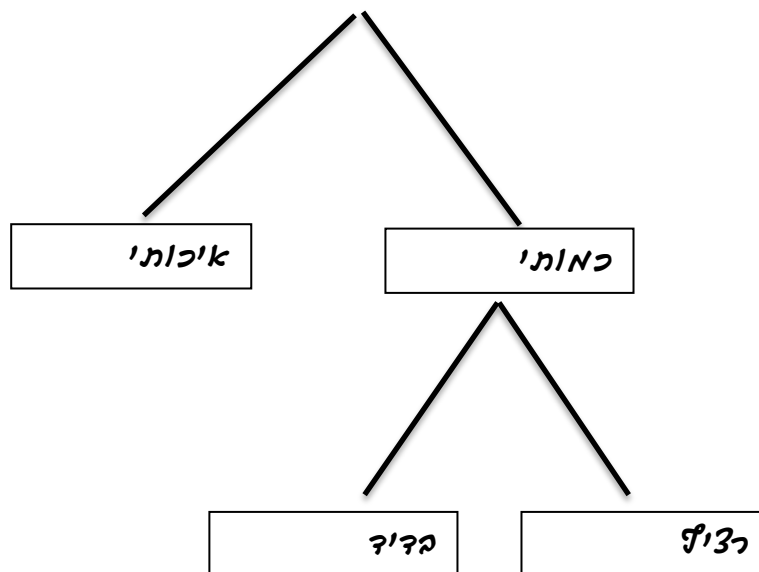
משתנה בדיד : משתנה שערכיו מתקבלים מתוך סידרה של ערכים אפשריים. כמו : מספר ילדים למשפחה (1,2,3..)

ציון בבחינה (מ 0 ועד 100 בקפיצות של 1)

משתנה רציף : משתנה שערכיו מתקבלים מתוך אינסוף ערכים בתחום מסוים , הערכים מתקבלים ברצף וללא קפיצות של ערכים .

כמו : גובה בס"מ – אם למשל, הגובה הנמוך ביותר הוא 150 ועד 190 ס"מ בקבוצה הגבהים הם ברצף. גם בין 160 ל 161 ס"מ יש רצף אינסופי של ערכים אפשריים לגובה (16.233 ס"מ הוא גם גובה אפשרי)

משקל בק"ג , מהירות בקמ"ש וכולי.



תרגילים:

1. לפניכם רשימה של משתנים :

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

ציינו באיזה סולם מדידה המשתנה הנחקר (שמי , סדר, רווחים או מנה)

2. להלן התפלגות מספר האיחורים לעבודה בחודש של העובדים בחברת "סטאר".
בחברה 200 עובדים.

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

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

3. לפניכם רשימה של משתנים כמותיים . ציין ליד כל משתנה אם הוא רציף או בדיד.

- א. שכר עובד בש"ח.
- ב. ציון בחינת בגרות.
- ג. תוצאה בהטלת קובייה.
- ד. מהירות ריצה בתחרות.
- ה. שיעור התמיכה בממשלה.

פרק 2 - סטטיסטיקה תיאורית - הצגה של נתונים

רקע:

דרכים להצגת נתונים שנאספו:

א. רשימה של תצפיות:

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

3 4 3 5 4

ב. טבלת שכיחויות בדידה:

| שם המשתנה - X | שכיחות - $f(X)$ | שכיחות יחסית באחוזים |
|---------------|------------------------|----------------------------|
| X_1 | f_1 | $\frac{f_1}{N} \times 100$ |
| X_2 | f_2 | $\frac{f_2}{N} \times 100$ |
| X_3 | f_3 | $\frac{f_3}{N} \times 100$ |
| \vdots | \vdots | \vdots |
| X_k | f_k | $\frac{f_k}{N} \times 100$ |
| סה"כ | $N = \sum_{i=1}^k f_i$ | 100% |

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

למשל, להלן התפלגות הציונים בכיתה מסוימת:

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

שכיחות מצטברת – צבירה של השכיחויות: F_i - השכיחות המצטברת נותנת כמה תצפיות קטנות או שוות לערך.

שכיחות יחסית (פרופורציה) – השכיחות מחולקת לכמות התצפיות הכללי: $\frac{f_i}{n}$ - איזה חלק מהתצפיות בקבוצה שוות לערך.

ג. טבלת שכיחויות במחלקות:

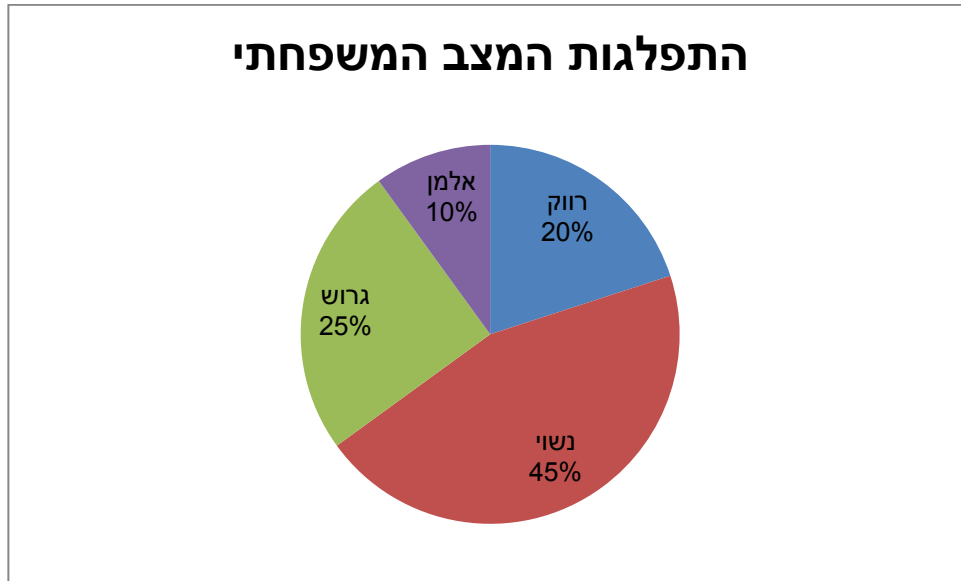
משתמשים שהמשתנה כמותי רציף או כאשר יש מספר ערכים רב במשתנה הבדיד וטבלת שכיחויות תהיה ארוכה מידי.

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

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

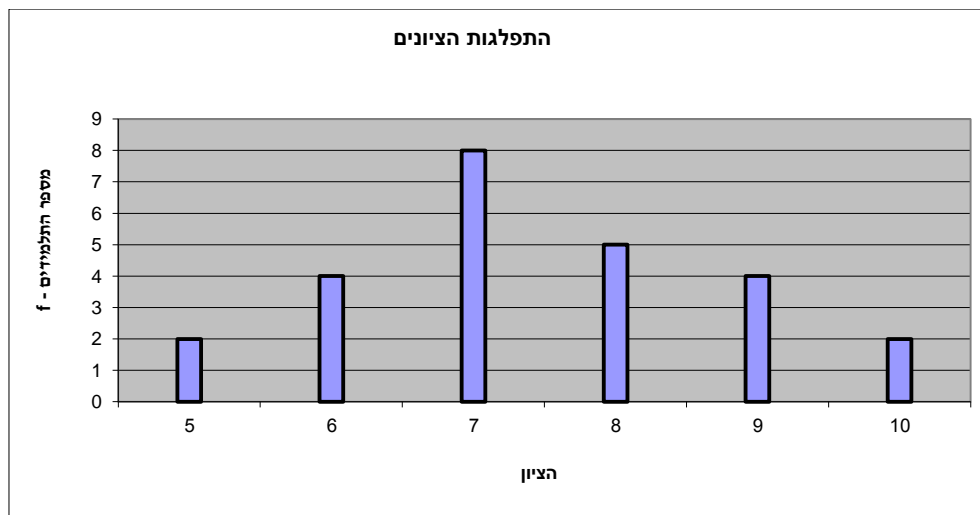
ד. **דיאגרמת עוגה :**

זהו התיאור הגרפי של משתנה איכותי. בדיאגרמת עוגה כל ערך במשתנה מקבל "נתח" יחסי מהעוגה. הנתח בעוגה פרופורציוני לשכיחות היחסית של ערך המשתנה בנתונים.



ה. **דיאגרמת מקלות :**

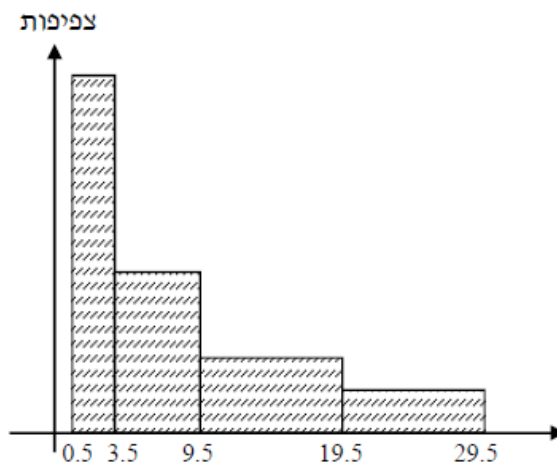
הציר האופקי הוא הציר של המשתנה הציר האנכי של השכיחות – הגובה של המקל מעיד על השכיחות .
 רלבנטי למשתנה כמותי בדיד. לא נהוג להשתמש בתיאור למשתנה איכותי וכמו כן לא למשתנה כמותי רציף .
 כמו כן בסולמות מדידה עבור משתנה מסולם סדר.



1. היסטוגרמה:

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

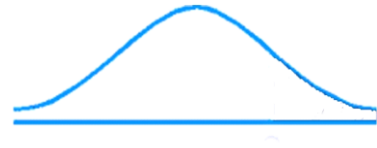
| צפיפות | מצטברת | שכיחות | אמצע | רוחב | X |
|--------|--------|--------|------|------|-------------|
| 6.6667 | 20 | 20 | 2 | 3 | 0.5 - 3.5 |
| 3 | 38 | 18 | 6.5 | 6 | 3.5 - 9.5 |
| 1.4 | 52 | 14 | 14.5 | 10 | 9.5 - 19.5 |
| 0.8 | 60 | 8 | 24.5 | 10 | 19.5 - 29.5 |



פוליגון-מצולעון: אם נחבר את אמצע קצה כל מלבן בקווים ישרים. נותן מראה חזותי לצורה של התפלגות המשתנה.

צורות התפלגות נפוצות

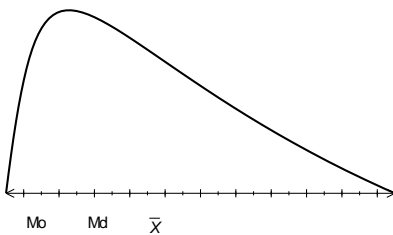
התפלגות סימטרית פעמונית- רוב התצפיות במרכז וככל שנתרחק מהמרכז יהיו פחות תצפיות באופן סימטרי. למשל, ציוני IQ.



ישנן התפלגויות סימטריות שאינן פעמוניות:

התפלגות אסימטרית ימנית (חיובית) – רוב התצפיות מקבלות ערכים נמוכים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים גבוהים קיצוניים. למשל, שכר במשק.

התפלגות א-סימטרית ימנית או חיובית



התפלגות אסימטרית שמאלית (שלילית) רוב התצפיות מקבלות ערכים גבוהים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים נמוכים קיצוניים. למשל, אורך חיים.

התפלגות א-סימטרית שמאלית או שלילית



תרגילים:

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

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

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

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

3. להלן התפלגות ההשכלה במקום עבודה מסוים:

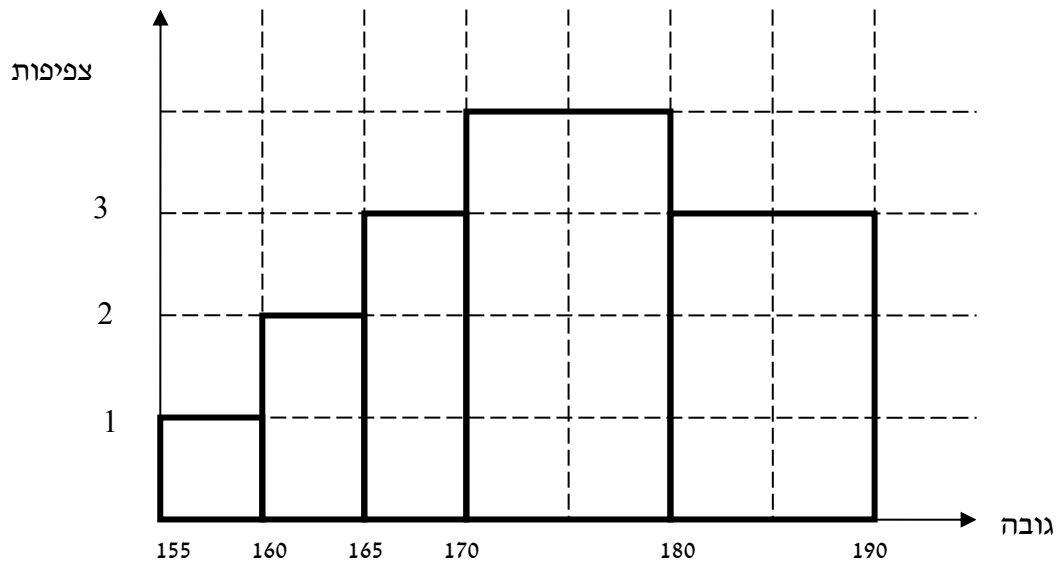
| מספר העובדים | השכלה |
|--------------|---------|
| 60 | נמוכה |
| 120 | תיכונית |
| 20 | אקדמאית |

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

4. להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
- 7, 6, 8, 9, 10, 6, 4, 5, 8, 7, 6, 7, 6, 8, 9, 6, 7, 8, 5, 6

- א. מהו המשתנה? האם הוא בדיד או רציף?
- ב. תאר את הרשימה בטבלת שכיחות.
- ג. הוסף שכיחות יחסיות לטבלה.
- ד. תאר את הנתונים באופן גרפי.

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



- א. מהו המשתנה הנחקר? האם הוא בדיד או רציף?
 ב. תאר את הנתונים בטבלת שכיחויות במחלקות.
 ג. הוסף שכיחות יחסית לטבלה.
 ד. הוסף את הצפיפות של כל מחלקה לטבלה.
 ה. מהי צורת ההתפלגות של הגבהים?

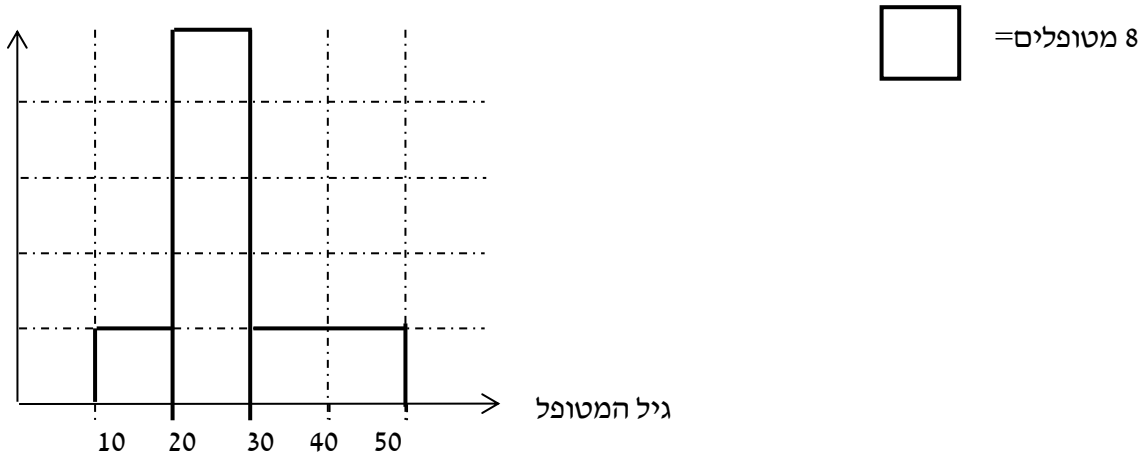
6. להלן התפלגות המשקל של קבוצה מסוימת בק"ג:

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

- א. תאר את ההתפלגות באופן גרפי.
 ב. מה ניתן להגיד על צורת ההתפלגות?

7. להלן גיל המטופלים של ד"ר שוורץ בשנים :

קנה מידה :



א. מה המשתנה הנחקר? האם הוא בדיד או רציף?

ב. מהי הקבוצה הנחקרת?

ג. תרגמו את ההסיטוגרמה לטבלת שכיחות.

ד. מהי הפרופורציה של המטופלים של ד"ר שוורץ בגילאים 20-30?

פרק 3 - סטטיסטיקה תיאורית - סכימה

רקע:

בסטטיסטיקה ישנה צורת רישום מקובלת כדי לרשום סכום של תצפיות:

$$\sum_{i=1}^n X_i$$

נסביר את צורת הרישום על ידי הדוגמה הבאה:

| i | X_i |
|---|-------|
| 1 | 5 |
| 2 | 0 |
| 3 | 1 |
| 4 | 3 |
| 5 | 2 |

תרגילים:

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

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

חשבו:

$$\sum_{i=1}^3 X_i$$

$$\sum_{i=1}^5 Y_i$$

$$\sum_{i=1}^4 X_i$$

$$\left(\sum_{i=1}^4 X_i\right)^2$$

$$\sum X_i$$

$$\sum X_i Y_i$$

$$\sum (X_i) \sum (Y_i)$$

2. נתון לוח ערכי המשתנים x_i ו- y_i כאשר: $i=1,2,\dots,6$

| | | | | | | |
|-------|---|---|---|----|----|---|
| i | 1 | 2 | 3 | 4 | 5 | 6 |
| x_i | 3 | 2 | 4 | -2 | 1 | 4 |
| y_i | 2 | 0 | 0 | 1 | -5 | 2 |

ונתונים הקבועים: $a=2$ $b=5$ חשבו את הנוסחאות הבאות:

$$\sum_{i=1}^4 y_i \quad \text{א.}$$

$$\sum_{i=1}^6 a \quad \text{ב.}$$

$$\sum_{i=1}^6 x_i y_i \quad \text{ג.}$$

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

$$\sum_{i=1}^6 x_i + a \quad \text{ה.}$$

3. קבע לכל זהות אם היא נכונה:

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

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

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

4. נתון: $\sum_{i=1}^{10} X_i = 80$ חשב: $\sum_{i=1}^{10} (X_i - 4)^2$ (פתרון: 1160)

$$\sum_{i=1}^{10} X_i^2 = 1640$$

פרק 4 - סטטיסטיקה תיאורית - מדדי מיקום מרכזי

רקע:

המטרה במדדי המיקום המרכזי למדוד את מרכז ההתפלגות של התצפיות.

השכיח – MODE

השכיח הוא הערך הנפוץ ביותר בהתפלגות.

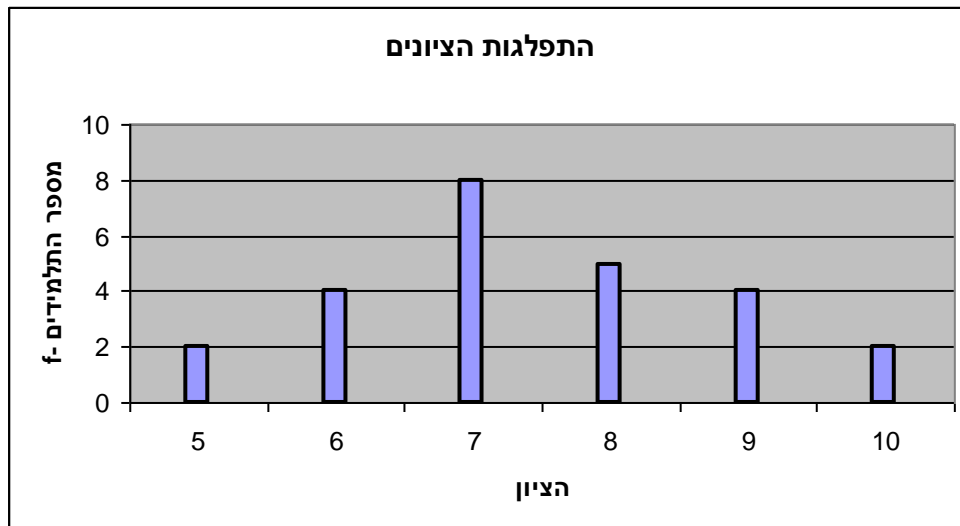
ברשימה : הערך החוזר על עצמו הכי הרבה פעמים .

7 9 4 8 4 10 6

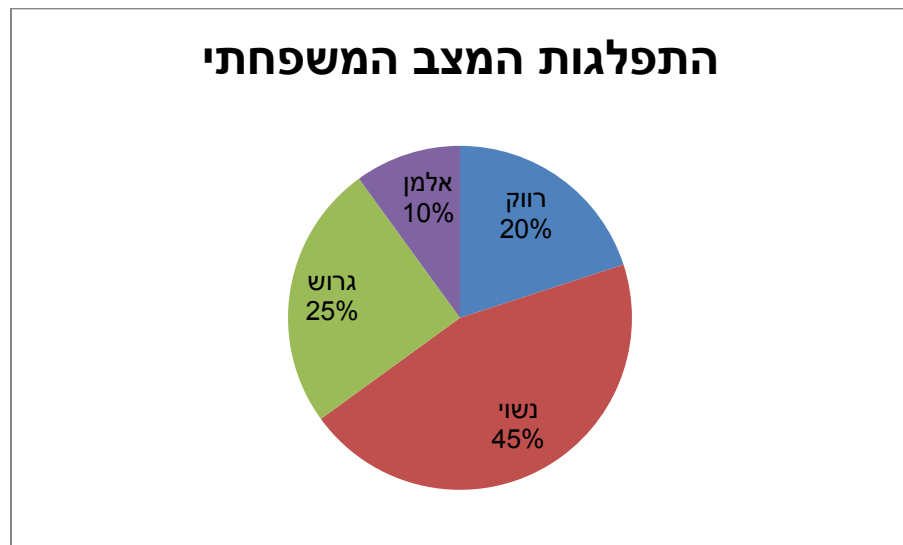
בטבלת שכיחויות בדידה : הערך שהשכיחות שלו היא הגבוהה ביותר.

| # תכניות החיסכון | $f(x)$ |
|------------------|--------|
| 0 | 100 |
| 1 | 75 |
| 2 | 25 |
| 3 | 25 |
| 4 | 25 |

בדיאגרמת מקלות : שיעור ה- X של המקל הגבוה ביותר.



בעוגה: הערך של הפלח הגדול ביותר.

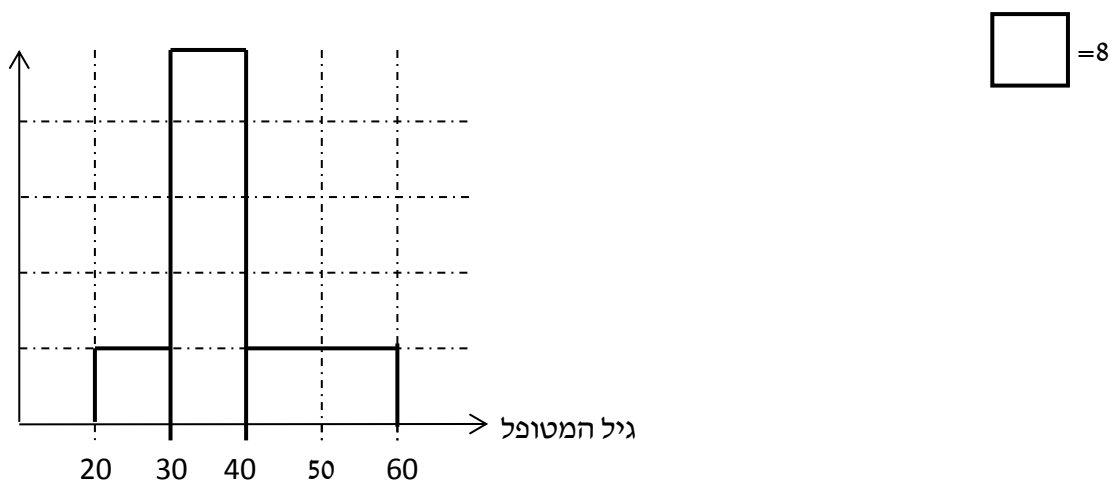


בטבלת שכיחויות במחלקות: אמצע המחלקה עם הצפיפות הגבוהה ביותר. התפלגות הציונים בכיתה.

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

בהיסטוגרמה: שיעור ה- X של אמצע המחלקה הגבוהה ביותר.

להלן גיל המטופלים של ד"ר שוורץ בשנים:



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

MIDRANGE – (טווח) – אמצע תחום

הממוצע בין התצפית הגבוהה ביותר לתצפית הנמוכה ביותר.

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

MEDIAN - החציון

החציון הוא ערך שמחצית מהתצפיות קטנות או שוות לו ומחצית מהתצפיות גדולות או שוות לו. ברשימה: נסדר את התצפיות בסדר עולה.

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

אם יש מספר זוגי של איברים החציון יהיה הממוצע של האיבר ה- $\frac{n}{2}$ והאיבר ה- $\frac{n}{2}+1$

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

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

בטבלת שכיחויות בדידה: נעשה תהליך דומה אך נעזר בשכיחות המצטברת.

דיאגרמת מקלות: נמיר לטבלת שכיחויות בדידה במטרה למצוא את החציון.

בטבלת שכיחויות במחלקות:

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

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

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

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

L_0 - גבול התחתון של המחלקה.

L_1 - גבול העליון של המחלקה.

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

כללי: החציון אינו רלבנטי למשתנה מסולם שמי ולא רלבנטי למשתנה איכותי.

הממוצע

הנו מרכז הכובד של ההתפלגות.

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

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

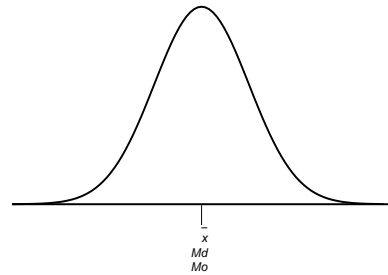
במחלקות: נשתמש באותה נוסחה רק נתייחס לאמצע המחלקה בתור ה X. הממוצע הזה יהיה ממוצע מקורב.

כללי: הממוצע רלבנטי רק למשתנה כמותי.

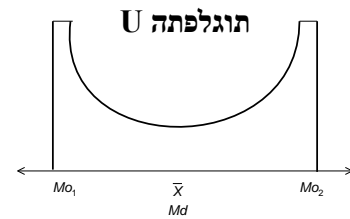
מדדי המיקום המרכזי בהתפלגויות המיוחדות:

בהתפלגות סימטרית פעמונית כל מדדי המרכז שווים זה לזה:

התפלגות סימטרית



בהתפלגות סימטרית השכיח לא חייב להיות במרכז:

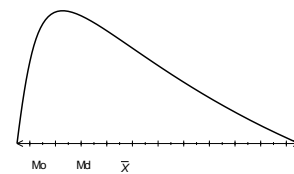


בהתפלגות אסימטרית

התפלגות
א-סימטרית
שמאלית או
שלילית



התפלגות א-סימטרית
ימנית או חיובית



תרגילים:

1. להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא :
 7, 6, 8, 9, 10, 6, 4, 5, 8, 7, 6, 7, 6, 8, 9, 6, 7, 8, 5, 6
 חשב את החציון, השכיח, והממוצע של הציונים.

2. בדקו את מספר החדרים לדירה בבניין בן 5 דירות והתקבל ממוצע 3.8
 לגבי 4 דירות נמצא מספר חדרים : 4, 3, 4, 5 .
 א. כמה חדרים יש בדירה החמישית?
 ב. מהו השכיח ומהו החציון?

3. להלן התפלגות מספר מקלטי הטלויזיה שנספרו עבור כל משפחה בישוב מסוים :

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

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

4. להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן"

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

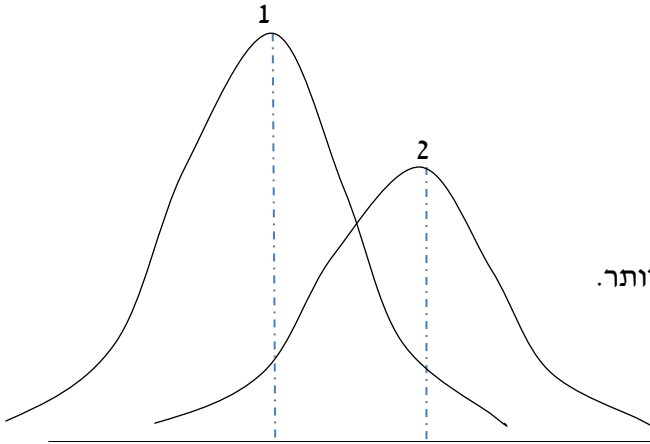
א. כמה משפחות יש בישוב?

ב. מה אחוז המשפחות בישוב עם לכל היותר 2 מכוניות?

ג. חשבו את הממוצע, החציון והשכיח.

הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם!

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



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

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

| מספר טלוויזיות | מספר משפחות |
|----------------|-------------|
| 0 | 28 |
| 1 | 62 |
| 2 | |
| 3 | |

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

7. להלן התפלגות המשקל של קבוצה מסוימת בק"ג :

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

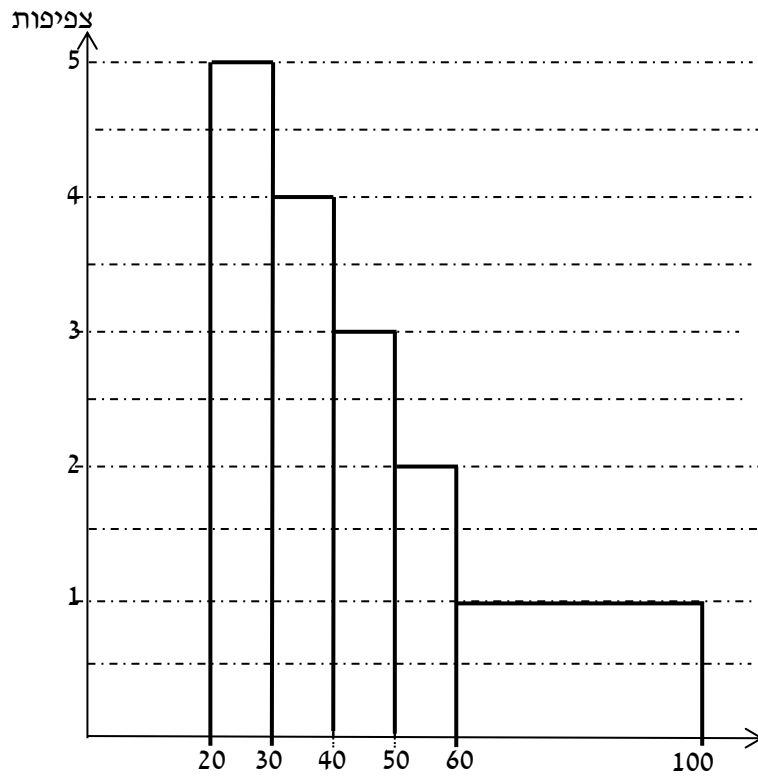
מה הממוצע והחציון של ההתפלגות?

8. להלן התפלגות הגבהים בס"מ בקבוצה מסוימת.

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

חשב את הממוצע, החציון והשכיח של הגבהים בקבוצה זו.

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



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

פתרונות:**שאלה 1:**

החציון : 7

השכיח : 6

הממוצע : 6.9

שאלה 2:

א. 3

ב. שכיח : 3,4 חציון : 4

שאלה 3:

א. הממוצע : 1.7

החציון : 1.5

השכיח : 1

ב. הממוצע יגדל ויתר המדדים לא ישתנו.

שאלה 4:

א. 630

ב. 34.13%

ג. שכיח וחציון : 3

ממוצע : 2.952

שאלה 5:

תשובה : ב

שאלה 6:

ב חציון : 2 שכיח : 2 אמצע טווח : 1.5

שאלה 7:

חציון וממוצע : 55

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

רקע:

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

הטווחותחום RANGE:

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

שונות וסטיית תקן:

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

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

דוגמה : נחשב את השונות של סדרת המספרים הבאה : 5,4,9

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

להלן התפלגות הציונים בכיתה מסוימת בה ממוצע הציונים הוא 7.44

| $x^2 \cdot f$ | השכיחות f | הציון X |
|---------------|-----------|---------|
| 50 | 2 | 5 |
| 144 | 4 | 6 |
| 392 | 8 | 7 |
| 320 | 5 | 8 |
| 324 | 4 | 9 |
| 200 | 2 | 10 |
| 1430 | | סה"כ |

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

$$s = \sqrt{s_x^2} = \sqrt{1.8464} = 1.3588$$

כשיש מחלקות נעזר באמצע המחלקה כדי לחשב את השונות.

תרגילים:

1. להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:

7, 6, 8, 9, 10, 6, 4, 5, 8, 7, 6, 7, 6, 8, 9, 6, 7, 8, 5, 6
 חשבו את השונות, סטיית התקן והטווח של הציונים.

2. להלן התפלגות מספר המכוניות למשפחה ב"הגורן"

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

א. חשבו סטיית התקן.

ב. חשבו את הטווח של הנתונים.

הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם!

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

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

ב. האם הממוצע יגדל/יקטן/לא ישתנה וסטיית התקן תגדל/תקטן/לא תשנה כאשר יתווספו שני

עובדים אשר אחד עם וותק של 0 שנים והשני עם וותק של 8 שנים להתפלגות?

4. נתונה רשימה של 5 תצפיות, אך רק עבור 4 מהן נרשמו הסטיות שלהן מהממוצע:

2, 3, 2, 1. - חשב את השונות של חמש התצפיות.

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

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

א. מה הממוצע של מספר החדרים לשכונה בדירה?

ב. חשבו את סטיית התקן של מספר החדרים לדירה.

ג. חלק מבעלי הדירות בנות 2 החדרים הפכו את דירתם לדירת חדר. כיצד הדבר ישפיע

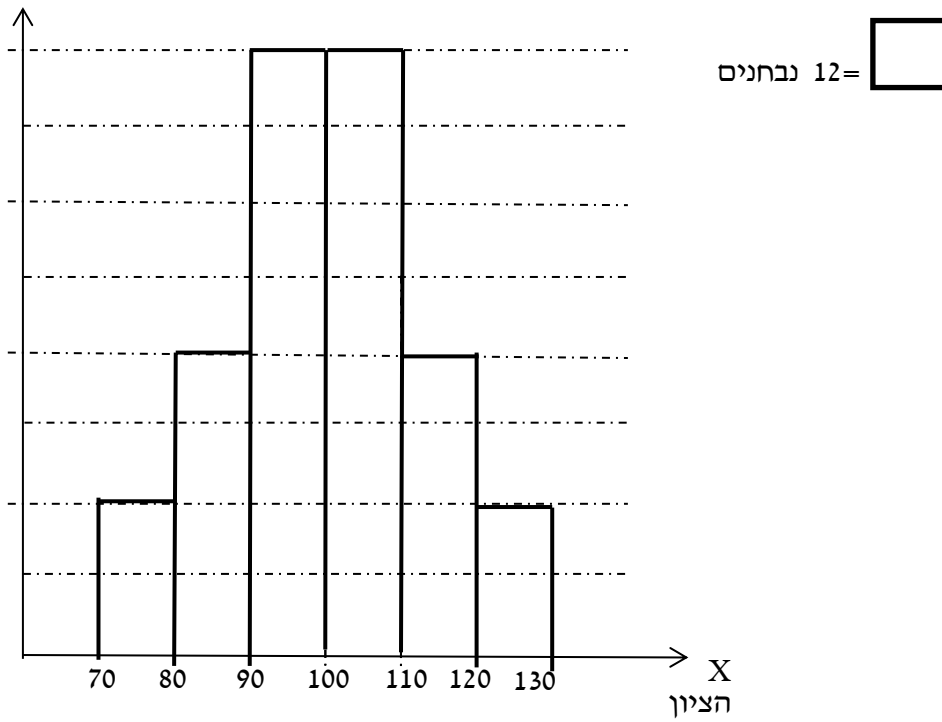
(יקטין, יגדל, לא ישנה) כל מדד שחישבתם בסעיפים הקודמים.

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

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

מהי סטיית התקן של התפלגות המשקל?

7. להלן התפלגות הציונים במבחן אינטליגנציה:



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

פתרונות :**שאלה 1:**

השונות : 2.19

סטיית תקן : 1.48

טווח : 6

שאלה 2:

א. סטיית תקן: 1.106

ב. טווח 4

שאלה 3:

א. ממוצע לא ישתנה, סטיית התקן תקטן.

ב. ממוצע לא ישתנה, סטיית התקן תגדל.

שאלה 4:

10.8

שאלה 5:

א. 3.05

ב. 1.16

שאלה 6:

7.73

שאלה 7 :

א. 100

ב. 12.96

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

רקע:

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

שלבים במציאת טווח בין-רבעוני במחלקות:

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

שלב א: נימצא את הרבעון התחתון (האחוזון ה-25) והרבעון העליון (האחוזון ה-75).

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

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

נוסחאות הרבעונים יהיו:

$$Q_1 = L_0 + \frac{\frac{n}{4} - F(x_{m-1})}{f(x_m)} \cdot (L_1 - L_0) \quad ; \quad Q_3 = L_0 + \frac{\frac{3n}{4} - F(x_{m-1})}{f(x_m)} \cdot (L_1 - L_0)$$

שלב ב: נחסר את הרבעונים:

$$IQR = Q_3 - Q_1$$

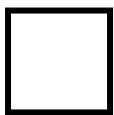
תרגילים:

1. להלן התפלגות המשקל של קבוצה מסוימת בק"ג:

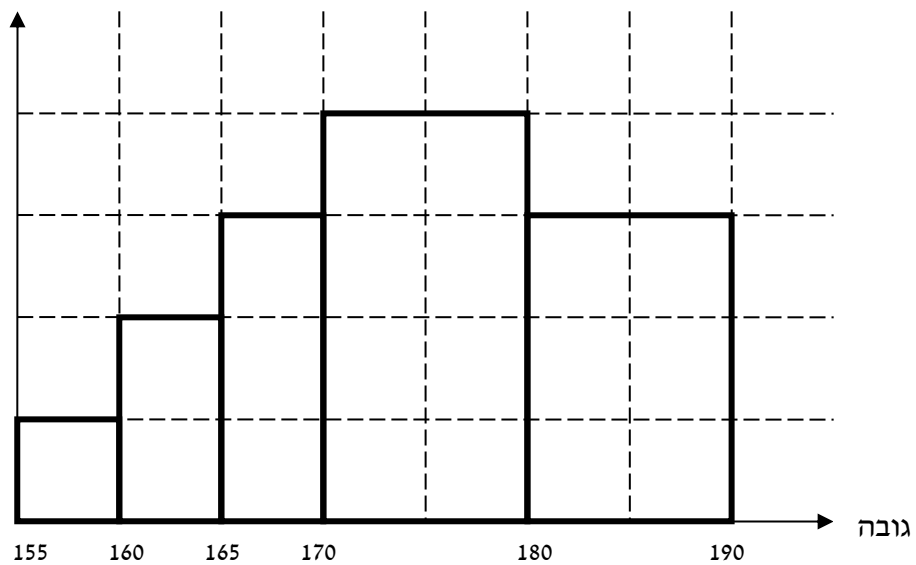
| משקל | מספר מקרים |
|-------|------------|
| 40-45 | 10 |
| 45-50 | 20 |
| 50-60 | 30 |
| 60-65 | 20 |
| 65-70 | 10 |

מצא את הטווח הבין-רבעוני.

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



אנשים = 5



מצא

את הטווח

הבין-

רבעוני.

פתרונות:

שאלה 1:

13.75

שאלה 2:

13.33

פרק 7 - סטטיסטיקה תיאורית - ממוצע משוקלל ושונות מצורפת

רקע:

מדובר על מצב שבו ישנן כמה קבוצות שרוצים לאחד אותן לקבוצה אחת גדולה. מתעניינים בממוצע והשונות של הקבוצה הגדולה המתקבלת מאיחוד הקבוצות הקטנות.

$$n_j = \text{מס' התצפיות בקבוצת ה-} j$$

$$= j \text{ אינדקס של הקבוצה.}$$

$$N = \text{מס' התצפיות בכל הקבוצות יחד (סכום כל ה-} n_j)$$

$$\bar{x}_j = \text{הממוצע בקבוצה ה-} j$$

$$S_j^2 = \text{השונות בקבוצה ה-} j$$

הנוסחאות לממוצע משוקלל ושונות מצורפת:

$$\bar{\bar{x}} = \frac{\sum_{j=1}^k \bar{x}_j n_j}{N} \quad ; \quad N = \sum_{j=1}^k n_j \quad ; \quad s_c^2 = \frac{\sum_{j=1}^k n_j s_j^2}{N} + \frac{\sum_{j=1}^k n_j (\bar{x}_j - \bar{\bar{x}})^2}{N}$$

דוגמה: (פתרון בהקלטה)

בחברה שני אגפים, אגף א' מונה עשרים עובדים, השכר הממוצע שם הוא 6,000 ₪ וסטיית התקן היא 2,000 ₪. באגף ב' עשרה עובדים השכר הממוצע הוא 12,000 ₪ וסטיית התקן היא 3,000 ₪.

מהו השכר הממוצע ומהי סטיית התקן של שכר העובדים בחברה?

תרגילים:

1. להלן נתונים לגבי ציונים במבחן באנגלית ב-3 כיתות מתוך שכבה י' בתיכון:

| כיתה | ממוצע | מס' תלמידים | סטיית תקן |
|------|-------|-------------|-----------|
| 1 | 76 | 40 | 12 |
| 2 | 68 | 20 | 15 |
| 3 | 82 | 30 | 10 |

א. חשב את הממוצע המשוקלל בשכבה.

ב. חשב את השונות המצורפת בשכבה.

2. נתונות שתי קבוצות:

בקבוצה I פי שתיים תצפיות מאשר בקבוצה II.

הממוצע בשתי הקבוצות הוא 70.

השונות בקבוצה I היא 100.

השונות בקבוצה II היא 400.

א. מצא את הממוצע של התצפיות לאחר שאוחדו שתי הקבוצות לקבוצה אחת.

ב. מצא את סטיית התקן של התצפיות לאחר שאוחדו שתי הקבוצות לקבוצה אחת.

פתרונות:**שאלה 2**

א. 70

ב. 14.14

שאלה 1

א. 76.22

ב. 173.5

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

רקע:

המטרה למדוד איך תצפית ממוקמות יחסית לשאר התצפיות בהתפלגות.

ציון תקן:

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

ציון התקן נותן כמה סטיות תקן סוטה התצפית מהממוצע.

כלומר, ציון התקן מעיד על כמה סטיות תקן התצפית מעל או מתחת לממוצע.

ציון תקן חיובי אומר שהתצפית מעל הממוצע.

ציון תקן שלילי אומר שהתצפית מתחת לממוצע.

ציון תקן אפס אומר שהתצפית בדיוק בממוצע.

דוגמה: (פתרון בהקלטה)

במקום עבודה מסוים ממוצע המשכורות 8 אלפי ₪ עם סטית תקן של 2 אלפי ₪ באותו מקום עבודה ההשכלה הממוצעת של העובדים הנה 14 שנים עם סטית תקן של 1.5 שנים. ערך מרויח במקום עבודה זה 11 אלף ₪ והשכלתו 16 שנים. מה ערך יותר באופן יחסי משכיל או משתכר ?

תרגילים

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

| המקצוע | ממוצע | סטיית תקן |
|---------|-------|-----------|
| לשון | 74 | 12 |
| מתמטיקה | 80 | 16 |

- עודד קיבל: 68 בלשון ו70 במתמטיקה.
 א. באיזה מקצוע עודד טוב יותר באופן יחסי לשכבה שלו?
 ב. איזה ציון עודד צריך לקבל במתמטיקה כדי שיהיה שקול לציונו בלשון?

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

| מספר פועלים | תפוקה | |
|-------------|-------|-----------|
| 15 | 48 | ממוצע |
| 2 | 10 | סטיית תקן |

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

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

פתרונות:

שאלה 1:

א. לשון

ב. 72

שאלה 2:

תשובה ב

שאלה 3:

א. משקל

ב. 70

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

רקע:

האחוזון (המאון) ה- p הוא הערך בנתונים המחלק את הנתונים בצורה כזאת שעד אליו יש $p\%$

מהנתונים. מסמנים את האחוזון ה- p ב- X_p .

למשל, המאון ה-25 הוא האחוזון ה-25 או הרבעון התחתון: ערך ש-רבע מהתצפיות קטנות

ממנו והשאר גבוהות ממנו. מסומן: $X_{0.25}$

מציאת מאון במחלקות:

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

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

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

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

L_0 - גבול התחתון של המחלקה.

L_1 - גבול העליון של המחלקה.

אם רוצים לחלץ את אחוז התצפיות שמתחת לערך מסוים נשתמש בנוסחה הבאה:

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

דוגמה : (פתרון בהקלטה)
להלן התפלגות השכר של עובדים בחברה מסוימת :

| שכר בש"ח | f(x) |
|-------------|------|
| 4000-6000 | 140 |
| 6000-10000 | 128 |
| 10000-15000 | 60 |
| 15000-20000 | 54 |
| 20000-40000 | 18 |

- א. מצאו את המאון ה-40.
ב. מהו אחוז העובדים שמשתכרים מתחת ל-5,000 ₪?

תרגילים:

1. להלן התפלגות השכר (באלפי שקלים) בחברה:

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

א. חשבו את המאון ה-60.

ב. מהו העשירון העליון?

ג. 20% מהמשכורות הגבוהות ביותר הן משכורות של הבכירים, מהי המשכורת המינימאלית לבכיר?

ד. מה אחוז האנשים שמשתכרים מתחת ל-7000 ₪?

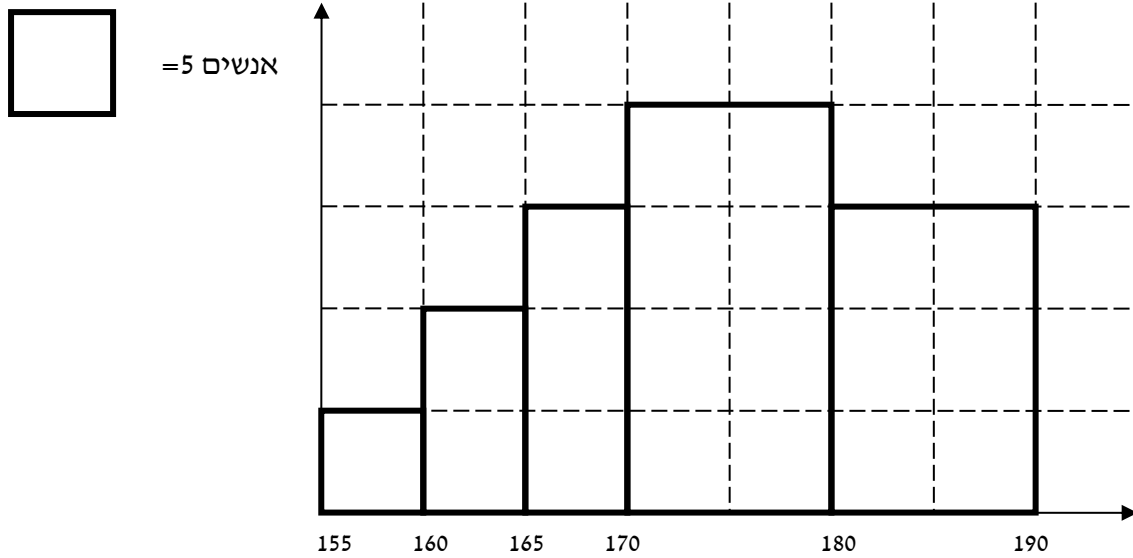
ה. איזה אחוז מהעובדים משתכרים מעל ל-25,000 ₪?

ו. איזה אחוז מהעובדים משתכרים בין 7000 ל-25,000 ₪?

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

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

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



- א. העשירון
התחתון.
ב. האחוזון ה-
.30
ג. הגובה ש-20%

- מהתצפית גדולות ממנו.
ד. את אחוז התצפיות מתחת לגובה 158 ס"מ.
ה. את אחוז התצפיות מעל לגובה 185 ס"מ.
ו. את אחוז התצפיות בין גובה 170 ס"מ ל-185 ס"מ.

פתרונות:**שאלה 1 :**

א. 13.23

ב. 22

ג. 17.2

ד. 8.82%

ה. 7.36%

ו. 83.82%

שאלה 3 :

א. 162.5

ב. 170

ג. 183.33

ד. 3%

ה. 15%

ו. 55%

פרק 10 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

רקע:

מצב שבו מבצעים שינוי מסוג הוספה של קבוע (או החסרה) והכפלה של קבוע (או חילוק) לכל

$$y = a \cdot x + b$$

וכך יושפעו המדדים השונים :

מדדי המרכז:

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

$$Mo_y = a \cdot Mo + b$$

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

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

מדדי הפיזור:

$$R_y = |a| R_x$$

$$s_y = |a| s_x$$

$$s_y^2 = a^2 s_x^2$$

מדדי המיקום היחסי:

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

$$Z_Y = \frac{a}{|a|} Z_X$$

שלבי העבודה:

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

דוגמה: (פתרון בהקלטה)

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

תרגילים:

1. עבור סדרת נתונים התקבל:

$$\bar{X} = 80$$

$$S = 15$$

$$MO = 70$$

הוחלט להכפיל את כל התצפיות פי-4 ולהחסיר מהתוצאה 5. חשב את המדדים הללו לאחר השינוי.

2. בחברה מסוימת השכר הממוצע הוא 40 ₪ לשעה עם סטיית תקן של 5 ₪ לשעה. הוחלט להעלות את כל המשכורות ב-10%, אך זה לא סיפק את העובדים ולכן הם קיבלו לאחר מכן תוספת של 2 ₪ לשעה. מה הממוצע ומהי השונות של השכר לשעה לאחר כל השינויים.

3. במבחן הציון החציוני היה 73, טווח הציונים היה 40 נקודות. והעשירון העליון היה הציון 87. כיוון שהציונים בבחינה היו נמוכים, המורה החליט לתת פקטור של 4 נק' לכל התלמידים. חשבו את המדדים לאחר הפקטור.

4. דגמו מקו ייצור 50 קופסאות של גפרורים. בדקו בכל קופסא בה יש 40 גפרורים את כמות הגפרורים הפגומים. קבלו שבממוצע יש 3 גפרורים פגומים בקופסא. עם סטיית תקן של 1.5 גפרורים. מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקנים בקופסא?

5. חברת בזק הציעה את החבילה הבאה:
שלושים שקלים דמי מנוי חודשיים קבועים. ובנוסף 10 אגורות לכל דקה של שיחה יוצאת, אדם בדק במשך שנה את דקות השיחות היוצאות שלו, וקיבל שבממוצע בחודש יש לו 600 דקות שיחות יוצאות עם שונות 2500 דקות רבועות, כמו כן בחודש ינואר ציון התקן היה 2. חשבו את המדדים הללו עבור חשבון הטלפון החודשי של אותו אדם בשקלים אם היה משתמש בחבילה המוצעת לו על ידי בזק.

6. הוכח שאם כל התצפיות בהתפלגות עברו טרנספורמציה לינארית:

$$Y_i = a \cdot X_i + b$$

אזי הממוצע והשונות של כלל התצפיות לאחר הטרנספורמציה יהיו בהתאמה:

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

$$s_y^2 = a^2 s_x^2$$

פתרונות :**שאלה 1 :**

הממוצע: 315
סטיית התקן: 60
השונות: 30.25
השכיח: 275

שאלה 3 :

טווח: 40
חציון: 77
עשירון עליון: 91

שאלה 4 :

ממוצע: 37
סטיית תקן: 1.5

שאלה 5 :

ממוצע: 90
שונות: 25
ציון תקן: 2

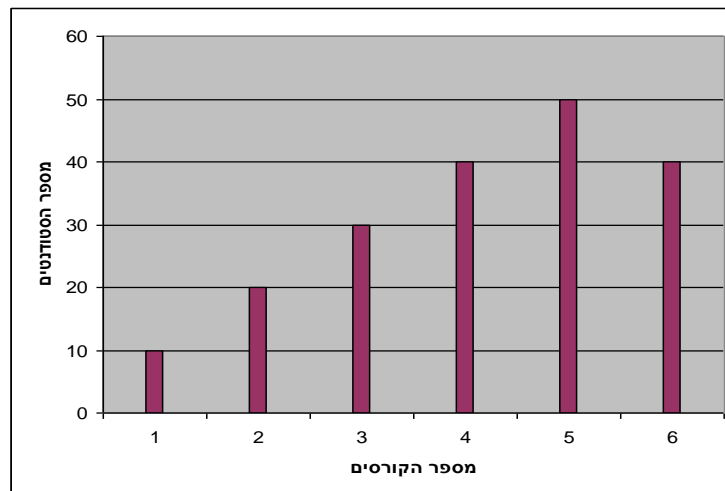
פרק 11 - סטטיסטיקה תיאורית - שאלות מסכמות

1. בדקו עבור 5 תלמידים את המשקל שלהם :

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

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

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



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

3. להלן התפלגות הציונים בבחינה בלשון שנעשתה עבור תלמידי כיתות ד'. השתתפו במחקר 150 תלמידים.

$$\text{ממוצע הציונים שהתקבל: } \bar{X} = 7\frac{1}{15}$$

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

- א. השלם את השכיחויות החסרות בטבלה.
 ב. חשב את הציון החציוני, השכיח.
 ג. חשב שונות וסטיית תקן להתפלגות הציונים.

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

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

- א. מה אחוז האנשים עם רמת שביעות רצון נמוכה?
- ב. מה המשתנה הנחקר ומאיזה סוג הוא?
- ג. מהי הדרך הגרפית המתאימה ביותר לתיאור הנתונים?
- i. היסטוגרמה.
- ii. דיאגרמת מקלות.
- iii. דיאגרמת עוגה.
- ד. חשבו את המדדים הבאים:
1. טווח
 2. שכיח
 3. חציון

5. להלן התפלגות מספר שעות העבודה לשבוע של העובדים בחברת "סטאר".
בחברה 200 עובדים.

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

א. השלם את הטבלה.

ב. חשב את החציון, השכיח, והממוצע של התפלגות מס' שעות העבודה בחברה.

ג. מהי סטיית התקן של מס' שעות העבודה?

ד. מה העשירון העליון של ההתפלגות?

ה. איזה אחוז מהעובדים עובדים מעל 45 שעות בשבוע?

ו. מה ציון התקן של רינה שעובדת 30 שעות בשבוע?

ז. כיצד ישתנה החציון, הממוצע וסטיית התקן אם מספר שעות העבודה המינימאלי אינו 10 אלא 15? הסבר.

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

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

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

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

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

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

8. התפלגות ציוני מבחן אינטיליגנציה היא סימטרית .

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

נתון שהעשירון העליון הוא 130 והרבעון התחתון הוא 90.

נתון שלמבחן נגשו 500 מועמדים.

א. השלימו את הטבלה.

ב. מהו הממוצע והחציון של ההתפלגות?

ג. מהו הציון ש 40% מהתלמידים קיבלו מעליו? באיזה אחוזון מדובר?

ד. אם יוחלט להעלות את כל הציונים ב-10 נקודות . כיצד הדבר ישפיע על הממוצע וסטיית התקן של הציונים?

9. להלן מספר טענות , עבור כל טענה ציין אם היא נכונה או לא נכונה ונמקד .

א. בסדרה שבה כל התצפיות שוות זו לזו השונות הינה 0.

ב. ציון התקן של החציון תמיד יהיה 0.

ג. ציון התקן של האחוזון ה-70 בהתפלגות אסימטרית ימנית (חיובית) תמיד יהיה חיובי.

ד. אם נוסיף תצפיות לסדרה של תצפיות, הדבר בהכרח יגדיל את הממוצע של הסדרה.

ה. בסדרה החציון הינו 80. הוספו שתי תצפיות אחת 79 ואחת 100 לכן החציון יגדל.

ו. אם נוסיף את הערך 4 לכל התצפיות אז סטיית התקן לא תשתנה.

ז. אם נחלק את כל התצפיות בהתפלגות ב-2 אז השונות תקטן פי 2.

ח. אם נגדיל את ממוצע המשכורות של עובדים בחברה אז גם השונות תגדל.

פתרונות:**שאלה 1:**

- א. המשתנה הנחקר כאן הוא משקל תלמיד בק"ג והוא משתנה כמותי רציף.
ב.

$$\bar{X} = 52$$

$$Md = X_{\frac{n+1}{2}} = X_3 = 58$$

השכיח הוא 58

$$R = 28 \quad \text{ג.}$$

$$s = 10.12$$

- ד. הוא חריג יותר בגובה כי שם ציון התקן בערך מוחלט יותר גבוה.
ה. הממוצע לא ישתנה אך סטיית התקן תקטן.

שאלה 2:

- א. מספר הקורסים. בדיד.
ב. התפלגות אסימטרית שמאלית
ד. השכיח: 5
הטווח: 5

שאלה 3:

- א. 20 תלמידים קיבלו ציון 6 ו-40
תלמידים קיבלו ציון 8.
החציון: 7
השכיח: 8
ג. השונות: 2.533
סטיית התקן: 1.592

שאלה 4:

- א. 20%
ב. שביעות רצון (סדר)
ג. 2
ד. טווח: 4 שכיח: 2 חציון: 2.5
ה. חציון: 4

שאלה 5:

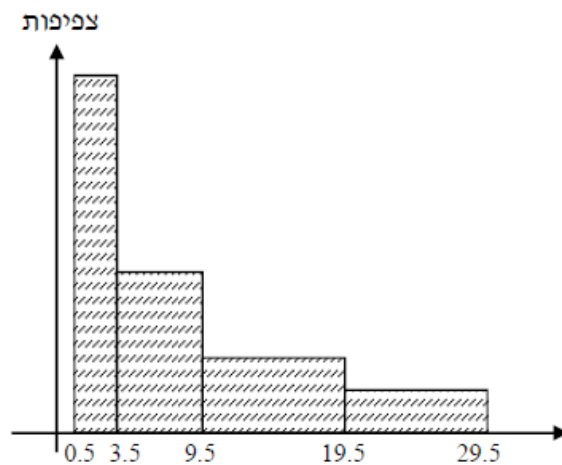
- ב. החציון : 35
 השכיח : 35
 הממוצע : 35
 ג. סטיית תקן : 12.65
 ד. 53.333
 ה. 25%
 ו. -0.395
 ז. חציון לא ישתנה, ממוצע יגדל סטיית התקן תקטן.

שאלה 6:

- א. 38%
 ב. 40%
 ג. 150
 ד. החציון : 100

שאלה 7:

א.



ב. ההתפלגות היא א-סימטרית ימנית .

ג. שכיח : 2 חציון : 6.83

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

שאלה 8:

א.

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

ב. 100

ג. 104

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

שאלה 9:

א. נכון

ב. לא נכון

ג. לא נכון

ד. לא נכון

ה. לא נכון

ו. נכון

ז. לא נכון

ח. לא נכון

פרק 12 - 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 :

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

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.
- הרכב את כל המילים האפשריות.
 - רשום את המקרים למאורע:
 - A במילה נמצאת האות E.
 - B במילה האותיות שונות.
 - ג. רשום את המקרים למאורע \bar{A} .
2. מטילים זוג קוביות.
- רשום את מרחב המדגם של הניסוי. האם המרחב מדגם הוא אחיד?
 - 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 בישוב מסוים:

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

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

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

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

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 13 - המשתנה המקרי הבדיד - פונקציית ההסתברות

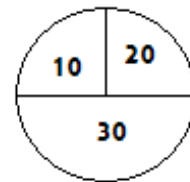
רקע:

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 מהיישוב, נגדיר את X להיות מספר המכוניות של המשפחה שנבחרה.
- בנו את פונקציית ההסתברות של X .
2. מהאותיות A, B, C יוצרים קוד דו תווי.
- א. כמה קודים ניתן ליצור?
 - ב. רשמו את כל הקודים האפשריים
 - ג. נגדיר את X להיות מספר הפעמים שהאות B מופיעה בקוד, בנו את פונקציית ההסתברות של X .

פרק 14 - המשתנה המקרי הבדיד - תוחלת, שונות וסטיית תקן

רקע:

$$E(X) = \sum_i x_i P(x_i) = \mu$$

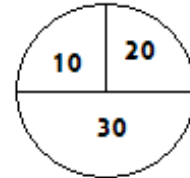
$$V(X) = \sum_i (x_i - \mu)^2 P(x_i) = \sum_i x_i^2 P(x_i) - \mu^2 = \sigma^2$$

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 :

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

מהי התוחלת, השונות וסטית התקן של X ?

2. בישוב מסוים שני סניפי בנק, בנק פועלים ובנק לאומי. מתוך האוכלוסייה הבוגרת בישוב ל-50% חשבון בנק בסניף הפועלים 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 $E(X)$

פתרונות:

שאלה 1

תוחלת : 2 שונות : 796

פרק 15 - 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. נוהה שמדובר בטרנספורמציה לינארית (שינוי קבוע לכל התצפיות).
2. נרשום את כלל הטרנספורמציה לפי נתוני השאלה.
3. נפשט את הכלל ונוהה 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 והשונוות 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?

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

פרק 16 - 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 :

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

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

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

$$E(X) = np \quad \text{תוחלת}$$

$$V(X) = npq \quad \text{שונות}$$

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

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

דוגמה : (פתרון בהקלטה)

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

פתרונות :**שאלה 7 :**

- א. 0.401
 ב. תוחלת : 8.025
 סטיית תקן : 2.193

שאלה 2 :

- ב. 0.2335
 ג. 0.1493
 ד. תוחלת : 7
 סטיית תקן : 1.449

שאלה 8 :

7.5

שאלה 3 :

- א. 0.5314
 ב. 0.0984
 ג. 0.1143
 ד. תוחלת : -18
 סטיית תקן : 14.697

שאלה 4 :

- א. 0.1789
 ב. 2

פרק 17 - 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. 30% מהמכונות בארץ הן בצבע לבן. בכל יום נכנסות לחניון 10 מכונות אקראיות.
 - א. מה ההסתברות שביום מסוים בדיוק מחצית מהמכונות בחניון יהיו לבנות?
 - ב. מה תוחלת מספר הימים שיעברו מהיום עד שלראשונה מחצית מהמכונות בחניון יהיו לבנות?

3. 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?

פתרונות :**שאלה 2:**

א. 0.1029

ב. 9.72

שאלה 3:

א. 0.06

ב. 0.7176

ג. 0.0729

ד. 9.487 משחקים

פרק 18 - התפלגויות רציפות מיוחדות - התפלגות אחידה

רקע:

זו התפלגות שפונקציית הצפיפות שלה קבועה בין a לבין b.

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

פונקציית הצפיפות:

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

$$a \leq x \leq b$$

פונקציית ההתפלגות המצטברת:

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

התוחלת:

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

השונות:

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

דוגמה: (הפתרון בהקלטה)

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

א. מה הסיכוי ש-X קטן מ-25?

ב. מה התוחלת והשונות של X?

תרגילים:

1. משך (בדקות) הפסקה בשיעור, X , מתפלג $U(13, 16)$.
 - א. מהי התוחלת ומהי סטית התקן של משך ההפסקה?
 - ב. מהי ההסתברות שהפסקה תמשך יותר מ-15 דקות?
 - ג. מהי ההסתברות שמשך ההפסקה יסטה מהתוחלת בפחות מדקה?

2. רכבת מגיעה לתחנה בשעות היום כל עשר דקות. אדם הגיע לתחנה בזמן אקראי.
 - א. הסבר כיצד מתפלג זמן ההמתנה לרכבת?
 - ב. אם זמן ההמתנה לרכבת ארך יותר מ-5 דקות, מהי ההסתברות שבסך הכל האדם ימתין לרכבת פחות מ-8 דקות?
 - ג. מה תוחלת מספר הימים שיעברו עד הפעם הראשונה שהאדם ימתין לרכבת יותר מ-9 דקות?

3. מכונה אוטומטית ממלאת גביעי גלידה. משקל הגלידה לגביע מתפלג אחיד בין 100-110 גרם (המשקל הוא של גלידה ללא הגביע).
 - א. מה ההסתברות שמשקל הגלידה בגביע יהיה מעל 108 גרם?
 - ב. נתון שהגלידה בגביע עם משקל נמוך מ-107 גרם. מה ההסתברות שמשקל הגלידה יהיה מעל 105 גרם?
 - ג. מה העשירון העליון של משקל הגלידה בגביע?

פתרונות:שאלה 2:א. $X \sim U(0,10)$

ב. 0.6

ג. 10

שאלה 1:

א. תוחלת: 14.5

שונות: 0.866

ב. $1/3$ ג. $2/3$ שאלה 3:

א. 0.2

ב. $\frac{2}{7}$

ג. 109

פרק 19 - 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 היולדו 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:

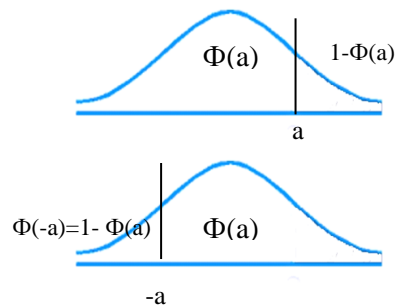
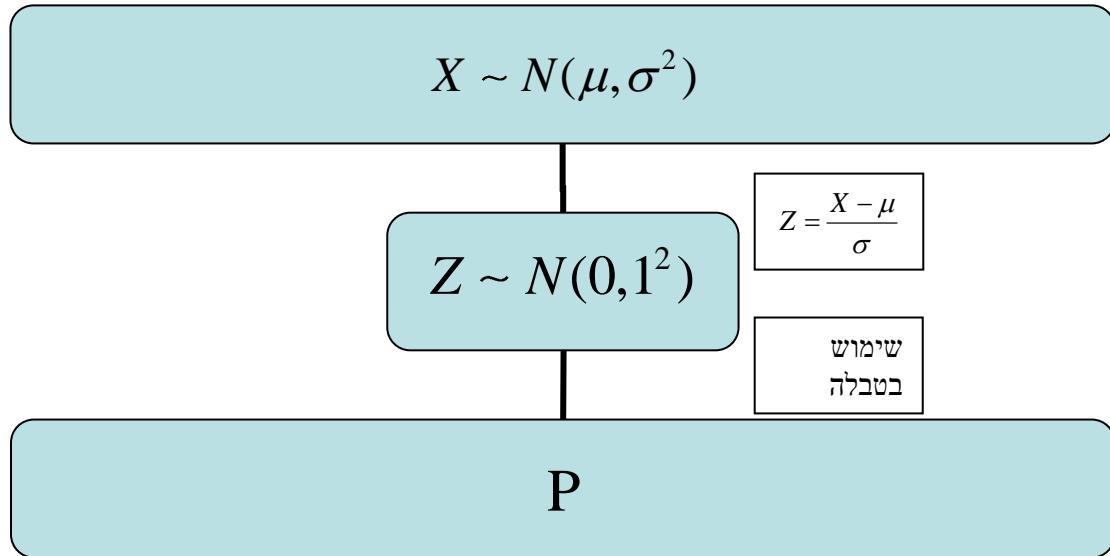
$$Z = \frac{X - \mu}{\sigma}$$

אחרי תקנון מקבלים ערך הנקרא ציון תקן.

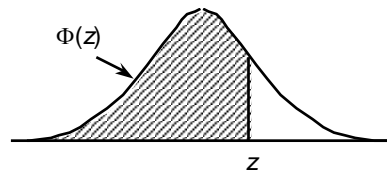
ציון התקן משמעו בכמה סטיות תקן הערך סוטה מהממוצע.

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 95.
- ג. מצאו את העשירון התחתון.

5. ציוני מבחן אינטליגנציה מתפלג נורמלית עם ממוצע 100 ושונויות 225 .
- מה העשירון העליון של הציונים במבחן האינטליגנציה?
 - מה העשירון התחתון של ההתפלגות?
 - מהו הציון ש- 20% מהנבחנים מקבלים מעליו?
 - מהו האחוזון ה- 20?
 - מהו הציון ש- 5% מהנבחנים מקבלים מתחתיו?
6. נפח משקה בבקבוק מתפלג נורמלית עם סטיית תקן של 20 מ"ל, נתון ש-33% מהבקבוקים הם עם נפח שעולה על 508.8 מ"ל.
- מה ממוצע נפח משקה בבקבוק ?
 - 5% מהבקבוקים המיוצרים עם הנפח הגבוה ביותר נשלחים לבדיקה, החל מאיזה נפח שולחים בבקבוק לבדיקה?
 - 1% מהבקבוקים עם הנפח הקטן ביותר נתרמים לצדקה, מהו הנפח המקסימלי לצדקה?
7. אורך חיים של מכשיר מתפלג נורמלית . ידוע שמחצית מהמכשירים חיים פחות מ- 500 שעות, כמו כן ידוע ש- 67% מהמכשירים חיים פחות מ- 544 שעות.
- מהו ממוצע אורך חיי מכשיר?
 - מהי סטיית בתקן של אורך חיי מכשיר?
 - מה הסיכוי שמכשיר אקראי יחיה פחות מ- 460 שעות?
 - מהו המאון העליון של אורח חיי מכשיר?
 - 1% מהמכשירים בעלי אורך החיים הקצר ביותר נשלח למעבדה לבדיקה מעמיקה. מהו אורך החיים המקסימלי לשליחת מכשיר למעבדה?

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



א. לאיזו התפלגות הממוצע הגבוה ביותר?
 ב. במה מבין המדדים הבאים התפלגות 1 ו 2 זהות?

- א. בעשירון העליון.
 ב. בממוצע.
 ג. בשונות.

ג. לאיזו התפלגות סטיית התקן הקטנה ביותר?

- א. 1
 ב. 2
 ג. 3
 ד. אין לדעת.

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

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

פתרונות :

| <u>שאלה 3</u> | <u>שאלה 1</u> |
|---------------|---------------|
| א. 26.43% | א. 89.25% |
| ב. 89.44% | ב. 2.28% |
| ג. 39.44% | ג. 0 |
| ד. 0.383 | ד. 50% |
| ה. 100% | |

| <u>שאלה 7</u> | <u>שאלה 5</u> |
|---------------|---------------|
| א. 500 | א. 119.2 |
| ב. 100 | ב. 80.8 |
| ג. 0.3446 | ג. 112.6 |
| ד. 733 | ד. 87.4 |
| ה. 267 | |

| <u>שאלה 9</u> | <u>שאלה 8</u> |
|---------------|---------------|
| א. 0.1587 | א. 3 |
| ב. 0.0228 | ב. בממוצע. |
| ג. 0.8563 | ג. 1 |
| ד. 0.3975 | |

פרק 20 - התפלגות הדגימה

ממוצע המדגם ומשפט הגבול המרכזי

רקע:

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

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

גדלים המתארים התפלגות כלשהי או אוכלוסייה כלשהי נקראים פרמטרים. להלן רשימה של פרמטרים החשובים לפרק זה:

ממוצע האוכלוסייה נסמן ב μ (נקרא גם תוחלת).

שונות אוכלוסייה נסמן ב- σ^2 .

סטיית תקן של אוכלוסייה: σ .

א. תכונות התפלגות

ממוצע כל ממוצעי המדגם האפשריים שווה לממוצע האוכלוסייה:

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

שונות כל ממוצעי המדגם האפשריים שווה לשונות האוכלוסייה מחולק ב- n . תכונה זו נכונה רק במדגם מקרי:

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

יש יחס הפוך בין גודל המדגם לבין שונות ממוצעי המדגם.

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

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

דוגמה: (פתרון בהקלטה)

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

א. מי אוכלוסיית המחקר? מהו המשתנה הנחקר?

ב. מהם הפרמטרים של האוכלוסייה?

ג. מה התוחלת ומהי סטיית 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 מקלטים | מספר המשפחות |
|-------------|---------------------------------------|
| 0 | 500 |
| 1 | 2500 |
| 2 | 3500 |
| 3 | 3000 |
| 4 | 500 |
| | סך הכול $N = 10000$ |

נגדיר את x להיות מספר המקלטים של משפחה אקראית.

א. בנו את 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.

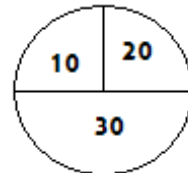
א. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה בדיוק 755 סמ"ק?

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 נורמאלית עם תוחלת 80 וסטיית תקן 4 .
 א. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?
 ב. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה שגודל המדגם הוא 16?
 ג. הסבר את ההבדל בתשובות של שני הסעיפים.
9. בקזינו ישנה רולטה. על הרולטה רשומים המס' הבאים כמוראה בשרטוט :



- אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.
 א. בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.
 ב. מה התוחלת ומה השונות של סכום הזכייה?
 ג. אם האדם ישחק את המשחק 5 פעמים מה התוחלת ומה השונות של ממוצע סכום הזכייה בחמשת המשחקים?
 ד. אם האדם משחק את המשחק 50 פעם מה ההסתברות שבסה"כ יזכה ב-1050 ₪ ומעלה?
10. לפי הערכות הלשכה המרכזית לסטטיסטיקה השכר הממוצע במשק הוא 8000 ₪ עם סטיית תקן של 3000 ₪. מה ההסתברות שבמדגם מקרי של 100 עובדים השכר הממוצע יהיה יותר מ-8500 ₪?
11. 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 בממוצע המדגם \bar{X} :

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

א. 0

ב. 0.5

ג. 1

ד. לא ניתן לדעת.

15. נתון ש X מתפלג כלשהו עם תוחלת μ ושונות σ^2 .

החליטו לבצע מדגם בגודל 200 מתוך ההפלגות הנתונה לפי משפט הגבול המרכזי מתקיים ש :
(בחר בתשובה הנכונה)

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

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

ג. $\bar{X} \sim N(\mu, \sigma^2)$

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

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

(בחר בתשובה הנכונה)

א. μ ו- \bar{X} יהיו משתנים מקריים.

ב. μ יהיה משתנה מקרי ו \bar{X} קבוע.

ג. \bar{X} יהיה משתנה מקרי ו μ קבוע.

ד. μ ו \bar{X} יהיו קבועים.

פתרונות:**שאלה 2**

א.

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

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

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

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

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

התפלגות סכום תצפיות המדגם ומשפט הגבול המרכזי

רקע:

$$T = \sum_{i=1}^n X_i$$

כעת נדון בסטיסטי המבטא את סכום התצפיות במדגם

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:

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

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

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

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)

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

דוגמה: (פתרון בהקלטה)

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

- מה התוחלת וסטיית התקן של סך המשכורות שיופקדו לסניף הבנק על ידי העובדים הללו?
- מה ההסתברות שלסניף יופקד פחות מ-780 אלף ₪ ע"י אותם עובדים? (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 מזוג 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?

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.

התפלגות מספר ההצלחות במדגם - הקרוב הנורמלי להתפלגות הבינומית

רקע:

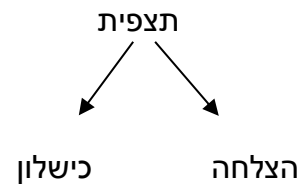
תזכורת על התפלגות בינומית

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

מספר ההצלחות במדגם נסמן ב- Y .

מחלקים כל תצפית במדגם להצלחה או כישלון.

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



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

מבצעים מדגם אקראי בגודל n .

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

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

תוחלת: $E(y) = np$

שוונות: $V(y) = npq$

קירוב נורמלי עבור התפלגות בינומית

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

$$1. n \cdot p \geq 5$$

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

$$y \rightsquigarrow N(np, npq)$$

$$Z_y = \frac{y - np}{\sqrt{npq}} \quad \text{אז :}$$

תיקון רציפות:

כאשר משתמשים בקירוב הנורמלי להתפלגות הבינומית יש לבצע תיקון רציפות .
הסיבה שעוברים כאן מהתפלגות בדידה להתפלגות נורמלית שהיא התפלגות רציפה.
על פי הכללים הבאים :

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

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

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

הערות:

- התנאים למעבר מבינומי לנורמלי הם נזילים, כלומר משתנים ממרצה אחד לשני. התנאי שהצגתי כאן הוא הפופולרי ביותר:

$$1. n \cdot p \geq 5$$

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

- ישנם מרצים שנותנים את התנאי המחמיר הבא:

$$1. n \cdot p \geq 10$$

$$2. n \cdot (1 - p) \geq 10$$

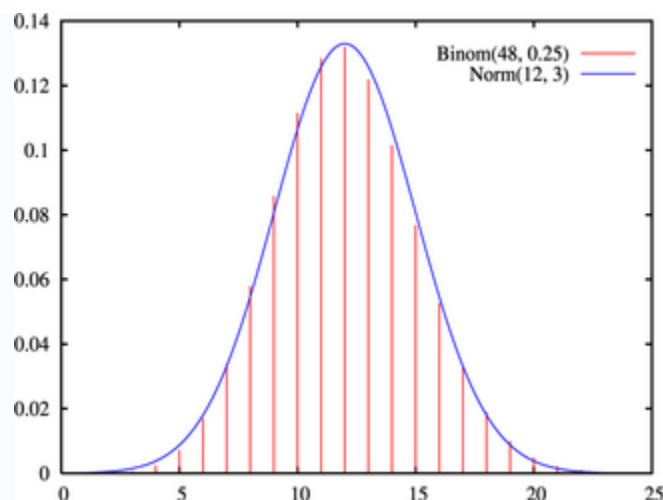
- וישנם מרצים שפשוט התנאי שהם נותנים הוא: $(n \geq 30)$.
- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור מהתפלגות בינומית לנורמלית.
- הערה נוספת היא לגבי תיקון רציפות. ישנם מרצים שלא מחייבים לבצע תיקון רציפות שהמדגמים גדולים (בדרך כלל מעל 100 תצפיות) אני בפתרונות שאציג תמיד אבצע תיקון רציפות במעבר מבינומי לנורמלי כיוון שכך הפתרון יהיה יותר מדויק (בכל מקרה שהמדגמים גדולים העניין זניח).

דוגמה: (הפתרון בהקלטה)

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

א. מה הסיכוי שבדיוק 14 מתוכם יהיו זקוקים למשקפיים?

ב. מה הסיכוי שלכל היותר 13 מתוכם זקוקים למשקפיים?



תרגילים:

1. נתון ש-20% מאוכלוסייה מסוימת אקדמאית. נבחרו באקראי 10 אנשים באותה אוכלוסייה.
- א. מה ההסתברות ששלושה מהם אקדמאים?
- ב. מה ההסתברות שלכל היותר אחד מהם אקדמאי?
- ג. מה התוחלת ומהי סטיית התקן של מספר האקדמאים במדגם?
2. במפעל 10% מהמוצרים פגומים. נלקחו 100 מוצרים באקראי מקו הייצור.
- א. מה ההסתברות שנדגמו לפחות 6 מוצרים פגומים?
- ב. מה ההסתברות שמספר המוצרים הפגומים יהיה לכל היותר 11 במדגם?
3. ציוני פסיכומטרי בקרב הנרשמים למוסד מסוים מתפלגים נורמאלית עם ממוצע 500 וסטיית תקן 100. למוסד מסוים הוחלט לקבל אך ורק סטודנטים שקיבלנו מעל 600 בפסיכומטרי. 100 סטודנטים אקראיים נרשמו למוסד. מה ההסתברות שלפחות 20 יתקבלו?
4. מטילים מטבע 50 פעמים.
- א. מה ההסתברות לקבל לכל היותר 30 עצים?
- ב. מה ההסתברות לקבל 28 עצים לפי התפלגות הבינומית ולפי הקירוב הנורמאלי?
5. במטוס מקום ל-400 נוסעים. נרשמו לטיסה 430 אנשים (overbooking). מנתונים סטטיסטיים ידוע שהסיכוי שאדם שנרשם לטיסה אכן יגיע הוא 0.9.
- א. מה ההסתברות שלא יהיו מקומות ישיבה לכל האנשים שהגיעו לטיסה?
- ב. מה צריך להיות גודל המטוס כדי שבסיכוי שלפחות 95% המטוס יספיק לכמות הנרשמים?
6. מפעל לייצור ארטיקים טוען ש הסיכוי שארטיק שהוא מייצר יהיה פגום הוא 0.01. מוכר הזמין 1000 ארטיקים מהמפעל. מה ההסתברות שהמוכר יקבל לפחות 980 ארטיקים תקינים אם טענת המפעל מוצדקת?
7. מהמר מטיל קובייה הוגנת 100 פעמים. בכל הטלה, אם מתקבל תוצאה זוגית בקובייה המהמר זוכה בשקל. אחרת, המהמר משלם שקל. המהמר הטיל את הקובייה 100 פעמים מה הסיכוי שהרווח של המהמר יהיה לכל היותר 10?

פתרונות:**שאלה 1**

א. 0.201

ב. 0.3758

ג. התוחלת : 2, סטיית התקן : 1.2649

שאלה 2

א. 0.9332

ב. 0.6915

שאלה 3

0.1611

שאלה 4

א. 0.9406

שאלה 5

א. 0.015

שאלה 6

0.9996

שאלה 7

0.8643

התפלגות פרופורציית ההצלחות במדגם

רקע:

בפרק זה נדון על התפלגות הדגימה של פרופורציית המדגם.

Y - מספר ההצלחות במדגם (למשל, מספר המובטלים במדגם)

$$\hat{p} = \frac{y}{n} \text{ - פרופורציית ההצלחות במדגם (למשל, שיעור המובטלים במדגם)}$$

למשל,

$$n = 200$$

מספר המובטלים : $Y = 20$

$$\hat{p} = \frac{20}{200} = 0.1 \text{ פרופורציית המובטלים במדגם}$$

נסמן ב- p את שיעור ההצלחה באוכלוסייה וב- q את שיעור הכישלונות באוכלוסייה.

נבצע מדגם מקרי (הנחה שהתצפיות בלתי תלויות זו בזו) ונתבונן בהתפלגות של פרופורציית המדגם.

התוחלת, השונות וסטיית התקן של פרופורציית המדגם:

$$E(\hat{P}) = p$$

$$V(\hat{P}) = \frac{pq}{n}$$

$$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}}$$

משפט הגבול המרכזי עבור הפרופורציה המדגמית:

אם $np \geq 5$ & $nq \geq 5$ אזי $\hat{p} \rightsquigarrow N(p, \frac{pq}{n})$

$$Z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

הערות:

- התנאים לקרוב הנורמאלי הם נזילים, כלומר משתנים ממרצה אחד לשני. התנאי שהצגתי כאן הוא הפופולרי ביותר:

$$1. \quad n \cdot p \geq 5$$

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

- ישנם מרצים שנותנים את התנאי המחמיר הבא:

$$1. \quad n \cdot p \geq 10$$

$$2. \quad n \cdot (1 - p) \geq 10$$

- וישנם מרצים המשתמשים בתנאי: $(n \geq 30)$.

- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור לנורמלית.
 - כיוון שפרופורציה אינה חייבת להיות מספר שלם בהכרח לא נהוג לבצע כאן תיקון רציפות.
- דוגמה:** (פתרון בהקלטה)

לפי נתוני משרד החינוך בעיר ירושלים ל-60% מתלמידי התיכון זכאים לתעודת בגרות. נדגמו 200 תלמידי תיכון.

- א. מה ההסתברות שהשכיחות היחסית \hat{p} של הזכאים לבגרות במדגם תעלה על 60%?
- ב. מה ההסתברות שפרופורציית הזכאים לבגרות במדגם תעלה על 70%?

תרגילים:

1. במדינה מסוימת 10% מכלל האוכלוסייה הינם מובטלים. נדגמו באקראי 140 אנשים מהמדינה.
 - א. מה התוחלת ומהי השונות של פרופורציות המובטלים שנדגמו?
 - ב. מה ההסתברות שבמדגם לפחות 10% יהיו מובטלים?
 - ג. מה ההסתברות שלכל היותר 9% מהמדגם יהיו מובטלים?

2. נניח כי 30% מהאוכלוסייה תומכים בהצעת חוק מסוימת. אם נדגום מהאוכלוסייה 200 איש.

חשבו את ההסתברויות הבאות:

 - א. לפחות 35% יתמכו בהצעת החוק במדגם.
 - ב. לכל היותר 25% יתמכו בהצעת החוק במדגם.
 - ג. יותר מ – 27% יתמכו בהצעת החוק במדגם.

3. לפי נתוני משרד התקשורת 40% מהאוכלוסייה מחזיקים בטלפון נייד מסוג "סמארטפון". נדגמו 400 אנשים מהאוכלוסייה.
 - א. מה ההסתברות שבמדגם לכל היותר ל 40% יש סמארטפון?
 - ב. מה ההסתברות שבמדגם לרוב יש סמארטפון?
 - ג. מה ההסתברות שפרופורציית בעלי הסמארטפון במדגם תסטה מהפרופורציה באוכלוסייה בלא יותר מ-4%?
 - ד. כיצד התשובה לסעיף הקודם הייתה משתנה אם הינו מגדילים את גודל המדגם?

4. נתון כי 80% מבתי האב מחוברים לאינטרנט. נדגמו 400 בתי אב אקראיים.
 - א. מה ההסתברות שלפחות 340 מהם מחוברים לאינטרנט?
 - ב. מה ההסתברות שפרופורציית המחוברים לאינטרנט במדגם תסטה מהפרופורציה האמתית ביותר מ-4%?
 - ג. כמה בתי אב יש לדגום כדי שהסטייה בין הפרופורציה המדגמית לפרופורציה האמתית לא תעלה על 3% בהסתברות של 90%?
 - ד. מהו העשירון התחתון של התפלגות פרופורציית המדגם?

5. נתון שציוני פסיכומטרי מתפלגים נורמלית עם תוחלת 500 וסטיית תקן 100. ל"מועדון ה-700" נכללים נבחנים שמקבלים ציון מעל 700 בפסיכומטרי. מה הסיכוי שבמועד בו נבחנו 2000 נבחנים אקראיים יהיו לפחות 3% המשתייכים למועדון?

6. נתון ש $X \sim B(n, p)$ נגדיר את המשתנה הבא : $\hat{P} = \frac{X}{n}$.

א. הוכיחו ש :

$$E(\hat{P}) = p$$

$$V(\hat{P}) = \frac{p(1-p)}{n}$$

ב. מה p המביא את $V(\hat{P})$ להיות במקסימום?

פתרונות:**שאלה 1**

א. התוחלת: 0.1, השונות: 0.00064

ב. 0.5

ג. 0.3446

שאלה 2

א. 0.0618

ב. 0.0618

ג. 0.8238

שאלה 3

א. 0.5

ב. 0

ג. 0.8968

ד. גדלה

שאלה 4

א. 0.0062

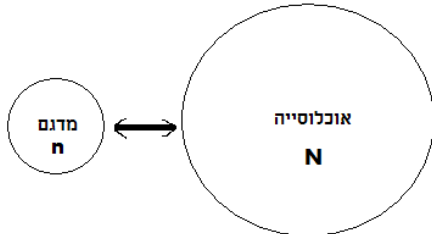
ב. 0.0456

פרק 21 - 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 סיכוי 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 מקלטים | מספר המשפחות |
|-------------|--------------------|
| 0 | 50 |
| 1 | 250 |
| 2 | 350 |
| 3 | 300 |
| 4 | 50 |
| | 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 22 - 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 פרופרציית 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 אומר חסר הטויה ל μ .

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

פרופורציה באוכלוסייה: p

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

$E(\hat{p}) = p$ לכן \hat{p} הינו אומר חסר הטויה 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 הטויה 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 וחישובו את הממוצע שלהם. מהי טעות התקן?
 א. סטיית התקן של האוכלוסייה.
 ב. סטיית התקן של ממוצע האוכלוסייה.
 ג. סטיית התקן של המדגם.
 ד. סטיית התקן של ממוצע המדגם.

6. משקל הממוצע של אוכלוסייה מסוימת הוא 75 ק"ג עם שונות של 25. אם יבחרו כל המדגמים האפשריים בגודל 10 מאוכלוסייה זו סטיית התקן של ממוצעי המדגמים תהייה:

א. 3

ב. 2.5

ג. 1.581

ד. אין מספיק נתונים לדעת.

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

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

פתרונות:

שאלה 3

א. 177.9

ב. 64.1

ג. 0.4

שאלה 4

א. 8.1

ב. 3.16

שאלה 5

התשובה היא ד.

שאלה 6

התשובה היא ג.

שאלה 7

התשובה היא ד.

שאלה 8

התשובה היא ד.

שאלה 9

התשובה היא ג.

פרק 23 - 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 האמתי הוא אפסי. מה שנהוג לעשות כדי לאמוד את ממוצע האוכלוסייה זה לבנות רווח סמך. נבנה מרווח בטחון שהסיכוי שהפרמטר μ ייכלל בתוכו הוא $1-\alpha$.

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

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

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 (שוונות 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 (שוונות 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.
 - 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.

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.

פתרונות :**שאלה 2**

$$4920.6 < \mu < 4979.4$$

שאלה 3

$$א. \quad 223.42 < \mu < 236.58$$

$$ב. \quad 222.16 < \mu < 237.84$$

שאלה 5

$$א. \quad 102$$

$$ב. \quad 3$$

$$ג. \quad 0.9544$$

שאלה 6

$$א. \quad 83.5 < \mu < 4.42$$

$$ב. \quad \text{יקטן פי 2}$$

$$ג. \quad \text{גדל}$$

שאלה 7

$$א. \quad 87$$

$$ב. \quad 5$$

$$ג. \quad 0.9544$$

שאלה 8

$$א. \quad 139$$

$$ב. \quad 21 < \mu < 25$$

שאלה 9

התשובה היא : ב

שאלה 10

התשובה היא : ג

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 נורמלית 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

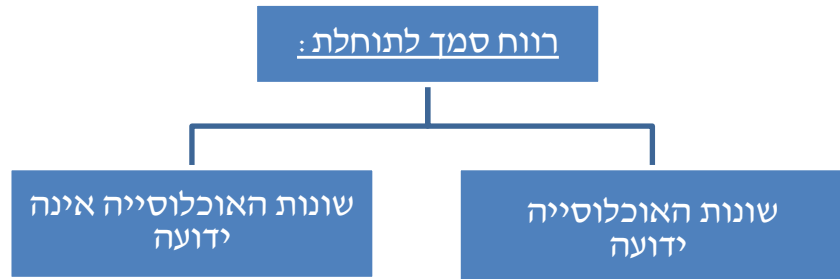
א. 139

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:



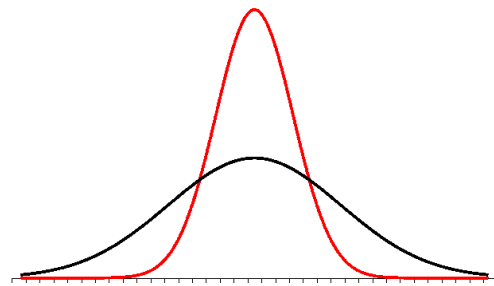
בפרק זה נעסוק במקרה ששונות האוכלוסייה (σ^2) אינה ידועה לנו. מקרה יותר פרקטי.

התנאי: $X \sim N$ או שהמדגם גדול

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

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

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 עלות ההתפלגות הופכת להיות יותר גבוהה וצרה. כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה : (פתרון בהקלטה)

הזמן שלוקח לפתור שאלה מסוימת בחשבון מתפלג אצל תלמידי כיתה 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: 89, 79, 84, 88, 84. הערה: לצורך פתרון הנח שקצב פעימות הלב מתפלג נורמאלי 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 נורמאלי. 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)

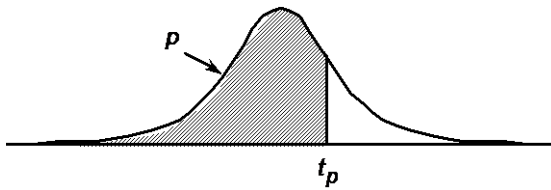
א. סטטיסטיקאי א

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 24 - רווח סמך לפרופורציה

רקע:

מטרה: לאמוד את P – פרופורציה באוכלוסייה.

האומד הנקודתי: $\hat{p} = \frac{y}{n}$ (Y- מספר ההצלחות שבמדגם)

רווח הסמך ל P: $\hat{p} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

התנאי לבנות את רווח הסמך הינו מדגם של לפחות 30 תצפיות (לעיתים נותנים תנאי של מספר הצלחות ומספר כשלונות לפחות 5 או לפחות 10)

האומד לטעות התקן: $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

מתקיים ש: $\hat{P} = \frac{A+B}{2}$ $L = 2\varepsilon$

דוגמה: (פתרון בהקלטה)

1. במטרה לאמוד את אחוז המובטלים במשק נדגמו 200 אזרחים. מתוכם התקבל ש 24 היו מובטלים.

א. בנו רווח סמך לאחוז המובטלים באוכלוסייה ברמת סמך של 95%.

ב. מהו האומד לטעות התקן?

פתרון:

א. $7.5\% < p < 16.5\%$

ב. 2.29%

תרגילים:

1. נדגמו 200 דירות בעיר חיפה. 48 מתוכן נמצאו כבעלות ממ"ד.
 - א. בנו רווח סמך ברמת סמך של 95% לאחוז הדירות בחיפה עם ממ"ד.
 - ב. על סמך סעיף א' מה ניתן לומר על שגיאת האמידה המקסימאלית?
 - ג. בהנחה ובחיפה 80 אלף דירות, בנו רווח סמך ברמת סמך של 95% למספר הדירות בחיפה עם ממ"ד בפועל.

2. במדגם של 300 אנשי היי-טק התקבל ש-180 מהם אקדמאים.
 - א. בנו רווח סמך לפרופורציית אקדמאים ברמת סמך של 95% (בקרב אנשי היי-טק).
 - ב. כיצד רווח הסמך של סעיף א' היה משתנה אם היינו מקטינים את רמת הסמך?
 - ג. כיצד רווח הסמך היה משתנה אם הינו מגדילים את גודל המדגם?

3. במדגם של 400 נהגים התקבל רווח סמך לפרופורציית הנהגים החדשים: $0.08 < p < 0.18$
 - א. כמה נהגים במדגם היו נהגים חדשים?
 - ב. מהי רמת הסמך של רווח הסמך שנבנה?

4. במסגרת מערכת הבחירות בארה"ב נשאלו 840 אנשים עבור איזה מועמד יצביעו.
 - א. 510 אנשים ענו כי יצביעו בעד ברק אובמה. בסקר פורסם שתתכן סטייה של $\pm 3\%$ מתוצאות האמת. באיזו רמת ביטחון הסקר השתמש?

5. במדגם של 300 נשים בגילאי 35-40 נמצא ש-140 היו נשואות, 80 היו גרושות, 60 רווקות והיתר אלמנות.
 - א. מצאו רווח סמך ברמה של 90% לאחוז הגרושות באוכלוסייה הנחקרת.
 - ב. מצאו רווח סמך ברמה של 99% לסיכוי שבאוכלוסייה הנחקרת תמצא אישה לא נשואה?

6. ביצעו מדגם באוכלוסייה. שיעור ההצלחות במדגם היה 10% ורווח הסמך ניבנה ברמת סמך של 95%. אורכו הינו 0.3156.
 - א. מהו גודל המדגם שנלקח?

פתרונות:**שאלה 3**

א. 52

ב. 0.997

שאלה 5א. $22.5\% < p < 30.9\%$ ב. $45.91\% < p < 60.72\%$ **שאלה 6**

200

קביעת גודל מדגם באמידת פרופורציה

רקע:

בפרק זה נדון איך קובעים גודל מדגם שבאים לאמוד פרופורציה באוכלוסייה מסוימת:

החוקר קובע מראש את רמת הסמך הרצויה: $1 - \alpha$.

החוקר קובע מראש את הטעות הסטטיסטית המרבית שבה הוא מעוניין: ε (או את אורך רווח הסמך).

$L = 2\varepsilon$ - אורך רווח הסמך.

ε - טעות אמידה מרבית: המרחק המקסימאלי (הסטייה) בין הפרמטר (p) לאומד (\hat{p}).

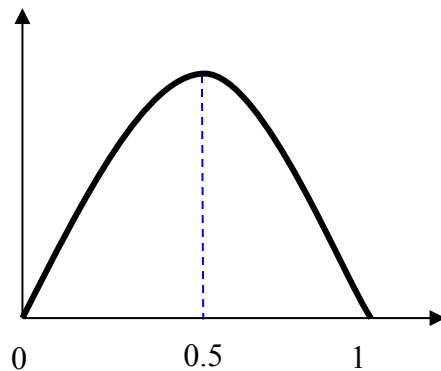
$$\varepsilon = z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

ויתעניין לדעת מהו גודל המדגם הרצוי לשם כך.

$$n \geq \left(\frac{2 \cdot z_{1-\alpha/2} \cdot \sqrt{\hat{p}(1-\hat{p})}}{L} \right)^2 \quad \text{נקבל ש:}$$

הבעיה שאין לנו יודעים את \hat{p} .

נתבונן בביטוי $\hat{p}(1-\hat{p})$:



כיוון שאין לנו ידע מוקדם על \hat{p} נציב את המקרה השמרני ביותר שממקסם את הביטוי עבור

$$\hat{p} = 0.5$$

$$n \geq \left(\frac{2 \cdot z_{1-\alpha/2} \cdot \sqrt{0.5 \cdot 0.5}}{L} \right)^2 \Rightarrow n \geq \left(\frac{z_{1-\alpha/2}}{L} \right)^2$$

אך אם תהיה לנו אינפורמציה מוקדמת על הפרופורציה נציב את הערך הקרוב ביותר ל-0.5 האפשרי.

דוגמה: (פתרון בהקלטה)

מעוניינים לאמוד את שיעור האבטלה במשק. האמידה צריכה להתבצע ברמת סמך של 90% ועם שגיאת אמידה שלא תעלה על 4%.
 א. מהו גודל המדגם המינימאלי שיש לקחת?
 ב. חזור לסעיף א' אם ידוע שהאבטלה לא אמורה לעלות על 20%.

פתרון:

א. 423

ב. 271

תרגילים:

לפתרון מלא בסרטון וידאו היכנסו ל- www.GooL.co.il

© כתב ופתר - ברק קנדל

1. הממשלה אומדת מדי חודש את אחוז התמיכה בה. מהו גודל המדגם אשר יש לקחת אם דורשים שהאומדן לא יסטה מהאחוז האמתי באוכלוסייה ביותר מ-3%, וזאת בביטחון של 95%?
2. משרד התקשורת מעוניין לדעת מה שיעור בתי האב עם אינטרנט.
 - א. כמה בתי אב יש לדגום אם מעוניינים שבביטחון של 90% אורך רווח הסמך לא יעלה על 8%?
 - ב. חזרו על סעיף א. אם ידעו שלפני חמש שנים ל-80% מבתי האב היה אינטרנט וכיום יש להניח שיש ליותר אינטרנט.
3. ערוץ טלוויזיה מעוניין לאמוד את הרייטינג של הערוץ בפריים טיים. המטרה שבביטחון של 95% הסטייה המרבית בין האומדן לרייטינג האמתי לא תעלה על 4%.
 - א. כמה מכשירי PEOPLE METER יש להתקין לצורך האמידה?
 - ב. לפי הערכה מוקדמת הרייטינג של הערוץ לא יכול לעלות על 20%. בהנחה ומכשיר כזה עולה 500 ₪ ליחידה מה החיסכון הכספי מאינפורמציה זאת?
4. השאלות הבאות מתייחסות לסעיף 4 :
 - א. כמה אזרחים יש לדגום כדי לאמוד את אחוז התמיכה בממשלה עם אורך רווח הסמך שלא עולה על 9% ברמת סמך של 90%?
 - ב. בהנחה ובוצע מדגם שאת גודלו חיבתם בסעיף א והתקבל שאחוז התמיכה בממשלה במדגם הנו 42%. בנו רווח סמך לאחוז התמיכה בממשלה ברמת סמך של 95%.
 - ג. על סמך סעיף ב'. האם תקבל את הטענה שמיעוט האוכלוסייה תומך הממשלה?
5. משרד הבריאות מתכנן לבצע מדגם שמטרתו לבדוק את הסיכוי לחלות בשפעת עם לקיחת חיסון נגד שפעת. הוא מעוניין שבסיכוי של 98% טעות האמידה לא תעלה על 3%.
 - א. כמה מחוסנים יש לדגום ?
 - ב. משרד הבריאות ביצע את המדגם שאת גודלו חיבתם בסעיף הקודם וקיבל ש 15% מבין אלה שקיבלו חיסון נגד שפעת בכל זאת חלו במשך החורף בשפעת. בנו ברמת סמך של 98% את הסיכוי לחלות בחורף בשפעת עם לקיחת חיסון נגד שפעת.
 - ג. בהמשך לסעיף הקודם. מהי טעות האמידה המרבית בביטחון של 98% ? מדוע הוא קטן מ-3% ?

פתרונות:

שאלה 1

1068

שאלה 3

א. 601

ב. 108000 ₪.

פרק 25 - בדיקת השערות על פרמטרים

הקדמה

רקע:

תהליך של בדיקת השערות הוא תהליך מאוד 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) ו**אזור 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 הברורות 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 :

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

הגדרת הטעויות:

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

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

דוגמה: (פתרון בהקלטה)

אדם חשוד בביצוע עבירה ונתבע בבית המשפט. אילו סוגי טעויות אפשריות בהכרעת הדין?

תרגילים:

1. לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 סמ"ק במדגם בגודל 25. בסופו של דבר הוחלט להכריע לטובת חברת המשקאות.
 - א. רשמו את השערות המחקר.
 - ב. מה מסקנת המחקר?
 - ג. איזו סוג טעות יתכן וביצעו במחקר?

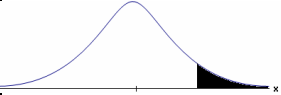


2. במחקר על פרמטר מסוים הוחלט בסופו של דבר לדחות את השערת האפס.
 - א. האם ניתן לדעת אם בוצע טעות במחקר?
 - ב. מה סוג הטעות האפשרית?

3. לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטיית תקן 0.4. ישנה טענה שכיום ממוצע מספר הילדים במשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה. על סמך תוצאות המדגם נקבע שלא ניתן לקבוע שבאופן מובהק תוחלת מספר הילדים למשפחה קטנה כיום.
 - א. מהי אוכלוסיית המחקר?
 - ב. מה המשתנה הנחקר?
 - ג. מה הפרמטר הנחקר?
 - ד. מה השערות המחקר?
 - ה. מה מסקנת המחקר?
 - ו. מהי סוג הטעות האפשרית במחקר?

פרק 26 - בדיקת השערות על תוחלת (ממוצע)

כאשר שונות האוכלוסיה ידועה

רקע:

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

סטטיסטי המבחן :

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

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.
- הגדלת רמת המובהקות לא תשתנה את מסקנת המחקר.
 - הגדלת רמת המובהקות תשנה את מסקנת המחקר.
 - הקטנת רמת המובהקות לא תשנה את מסקנת המחקר.
 - 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)
- השערת האפס הייתה נדחית.
 - השערת האפס הייתה לא נדחית.
 - 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.
- אם חוקר א' החליט לדחות את H_0 , מה יחליט חוקר ב'? נמקו.
 - אם חוקר א' יחליט לא לדחות את H_0 , מה יחליט חוקר ב'? נמקו.

פתרונות :**שאלה 1:** H_0 נקבל**שאלה 2:** H_0 נדחה**שאלה 3:** H_0 נדחה**שאלה 4:** H_0 נדחה**שאלה 5:** H_0 נקבל**שאלה 6:**

א

שאלה 7:

ג

שאלה 8:

א. אותה מסקנה

ב. לא ניתן לדעת.

סיכוי לטעויות ועוצמה כאשר שונות האוכלוסייה ידועה

רקע:

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

נגדיר את ההסתברויות הבאות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות)

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

הסיכוי לבצע טעות מסוג 2:

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

רמת בטחון:

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

עוצמה:

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

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 | | | תנאים : |
| 4. $X \sim N$ או מדגם מספיק גדול | | | |
| $\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$ | $\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$ | $\bar{X} > \mu_0 + Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ | כלל ההכרעה : אזור הדחייה של H_0 : |
| $P_{H_1}(\bar{X} < \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}})$ | $P_{H_1}(\bar{X} > \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}})$ | $P_{H_1}(\mu_0 - Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + Z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$ | 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:

לפתרון מלא בסרטון וידאו היכנסו ל- www.GooL.co.il

כתב ופתר - ברק קנדל ©

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 ברריות. 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 : $\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n})$

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

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 נורמאלית עם תוחלת אליה כוונה המכונה וסטיות 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 כיוול.

שאלה 5:

נכון

שאלה 6:

תשובה: א

שאלה 7:

תשובה: א

שאלה 8:

תשובה: ג

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 באוכלוסייה 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 פנצלין 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:

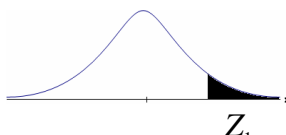
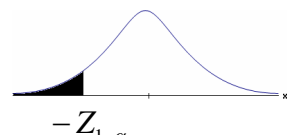
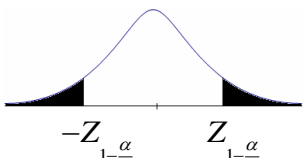
א. $191.8 \leq \mu \leq 200.2$

ב. נכריע שיש אמת בפרסום.

פרק 27 - בדיקת השערות על פרופורציה

התהליך

רקע:

| | | | |
|---|---|--|--|
| $H_0 : p = p_0$ $H_1 : p > p_0$ | $H_0 : p = p_0$ $H_1 : p < p_0$ | $H_0 : p = p_0$ $H_1 : p \neq p_0$ | השערות האפס : השערה אלטרנטיבית: |
| $np_0 \geq 5 \ \& \ n(1-p_0) \geq 5$ | | | תנאים: |
| $Z_{\hat{p}} > Z_{1-\alpha}$  $Z_{1-\alpha}$ - דוחים את H_0 | $Z_{\hat{p}} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ - דוחים את H_0 | או $Z_{\hat{p}} < -Z_{1-\frac{\alpha}{2}}$ $Z_{\hat{p}} > Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}}$ $Z_{1-\frac{\alpha}{2}}$ - דוחים את H_0 | כלל ההכרעה: אזור הדחייה של : |

סטטיסטי המבחן :

$$Z_{\hat{p}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

חלופה אחרת לכלל הכרעה:

| | | | |
|--|--|--|--|
| $\hat{p} > p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ | $\hat{p} < p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ | $\hat{p} > p_0 + Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ או $\hat{p} < p_0 - Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ | כלל ההכרעה: אזור הדחייה של H_0 |
|--|--|--|--|

דוגמה: (פתרון בהקלטה)

בחודש ינואר השנה פורסם שאחוז האבטלה במשק הוא 8% במדגם עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים. בדקו ברמת מובהקות של 5% האם כיום אחוז האבטלה הוא כמו בתחילת השנה.

תרגילים:

1. במשך שנים אחוז המועמדים שהתקבל לפקולטה מסוימת היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. ברמת מובהקות של 5% האם השנה הקשו על תנאי הקבלה?
2. במדגם של 300 אזרחים 57% מתנגדים להצעת חוק מסוימת. לאור נתונים אלה האם רוב האזרחים מתנגדים להצעת החוק? בדקו ברמת מובהקות של 10%.
3. הטילו מטבע 50 פעמים וקיבלו 28 פעמים עץ. האם המטבע הוגן ברמת מובהקות של 5%?
4. קפיטריה במכללה מסוימת מעריכה כי אחוז הסטודנטים שקונים קפה בקפיטריה הינו 20%. נערך סקר אשר כלל 200 סטודנטים. התברר כי 33 מהם רוכשים קפה בקפיטריה. מטרת הסקר הייתה לבדוק את אמיתות הערכה של הקפיטריה.
 - א. רשמו את ההשערות.
 - ב. בדוק את ההשערות ברמת מובהקות של 10%.
 - ג. מה תהיה המסקנה אם נקטין את רמת המובהקות?
5. חבר כנסת רוצה להעביר חוק. לצורך כך הוא דוגם 400 אזרחים במטרה לבדוק האם רוב האזרחים תומכים בחוק. במדגם התקבל ש-276 אזרחים תומכים בחוק.
 - א. מה מסקנתכם ברמת מובהקות של 5%?
 - ב. האם ניתן לדעת מה תהיה המסקנה אם רמת המובהקות תהיה גדולה יותר? הסבירו.
6. שני חוקרים בדקו את ההשערות הבאות:

$$H_0 : p = p_0$$

$$H_1 : p > p_0$$

חוקר א השתמש ברמת מובהקות α_1 וחוקר ב ברמת מובהקות α_2 החוקר הראשון דחה את H_0 ואילו החוקר השני קיבל את H_0 . שניהם התבססו על אותם תוצאות של מדגם. בחר בתשובה הנכונה:

 - א. $\alpha_1 = \alpha_2$
 - ב. $\alpha_1 > \alpha_2$
 - ג. $\alpha_1 < \alpha_2$
 - ד. המצב המתואר לא אפשרי.

פתרונות :**שאלה 1:** H_0 נדחה**שאלה 2:** H_0 נדחה**שאלה 3:** H_0 נקבל**שאלה 4:**ב. נקבל H_0

ג. המסקנה לא תשתנה.

שאלה 5:א. נדחה H_0

ב. המסקנה לא תשתנה.

שאלה 6:

התשובה היא : ג.

סיכוי לטעויות ועוצמה**רקע:**

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

נגדיר את ההסתברויות הבאות:**הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):**

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

הסיכוי לבצע טעות מסוג 2:

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

רמת בטחון:

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

עוצמה:

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

התהליך לחישוב סיכוי לטעות מסוג שני:

| | | | |
|---|---|--|---|
| $H_0 : p = p_0$ $H_1 : p > p_0$ | $H_0 : p = p_0$ $H_1 : p < p_0$ | $H_0 : p = p_0$ $H_1 : p \neq p_0$ | השערת האפס : השערה אלטרנטיבית: |
| $np_0 \geq 5 \ \& \ n(1-p_0) \geq 5$ | | | תנאים: |
| $\hat{p} > p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ | $\hat{p} < p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ | $\hat{p} > p_0 + Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ או $\hat{p} < p_0 - Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ | כלל ההכרעה: אזור הדחייה של H_0 |
| $P_{H_1}(\hat{p} < p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}})$ | $P_{H_1}(\hat{p} > p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}})$ | $P_{H_1}(p_0 - Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} < \hat{p} < p_0 + Z_{1-\alpha/2} \cdot \sqrt{\frac{p_0(1-p_0)}{n}})$ | חישוב β : |

כאשר : $\hat{P} \sim N(p, \frac{p(1-p)}{n})$

והתקנון : $Z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$

דוגמה: (פתרון בהקלטה)

רופאי שיניים טוענים שיותר ממחצית האוכלוסייה הבוגרת בארץ אינם מבקרים אצל רופא שיניים באופן קבוע, כנדרש. כדי לבדוק טענה זו, נערך סקר בקרב 150 אנשים בוגרים.

א. רשמו את ההשערות וכלל הכרעה ברמת מובהקות של 10%.

ב. מהי עוצמת המבחן אם מסתבר ש 60% מהאוכלוסייה אינם מבקרים אצל רופא שיניים באופן קבוע.

תרגילים:

1. משרד הבריאות פרסם ש 10% מתושבי המדינה סובלים ממחלת האסטמה. מחקר דורש לבדוק האם בחיפה, בגלל זיהום האוויר, שיעור הסובלים מאסטמה גבוה יותר. לצורך המחקר נבדקו 260 מתושבי חיפה.
 - א. רשמו את השערות המחקר, וצרו מבחן ברמת מובהקות של 5% לבדיקתן.
 - ב. מהי עצמת המבחן של סעיף א' בהנחה ובחיפה 16% מהתושבים סובלים מאסטמה?
 - ג. כיצד תשנה התשובה לסעיף ב' אם מסתבר שבחיפה 18% סובלים מאסטמה?
 - ד. בהמשך לסעיף א' האם נכון לומר שבהסתברות של 5% ההשערה שבחיפה 10% מהתושבים סובלים מאסטמה אינה נכונה?

2. אחוז הסובלים מתופעות הלוואי מתרופה מסוימת הוא 15%. חברת תרופות טוענת שפיתחה תרופה שאמורה לצמצם את אחוז הסובלים מתופעות לוואי. לצורך בדיקת הטענה הוחלט לבצע מחקר שיכלול 120 חולים שיקבלו את התרופה הנבדקת.
 - א. נניח שהתרופה נבדקת אכן מורידה את פרופורציות הסובלים מתופעות הלוואי ל-10% מהי עצמת המבחן עבור רמת מובהקות של 5%?

3. בעיר מסוימת היו 20% אקדמאים. בעקבות פתיחת מכללה בעיר לפני כמה שנים מעוניינים לבדוק האם אחוז האקדמאים גדל. מעוניינים שהמחקר יכלול 200 אנשים והוא יהיה ברמת מובהקות של 5%.
 - א. חשבו את הסיכוי לבצע טעות מסוג שני בהנחה והיום יש 28% אקדמאים.
 - ב. כיצד התשובה לסעיף הקודם תשתנה אם נגדיל את רמת המובהקות?

4. מעוניינים לבדוק האם בפקולטה מסוימת ישנה העדפה לגברים. הוחלט לדגום 200 מתקבלים ועל סמך מספר הבנים לקבוע אם טענת המחקר מתקבלת.

חוקר א' קבע רמת מובהקות של 5% וחוקר ב' החליט לקבל את טענת המחקר אם במדגם יהיו לפחות 120 בנים. למי מבין החוקרים רמת מובהקות גדולה יותר?

5. חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן (בחר בתשובה הנכונה)
 - א. השערת האפס נכונה.
 - ב. השערת האפס נדחתה.
 - ג. השערת האפס לא נדחתה.
 - ד. אף אחת מהתשובות לא נכונה בהכרח.

6. קבע אם הטענה הבאה נכונה:

"בבדיקת השערות לא ניתן לבצע בו זמנית טעות מסוג ראשון וטעות מסוג שני"

פתרונות:**שאלה 1:**

ב. 0.9015

ג. תגדל

ד. טענה לא נכונה.

שאלה 2:

0.4404

שאלה 3:

א. 0.1446

ב. תקטן.

שאלה 4:

חוקר א.

שאלה 5:

התשובה הנכונה היא ג.

שאלה 6:

נכונה.

מובהקות התוצאה

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.

| | | | |
|--------------------------------------|------------------------------------|--|---|
| $H_0 : p = p_0$ $H_1 : p > p_0$ | $H_0 : p = p_0$ $H_1 : p < p_0$ | $H_0 : p = p_0$ $H_1 : p \neq p_0$ | השערת האפס : השערה אלטרנטיבית: |
| $np_0 \geq 5 \ \& \ n(1-p_0) \geq 5$ | | | תנאים: |
| $P_{H_0}(\hat{P} \geq \hat{p})$ | $P_{H_0}(\hat{P} \leq \hat{p})$ | אם $2 \cdot P_{H_0}(\hat{P} \geq \hat{p}) \Leftarrow \hat{p} > p_0$ אם $2 \cdot P_{H_0}(\hat{P} \leq \hat{p}) \Leftarrow \hat{p} < p_0$ | p-value |

$$\hat{P} \sim N(p_0, \frac{p_0(1-p_0)}{n}) : \text{ כאשר בהנחת השערת האפס}$$

והתקנון:

$$Z_{\hat{p}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

דוגמה: (פתרון בהקלטה)

ישנה טענה שיש הבדל בין אחוז הבנים ואחוז הבנות הפונים ללמוד להנדסאי מחשבים. לשם כך נלקח מדגם מקרי של 200 תלמידים הלומדים מחשבים והתברר כי 112 מהם בנים.

א. מהי מובהקות התוצאה?

ב. מה המסקנה ברמת מובהקות של 5%?

תרגילים:

1. במשך שנים אחוז המועמדים שהתקבל לפקולטה מסוימת היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. רוצים לבדוק האם השנה הקשו על תנאי הקבלה.
 - א. מהי מובהקות התוצאה?
 - ב. מה תהיה המסקנה ברמת מובהקות של 1% וברמת מובהקות של 5%?

2. נהוג לחשוב ש 60% מהילדים בגיל שלוש קמים מהמיטה במהלך הלילה לפחות פעם אחת. ישנה טענה שללא שנת צהריים פחות מ-60% מהילדים בגיל זה יקומו לפחות פעם אחת במהלך הלילה. נדגמו 80 ילדים בגיל 3 אשר אינם ישנים בצהריים מתוכם התקבל ש 41 קמו במהלך הלילה.
 - א. מהי רמת המובהקות המינימלית עבורה תתקבל הטענה במחקר?
 - ב. מהי רמת המובהקות המקסימלית עבורה לא תתקבל טענת המחקר?
 - ג. עבור אילו רמות מובהקות נקבל את טענת המחקר?
 - ד. מה תהיה מסקנת המחקר ברמת מובהקות של 6%?

3. במטרה לבדוק האם מטבע הוא הוגן מטילים אותו 80 פעמים. התקבל ש 60 מההטלות הראו עץ. רשמו את השערות המחקר, חשבו את מובהקות התוצאה והסיקו מסקנה ברמת מובהקות של 5%.

4. בבדיקת השערות על פרופורציה התקבל שה- $p\text{-value}=0.02$.
 - א. יקבל את השערת האפס
 - ב. ידחה את השערת האפס.
 - ג. לא ניתן לדעת כי אין מספיק נתונים.

5. קבע אם הטענה הבאה נכונה:

"במבחן לבדיקת השערות חד-צדדי התקבל ערך $p\text{-value}$ של 3% לכן אם היינו מבצעים מבחן דו-צדדי (כאשר יתר הנתונים ללא שינוי) היינו מקבלים ערך $p\text{-value}$ של 6%"

6. במפעל 10% מהעובדים נפגעים לפחות פעם אחת בשנה מתאונות עבודה. לאור זאת, המפעל החליט לצאת בתוכנית לצמצום שיעור הנפגעים. תכנית זו נוסתה על 100 עובדים. מתוכם 12 נפגעו בתאונות עבודה במשך השנה. מהי רמת המובהקות הקטנה ביותר עבורה יוחלט שהתכנית יעילה?

פתרונות :**שאלה 1:**

א. 0.0455

שאלה 2:

א. 0.0548

ב. 0.0548

ג. מעל 0.0548

ד. נכריע לטובת טענת המחקר.

שאלה 3:

$$p_v = 0$$

שאלה 4:

התשובה הנכונה : ב

שאלה 5:

הטענה נכונה

שאלה 6:

0.7486