

סטטיסטיקה תיאורית מסלול ניצבים א



$$\{\sqrt{x}\}^2$$



תוכן העניינים

1. סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה 1
2. סטטיסטיקה תיאורית- הצגה של נתונים 5
3. סטטיסטיקה תיאורית- סכימה 13
4. סטטיסטיקה תיאורית- מדדי מרכז (ללא ספר) 13
5. סטטיסטיקה תיאורית - מדדי פיזור - הטווח, השונות וסטיית התקן 17
6. סטטיסטיקה תיאורית - מדדי פיזור - טווח בין רבעוני 20
7. סטטיסטיקה תיאורית- מדדי מיקום יחסי-ציון תקן 24
8. סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה 26
9. סטטיסטיקה תיאורית - טרנספורמציה לינארית 28
10. סטטיסטיקה תיאורית - מדדי אסימטריה 30
11. סטטיסטיקה תיאורית- ניתוח פלטים 34
12. סטטיסטיקה תיאורית שאלות אמריקאיות 36
13. התפלגויות רציפות מיוחדות - התפלגות נורמלית 43

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 1 - סטטיסטיקה תיאורית- סיווג משתנים וסולמות מדידה

תוכן העניינים

1. כללי..... 1

סטטיסטיקה תיאורית – סיווג משתנים וסולמות מדידה:

רקע:

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

מיון משתנים לפי סולמות המדידה:

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

סולם מנה/יחס:

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

סוגי משתנים:

נבצע סיווג של המשתנים:

משתנה איכותי

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

משתנה כמותי

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

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

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



שאלות:

- (1) באיזה סולם מדידה המשתנים הבאים נחקרים (שמי/סדר/רווחים/מנה):
- גובה (בס"מ).
 - מספר ילדים למשפחה.
 - מידת החרדה לפני מבחן.
 - שביעות רצון משירות לקוחות בסקלה מ-1 עד 7 (1 - כלל לא מרוצה עד 7 - מרוצה מאד)
 - השכלה.
 - מספר אוטובוס.
 - מקום מגורים.
 - מין (1=גבר ; 2=אישה).
 - מידת נעליים.

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

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

בחברה 200 עובדים.

- מהו המשתנה הנחקר כאן?
 - האם מדובר במשתנה איכותי או כמותי?
אם הוא כמותי האם הוא בדיד או רציף?
באיזה סולם מדידה המשתנה?
- (3) להלן רשימה של משתנים כמותיים. ציינו האם הוא משתנה רציף/בדיד:
- שכר ב-ש.
 - ציון בחינת בגרות.
 - תוצאה של הטלת קובייה.
 - מהירות ריצה בתחרות.
 - שיעור התמיכה בממשלה.

תשובות סופיות:

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

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 2 - סטטיסטיקה תיאורית- הצגה של נתונים

תוכן העניינים

1. כללי 5

סטטיסטיקה תיאורית – הצגה של נתונים:

רקע:

דרכים להצגת נתונים שנאספו:

רשימה של תצפיות:

התצפית היא הערך שנצפה עבור ישות מסוימת בקבוצה. רושמים את התצפיות שהתקבלו כרשומה, יעיל שיש מספר מועט של תצפיות. ההצגה הזו רלבנטית לכל סוגי המשתנים. למשל, להלן מספר החדרים בבניין בן 5 דירות: 4, 3, 5, 4, 3.

טבלת שכיחויות בדידה:

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

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

דוגמה:

להלן התפלגות הציונים בכיתה מסוימת:

$\frac{f_i}{n}$	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.

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

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

$$\frac{f_i}{n} - \text{איזה חלק מהתצפיות בקבוצה שוות לערך.}$$

טבלת שכיחויות במחלקות:

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

דוגמה:

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

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

דיאגרמת עוגה:

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

התפלגות המצב המשפחתי



דיאגרמת מקלות:

הציר האופקי הוא הציר של המשתנה והציר האנכי של השכיחות, כך שהגובה של המקל מעיד על השכיחות. רלבנטי למשתנה כמותי בדיד. לא נהוג להשתמש בתיאור למשתנה איכותי וכמו כן לא למשתנה כמותי רציף, וכן בסולמות מדידה עבור משתנה מסולם סדר.



היסטוגרמה:

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

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



פוליגון – מצולעון:

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

צורות התפלגות נפוצות:

התפלגות סימטרית פעמונית

רוב התצפיות במרכז, וככל שנתרחק מהמרכז יהיו פחות תצפיות באופן סימטרי. לדוגמה, ציוני IQ.

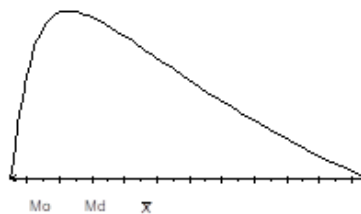


ישנן התפלגויות סימטריות שאינן פעמוניות, כגון:

התפלגות אסימטרית ימנית (חיובית)

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

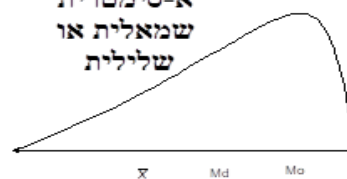
התפלגות א-סימטרית ימנית או חיובית



התפלגות אסימטרית שמאלית (שלילית)

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

התפלגות א-סימטרית שמאלית או שלילית



שאלות:

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

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

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

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

- 3) להלן התפלגות ההשכלה במקום עבודה מסוים:

השכלה	מספר העובדים
נמוכה	60
תיכונית	120
אקדמאית	20

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

- 4) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
- 6, 5, 8, 7, 6, 7, 6, 8, 7, 8, 5, 4, 6, 10, 9, 8, 6, 7.
- א. מהו המשתנה? האם הוא בדיד או רציף?
- ב. תארו את הרשימה בטבלת שכיחויות.
- ג. הוסיפו שכיחויות יחסיות לטבלה.
- ד. תארו את הנתונים באופן גרפי.

תשובות סופיות:

ב. עיין גרף מלא בסרטון הוידאו.

1) א. להלן טבלה:

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

ב. 19.6%.

2) א. מקצוע מועדף.

ב. עיין גרף מלא בסרטון הוידאו.

3) א. משתנה נחקר: השכלה, סוג: סדר.

ב+ג. להלן טבלה:

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

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

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 3 - סטטיסטיקה תיאורית- סכימה

תוכן העניינים

1. כללי.....13

סטטיסטיקה תיאורית – סכימה:

רקע:

בסטטיסטיקה ישנה צורת רישום מקובלת לסכום של תצפיות: $\sum_{i=1}^n X_i$.

נסביר את צורת הרישום על ידי הדוגמה הבאה:

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

(הסבר מלא מופיע בסרטונים באתר).

שאלות:

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

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

א. $\sum_{i=1}^3 X_i$

ב. $\sum_{i=1}^5 Y_i$

ג. $\sum_{i=1}^4 X_i$

ד. $\left(\sum_{i=1}^4 X_i\right)^2$

ה. $\sum X_i$

ו. $\sum X_i Y_i$

ז. $\sum(X_i) \sum(Y_i)$

- (2) נתון לוח ערכי המשתנים X_i ו- Y_i , כאשר: $i = 1, 2, \dots, 6$, ונתונים הקבועים:
 $a = 2$, $b = 5$. חשבו את הנוסחאות הבאות:

i	1	2	3	4	5	6
X_i	3	2	4	-2	1	4
Y_i	2	0	0	1	-5	2

$$\text{א. } \sum_{i=1}^4 y_i$$

$$\text{ב. } \sum_{i=1}^6 a$$

$$\text{ג. } \sum_{i=1}^6 x_i y_i$$

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

$$\text{ה. } \sum_{i=1}^6 x_i + a$$

- (3) קבעו לכל זהות האם היא נכונה:

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

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

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

תשובות סופיות:

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

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 4 - סטטיסטיקה תאורית- מדדי מרכז

תוכן העניינים

1. כללי (ללא ספר)

סטטיסטיקה תיאורית מסלול ניצבים א

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

תוכן העניינים

1. כללי 17

סטטיסטיקה תיאורית – מדדי פיזור – הטווח, השונות וסטיית התקן:

רקע:

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

הטווח / תחום (RANGE):

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

שונות וסטיית תקן:

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

$$S_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2$$

עבור סדרת נתונים:

דוגמאות:

(1) נחשב את השונות של סדרת המספרים הבאה: 5, 4, 9.

$$S_x^2 = \frac{\sum (x - \bar{x})^2 f}{n} = \frac{\sum x^2 \cdot f}{n} - \bar{x}^2$$

עבור טבלת שכיחויות:

(2) להלן התפלגות הציונים בכיתה מסוימת בה ממוצע הציונים הוא 7.44.

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

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

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

כשיש מחלקות נעזר באמצע המחלקה כדי לחשב את השונות.

שאלות:

(1) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
6, 5, 8, 7, 6, 8, 6, 7, 8, 5, 6, 4, 5, 8, 7, 6, 4, 5, 6, 8, 9, 10, 6, 8, 9, 7.
חשבו את השונות, סטיית התקן והטווח של הציונים.

(2) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

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

א. חשבו סטיית התקן.

ב. חשבו את הטווח של הנתונים.

הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם.

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

א. האם הממוצע יגדל/יקטן/לא ישתנה וסטיית התקן תגדל/תקטן/לא

תשנה כאשר יתווספו שני עובדים עם וותק של 4 שנים להתפלגות?

ב. האם הממוצע יגדל/יקטן/לא ישתנה וסטיית התקן תגדל/תקטן/לא

תשנה כאשר יתווספו שני עובדים אשר אחד עם וותק של 0 שנים והשני

עם וותק של 8 שנים להתפלגות?

(4) נתונה רשימה של 5 תצפיות, אך רק עבור 4 מהן נרשמו הסטיות שלהן

מהממוצע: 2, 3, 2, -1. חשבו את השונות של חמש התצפיות.

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

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

א. מה הממוצע של מספר החדרים לשכונה בדירה?

ב. חשבו את סטיית התקן של מספר החדרים לדירה.

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

הדבר ישפיע(יקטין, יגדל, לא ישנה)

על כל מדד שחישבתם בסעיפים הקודמים

תשובות סופיות:

- (1) שונות: 2.19, סטיית תקן: 1.48, טווח: 6.
- (2) א. סטיית תקן: 1.106. ב. טווח: 4.
- (3) א. ממוצע לא ישתנה, סטיית התקן תקטן.
ב. ממוצע לא ישתנה, סטיית התקן תגדל.
- (4) 10.8.
- (5) א. 3.05. ב. 1.16. ג. ממוצע: יקטן, סטיית התקן: תגדל.

סטטיסטיקה תיאורית מסלול ניצבים א

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

תוכן העניינים

1. טווח בינרבעוני.....20

סטטיסטיקה תיאורית – מדדי פיזור – טווח בין רבעוני:

רקע:

הטווח הבין-רבעוני (יש הקוראים לו התחום הבין-רבעוני) נותן את הטווח בין הרבעונים בו נמצאים 50% מהתצפיות המרכזיות. הרעיון ליצור מדד פיזורי שלא רגיש לתצפיות החריגות ביותר. כדי לחשב את הטווח הבין-רבעוני יש למצוא את הרבעון התחתון והעליון של התפלגות התצפיות.

רבעון תחתון – ערך שמחלק את ההתפלגות לשניים.
25% מהמקרים נמוכים ממנו או שווים לו ו-75% מהמקרים גבוהים או שווים לו.
סימון: Q_1 .

רבעון עליון – ערך שמחלק את ההתפלגות לשניים.
75% מהמקרים נמוכים ממנו או שווים לו ו-25% מהמקרים גבוהים או שווים לו.
סימון: Q_3 .

הטווח הבין רבעוני הוא הפער בין שני הרבעונים: $IQR = Q_3 - Q_1$.

שלב ב' במציאת טווח בין-רבעוני בטבלת שכיחויות:

שלב א': נמצא את הרבעון התחתון: הוא הערך שהשכיחות היחסית המצטברת באחוזים עברה לראשונה את 25%.

שלב ב': נמצא את הרבעון העליון: הוא הערך שהשכיחות היחסית המצטברת באחוזים עברה לראשונה את 75%.

שלב ג': נמצא את הטווח הבין-רבעוני: נחסר את הרבעונים: $IQR = Q_3 - Q_1$.

דוגמה (פתרון בהקלטה):

בסניף בנק 250 לקוחות. ספרו לכל לקוח את מספר תוכניות החיסכון שלו. מהו הטווח הבין-רבעוני של מספר תוכניות החיסכון בסניף?

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

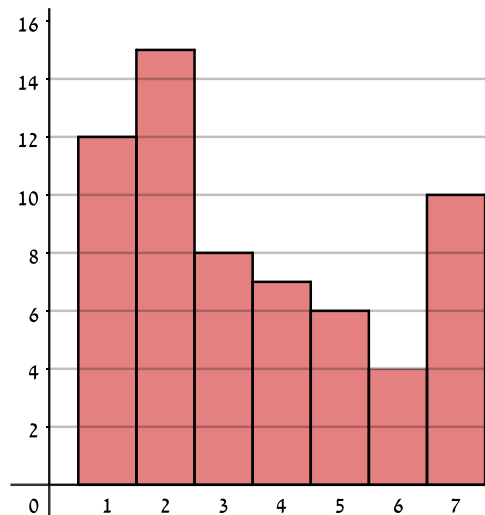
שאלות:

1) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

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

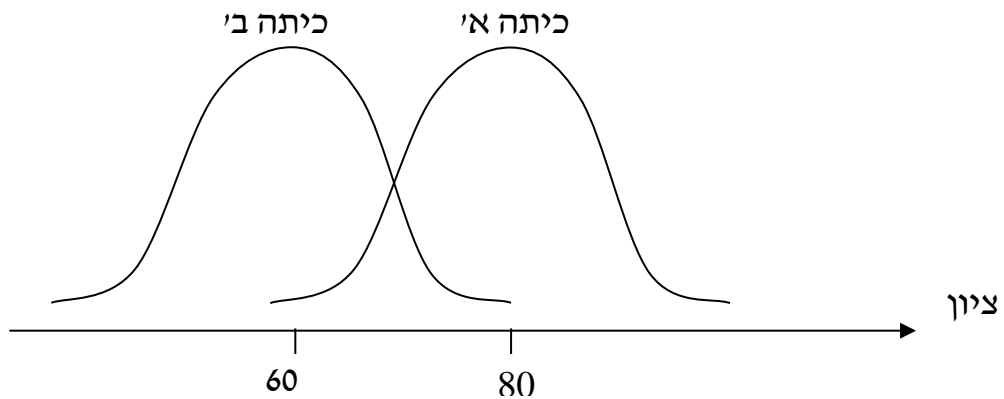
מהו הטווח הבין-רבעוני של מספר המכוניות למשפחה בישוב "הגורן"?

2) בסקר שנעשה בדקו את מספר ימי המחלה השנתיים של מורים בארץ.



- א. מה מייצגים הערכים בציר האופקי?
- ב. מהו הטווח הבין-רבעוני של מספר ימי המחלה של המורים?
- ג. אם נוסף 25 מורים אשר הצהירו שמספר ימי המחלה השנתיים שלהם הוא 4 ימים, כיצד הדבר ישנה את הטווח הבין-רבעוני? הסבירו.
- ד. אם מסתבר שחלק מהמורים בסקר הצהירו שהם חלו 7 ימים בשנה אבל בפועל הם חלו 8 ימים, כיצד הדבר ישנה את הטווח הבין-רבעוני? הסבירו.

3) לפניך שתי עקומות המתארות את התפלגות הציונים בכל כיתה. באיזו כיתה הטווח הבין-רבעוני גדול יותר?



- א. כיתה א.
- ב. כיתה ב'.
- ג. לשתיהן אותו טווח בין-רבעוני.
- ד. לא ניתן לדעת, אין מספיק נתונים.

4) הוספת גודל קבוע לכל תצפיות סדרת נתונים :

- א. תגדיל את הטווח הבין-רבעוני.
- ב. תקטין את הטווח הבין-רבעוני.
- ג. לא תשנה הטווח הבין-רבעוני.
- ד. לא ניתן לדעת מה יקרה לטווח הבין-רבעוני.

5) חושב הטווח הבין-רבעוני עבור התפלגות מסוימת והתקבלה התוצאה אפס. לכן :

- א. לפחות 50% מהתצפיות זהות.
- ב. סטיית התקן היא אפס.
- ג. ההתפלגות היא סימטרית.
- ד. מצב זה כלל לא יתכן.

- 6) סניף מספר 543 של בנק "רווה" בדק ל-80 לקוחות את מספר הפעמים שכל לקוח נכנס לסניף הבנק במשך שבוע. התוצאות שהתקבלו הן:
- 50 אנשים נכנסו 0 פעמים לסניף.
 - 20 אנשים נכנסו פעם אחת לסניף.
 - 5 אנשים נכנסו פעמיים לסניף.
 - 5 אנשים נכנסו יותר מפעמיים.
- מהו הטווח הבין-רבעוני?
- א. 60.
 - ב. 2.
 - ג. 50.
 - ד. 1.

- 7) התפלגות הציונים במבחן ווקסלר היא סימטרית לכן:
- א. טווח הציונים הוא אפס.
 - ב. הטווח הבין רבעוני של הציונים אפס.
 - ג. סעיפים א ו-ב הם נכונים.
 - ד. אף סעיף אינו נכון.

תשובות סופיות:

- 1) 2.
- 2) א. מספר ימי המחלה השנתיים. ב. 3. ג. יקטן. ד. לא ישתנה.
- 3) ג.
- 4) ג.
- 5) א.
- 6) ד.
- 7) ד.

סטטיסטיקה תיאורית מסלול ניצבים א

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

תוכן העניינים

1. כללי 24

סטטיסטיקה תיאורית – מדדי מיקום יחסי – ציון תקן:

רקע:

המטרה למדוד איך תצפית ממוקמת ביחס לשאר התצפיות בהתפלגות.

ציון תקן:

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

- ציון התקן נותן כמה סטיות תקן סוטה התצפית מהממוצע. כלומר, ציון התקן מעיד על כמה סטיות תקן התצפית מעל או מתחת לממוצע:
- ציון תקן חיובי אומר שהתצפית מעל הממוצע.
 - ציון תקן שלילי אומר שהתצפית מתחת לממוצע.
 - ציון תקן אפס אומר שהתצפית בדיוק בממוצע.

דוגמה (פתרון בהקלטה):

במקום עבודה מסוים, ממוצע המשכורות הוא 8 אלף ₪, עם סטית תקן של אלפיים ₪. באותו מקום עבודה ההשכלה הממוצעת של העובדים הנה 14 שנים, עם סטית תקן של 1.5 שנים. ערן מרוויח במקום עבודה זה 11 אלף ₪ והשכלתו 16 שנים. מה ערן יותר, באופן יחסי, משכיל או משתכר?

שאלות:

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

המקצוע	ממוצע	סטיית תקן
לשון	74	12
מתמטיקה	80	16

עודד קיבל: 68 בלשון ו-70 במתמטיקה.

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

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

מספר פועלים	תפוקה	ממוצע
15	48	
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?
 א. התפוקה.
 ב. כמות הפועלים.
 ג. חריגים באותה מידה.
 ד. חסרים נתונים כדי לדעת זאת.

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

תשובות סופיות:

- 1) א. לשון. ב. 72.
 2) ב'.
 3) א. משקל. ב. 70.

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 8 - סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה

תוכן העניינים

1. כללי.....26

סטטיסטיקה תיאורית – מדדי מיקום יחסי – אחוזונים בטבלה בדידה:

רקע:

האחוזון (המאון) ה- p הוא הערך בנתונים המחלק את הנתונים בצורה כזאת, שעד אליו (כולל) יש $p\%$ מהנתונים. מסמנים את האחוזון ה- p ב- X_p .

חישוב האחוזון מתוך נתונים בטבלת שכיחויות בדידה:

האחוזון הוא הערך שבו בפעם הראשונה השכיחות היחסית המצטברת (באחוזים) גדולה או שווה ל- $p\%$.

דוגמה (פתרון בהקלטה):

בסניף בנק 250 לקוחות. ספרו לכל לקוח את מספר תוכניות החיסכון שלו:

שכיחות יחסית מצטברת	שכיחות מצטברת	$F(x)$	# תוכניות החיסכון
		100	0
		75	1
		25	2
		25	3
		25	4

א. מצאו את האחוזון ה-25.

ב. מצאו את הערך ש-20% מהמקרים מעליו.

שאלות:

(1) להלן התפלגות של משתנה כלשהו:

$F(x)$	X
10	0
40	1
30	2
15	3
5	4

מצאו להתפלגות את:

- האחוזון ה-60.
- המאון ה-40.
- העשירון העליון.
- הטווח בין הרבעונים.

(2) להלן התפלגות מספר המכונניות למשפחה בישוב "הגורן":

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

חשבו את:

- העשירון התחתון.
- האחוזון ה-30.
- הערך ש-20% מהתצפית גדולות ממנו.
- רבעון עליון.

תשובות סופיות:

- (1) א. 2 ב. 1 ג. 3 ד. 1
- (2) א. 1 ב. 2 ג. 4 ד. 4

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 9 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

תוכן העניינים

1. כללי 28

סטטיסטיקה תיאורית – טרנספורמציה לינארית:

רקע:

מצב שבו מבצעים שינוי מסוג הוספה (או החסרה) של קבוע, והכפלה (או חילוק) של קבוע, לכל התצפיות: $y = a \cdot x + b$. כך יושפעו המדדים השונים:

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

$$MO_y = a \cdot MO_x + b$$

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

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

מדדי המרכז:

$$R_y = |a| R_x$$

$$S_y = |a| S_x$$

$$S_y^2 = a^2 S_x^2$$

מדדי הפיזור:

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

$$Z_y = \frac{a}{|a|} Z_x$$

מדדי המיקום היחסי:

שלבי העבודה:

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

דוגמה (פתרון בהקלטה):

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

שאלות:

- (1) עבור סדרת נתונים התקבל: $\bar{x} = 80, S = 15, MO = 70$. הוחלט להכפיל את כל התצפיות ב-4 ולהחסיר מהתוצאה 5. חשבו את המדדים הללו לאחר השינוי.
- (2) בחברה מסוימת השכר הממוצע הוא 40 ₪ לשעה עם סטיית תקן של 5 ₪ לשעה. הוחלט להעלות את כל המשכורות ב-10%, אך זה לא סיפק את העובדים ולכן הם קיבלו לאחר מכן תוספת של 2 ₪ לשעה. מה הממוצע ומהי השונות של השכר לשעה לאחר כל השינויים.
- (3) במבחן מסוים הציון החציוני היה 73, טווח הציונים היה 40 נקודות והעשירון העליון היה הציון 87. כיוון שהציונים בבחינה היו נמוכים, המורה החליט לתת פקטור של 4 נק' לכל התלמידים. חשבו את המדדים לאחר הפקטור.
- (4) דגמו מקו ייצור 50 קופסאות של גפרורים. בדקו בכל קופסא בה יש 40 גפרורים את כמות הגפרורים הפגומים. התקבל שבממוצע יש 3 גפרורים פגומים בקופסא, עם סטיית תקן של 1.5 גפרורים. מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקינים בקופסא?
- (5) חברת בזק הציעה את ההצעה הבאה: שלושים שקלים דמי מנוי חודשיים קבועים וכן 10 אגורות לכל דקה של שיחה יוצאת. אדם בדק במשך שנה את דקות השיחות היוצאות שלו, וקיבל שבממוצע חודשי יש לו 600 דקות שיחות יוצאות עם שונות של 2500 דקות רבועות, כמו כן בחודש ינואר ציון התקן היה 2. חשבו את המדדים הללו עבור חשבון הטלפון החודשי של אותו אדם בשקלים אם היה משתמש בחבילה המוצעת לו על ידי בזק.

תשובות סופיות:

- (1) ממוצע: 315, סטיית תקן: 60, שכיח: 275.
- (2) ממוצע: 46, שונות: 30.25.
- (3) טווח: 40, חציון: 77, עשירון עליון: 91.
- (4) ממוצע: 37, סטיית תקן: 1.5.
- (5) ממוצע: 90, שונות: 25, ציון תקן: 2.

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 10 - סטטיסטיקה תיאורית - מדדי אסימטריה

תוכן העניינים

1. מדד אסימטריה המבוסס על המרחק בין השכיח לממוצע (מקדם פירסון הראשון לצידוד). 30

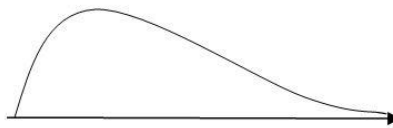
מדד אסימטריה המבוסס על המרחק בין השכיח לממוצע (מקדם פירסון הראשון לצידוד):

רקע:

המטרה היא למדוד עד כמה ההתפלגות היא אסימטרית. צידוד (או באנגלית skewness) הוא מידת האסימטריה של ההתפלגות. המדד שנלמד כאן נקרא מקדם פירסון הראשון לצידוד (Pearson's first coefficient of skewness). מדד זה רלבנטי רק במדידת אסימטריה בהתפלגות חד-שיאית (שכיח אחד) והוא מתבסס על המרחק בין השכיח לממוצע של הנתונים.

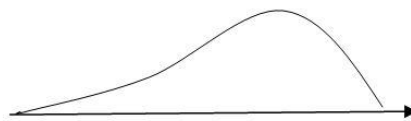
התפלגות אסימטרית חיובית/ימנית:

רוב התצפיות נמצאות בערכים הנמוכים וככל שהערכים גדלים יש פחות ופחות מקרים. בהתפלגות כזו הממוצע גדול מהשכיח.



התפלגות אסימטרית שלילית/שמאלית:

רוב התצפיות נמצאות בערכים הגבוהים וככל שהערכים קטנים יש פחות ופחות מקרים. בהתפלגות כזו הממוצע קטן מהשכיח.



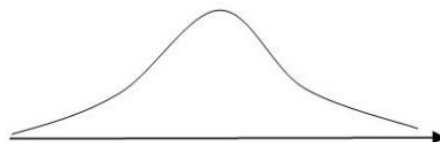
התפלגות סימטרית פעמונית:

מתקיים שרוב התצפיות במרכז ההתפלגות וככל שהערכים מתרחקים מהמרכז יש פחות מקרים באופן סימטרי. בהתפלגות כזו הממוצע שווה לשכיח.

$$S_{K1} = \frac{\bar{X} - MO}{S}$$

המדד מחושב באופן הבא:

החלוקה בסטיית התקן מטרתה לנטרל את היחידות ולהשוות בין התפלגויות שונות.



אם ההתפלגות היא סימטרית פעמונית יתקבל: $S_{K1} = 0$.

אם ההתפלגות היא אסימטרית ימנית יתקבל: $S_{K1} > 0$.

אם ההתפלגות היא אסימטרית שמאלית יתקבל: $S_{K1} < 0$.

ככל שהמדד יותר קרוב בערכו לאפס, נגיד שההתפלגות יותר סימטרית וככל שהמדד מתרחק מהאפס נאמר שההתפלגות היא יותר אסימטרית.

דוגמה (פתרון בהקלטה):

בבניין 10 דירות. ספרו לכל דירה את מספר המחשבים שיש בה. להלן התוצאות שהתקבלו:

מספר דירה	מספר מחשבים
1	5
2	7
3	5
4	3
5	2
6	6
7	0
8	5
9	1
10	4

חשבו את מקדם פירסון הראשון לצידוד.
האם ההתפלגות היא אסימטרית ולאיזה כיוון הצידוד?

שאלות:

- (1) במחקר על 300 נערים ונערות בדקו את מספר המילים שהם מקלידים ביום. להלן התוצאות שהתקבלו:

מספר המילים	מספר הנערים והנערות
0-200	90
200-400	88
400-600	50
600-800	40
800-1000	25
1000-1200	7

- א. מצאו את השכיח והממוצע של הנתונים.
 ב. חשבו את סטיית התקן של הנתונים (השתמשו באמצע מחלקה).
 ג. חשבו את מדד האסימטריה, S_{K1} , ונתחו האם ההתפלגות היא סימטרית או אסימטרית ולאיזה כיוון ההטיה?
 (2) בשכבה שלוש כיתות לימוד. להלן נתונים לגבי התפלגות הציונים בכל כיתה:

הכיתה		
3	2	1
-1	0	0.7
S_{K1}		

- א. דרגו את הכיתות לפי מידת האסימטריה.
 ב. באיזו כיתה רוב הסטודנטים קיבלו ציונים גבוהים יחסית לשאר הכיתה?
 (3) נתון שעבור נתונים מסוימים התקבל: $S_{K1} = 1$. איזה מהמשפטים הבאים נכון?
 א. ההתפלגות היא סימטרית.
 ב. ההתפלגות היא אסימטרית שלילית.
 ג. ההתפלגות היא עם זנב שמאלי.
 ד. ההתפלגות היא עם זנב ימני.

- (4) בהתפלגות מסוימת התקבל שהטווח הוא 0. מה ניתן להגיד על מדד S_{K1} ?

- א. 0
 ב. 1
 ג. 0.5

ד. המדד אינו מוגדר במקרה זה.

- 5) רוצים להשוות בין מדינה A למדינה B מבחינת אסימטריה בשכר. באיזו מדינה קיים אסימטריה גדולה יותר בשכר?
- א. במדינה שבה מדד הצידוד יותר גדול.
 ב. במדינה שבה מדד הצידוד הוא חיובי.
 ג. במדינה שבה ערכו של מדד הצידוד יותר רחוק מהאפס.
 ד. במדינה שבה מדד הצידוד יותר קרוב לערך 0.5.
- 6) בהתפלגות מספר ימי האשפוז במחלקה מסוימת התקבל שהשכיח גדול מהממוצע. מהי התשובה הנכונה בהכרח לגבי מקדם פירסון הראשון לצידוד?
- א. 0.
 ב. חיובי.
 ג. שלילי.
 ד. לא ניתן לדעת.

תשובות סופיות:

- 1) א. ממוצע: 395.3, שכיח: 100. ב. סטיית תקן: 275.4. ג. 1.072.
- 2) א. כיתה 3 < כיתה 2 < כיתה 1. ב. כיתה 3.
- 3) ד'.
 4) ד'.
 5) ג'.
 6) ג'.

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 11 - סטטיסטיקה תיאורית- ניתוח פלטים

תוכן העניינים

1. כללי..... 34

סטטיסטיקה תיאורית – ניתוח פלטים:

שאלות:

1) להלן פלט על התפלגות הגילאים באוכלוסייה מסוימת:

		Statistic
Age of Respondent	Mean	45.63
	Median	41.00
	Variance	317.140
	Std. Deviation	a
	Minimum	18
	Maximum	b
	Range	71
	Interquartile Range	28

- א. מצאו את הערכים בטבלה המסומנים ב- a ו- b .
 ב. נתון שההתפלגות היא אסימטרית, האם היא נוטה ימינה או שמאלה?

2) להלן התפלגות ההשכלה של העובדים בחברת "מתאר":

		Statistic
years of education	Mean	?
	Median	12.0000
	Variance	?
	Std. Deviation	2.54786
	Minimum	?
	Maximum	?
	Range	?
	Interquartile Range	?

מלאו את הערכים המסומנים בסימני שאלה.

תשובות סופיות:

- (1) א. $b = 89$, $a = 17.81$. ב. אסימטרית ימנית.
- (2) ממוצע: 11.909, שונות: 6.492, טווח: 10, טב"ר: 3.

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 12 - סטטיסטיקה תיאורית שאלות אמריקאיות

תוכן העניינים

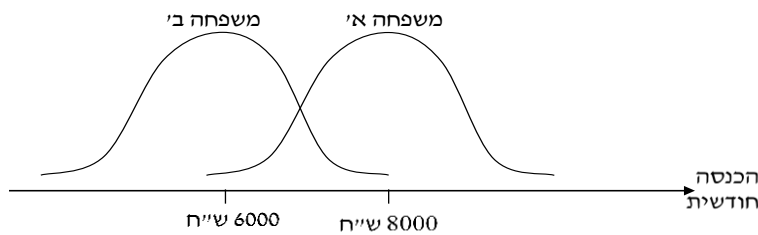
1. כללי..... 36

סטטיסטיקה תיאורית – שאלות אמריקאיות:

שאלות:

שאלות 1-3 מתייחסות לקטע הבא:

להלן שתי עקומות המתארות את התפלגות ההכנסות החודשיות של שתי משפחות שנבחרו באקראי:



- (1) לאיזו משפחה הכנסה שכיחה גבוהה יותר?
- משפחה א׳.
 - משפחה ב׳.
 - לשתיהן אותה הכנסה שכיחה.
 - לא ניתן לדעת – אין מספיק נתונים.
- (2) באיזו משפחה ההכנסה החציונית שווה להכנסה הממוצעת?
- משפחה א׳.
 - משפחה ב׳.
 - בשתיהן ההכנסה החציונית שווה להכנסה הממוצעת.
 - לא ניתן לדעת – אין מספיק נתונים.
- (3) באיזו משפחה סטית התקן של ההכנסה החודשית גבוהה יותר?
- משפחה א׳.
 - משפחה ב׳.
 - לשתיהן אותה סטית תקן.
 - לא ניתן לדעת – אין מספיק נתונים.

הנתונים הבאים מתייחסים לשאלות 4-6:

להלן נתונים חלקיים של טבלת שכיחויות:
 כמו כן, נתון כי הממוצע הוא 1.66.

$F(x)$	x
?	0
10	1
6	2
15	3
?	4
50	סה"כ

(4) השכיח של הנתונים הוא:

- א. 0.
- ב. 15.
- ג. ישנם שני שכיחים: 0 ו-3.
- ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של השכיח.

(5) חציון הנתונים הוא:

- א. 2.
- ב. 1.5.
- ג. 25.5.
- ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של החציון.

(6) הטווח של הנתונים:

- א. 11.
- ב. 3.
- ג. 4.
- ד. על סמך הנתונים החלקיים אי אפשר לקבוע מה יהיה ערכו של החציון.

(7) בהתפלגות אסימטרית ימנית של משתנה כמותי רציף, הערך המתאים למאון ה-30, ציון התקן שלו הוא בהכרח:

- א. שלילי.
- ב. חיובי.
- ג. אפס.
- ד. לא ניתן לדעת ללא הנתונים.

- 8) סדרת נתונים סטטיסטיים מונה 10 תצפיות. נתון כי סדרת הנתונים סימטרית סביב הממוצע. ממוצע הסדרה-40 ושונות הסדרה-100. בשלב מאוחר יותר נוספו שתי תצפיות נוספות לסדרה : 50 ו-30. השונות של 12 התצפיות :
- א. תקטן.
 - ב. תגדל.
 - ג. לא תשתנה.
 - ד. לא ניתן לחשב את השונות ללא ידיעת התצפיות.

הנתונים הבאים מתייחסים לשאלות 9-10:

בחברת "טיק" המשכורת הממוצעת היא 4,600 ₪ וסטיית התקן של משכורת זו הינה 200 ₪. לאחר מו"מ עם ועד עובדי ההנהלה סוכם כי המשכורת תוכפל פי 1.5.

- 9) מהי המשכורת הממוצעת החדשה (ב-₪)?

- א. 2,300.
- ב. 6,900.
- ג. 4,650.
- ד. 4,600.
- ה. חסרים נתונים כדי לדעת.

- 10) מהי סטיית התקן של המשכורת לאחר יישום המו"מ לגבי השכר (ב-₪)?

- א. 200.
- ב. 300.
- ג. 675.
- ד. לא ניתן לדעת.

- 11) הוספת גודל קבוע לכל תצפיות סדרת נתונים :

- א. תגדיל את סטיית התקן.
- ב. תקטין את סטיית התקן.
- ג. לא תשנה את סטיית התקן.
- ד. לא ניתן לדעת.

הנתונים הבאים מתייחסים לשאלות 12-13:

להלן נתונים על ציוני תלמידים שנבחנו במועדים שונים בסטטיסטיקה:

שם התלמיד	ציון	ממוצע הציונים במועד בו נבחן	סטיית התקן של הציונים במועד בו נבחן
צבי	50	50	12
סטף	82	80	5
שרית	65	60	15
לובה	60	63	1.5
מיטב	70	70	10

12 התלמיד הטוב ביותר ביחס לנבחנים באותו מועד בו נבחן הוא:

- א. מיטב.
- ב. צבי.
- ג. לובה.
- ד. שרית.
- ה. סטף.

13 פנינה נבחנה עם סטף וציון התקן שלה שווה לציון התקן של שרית לכן ציונה הוא:

- א. 80.55.
- ב. 65.
- ג. 80.
- ד. 81.66.

הנתונים הבאים מתייחסים לשאלות 14-17:

בבדיקת פתע של משרד הבריאות במפעל שוקולד, נמצא ש:

שוקולד פגום	0	1	2	3	4	5	6	7
מס' קופסאות	35	63	48	12	13	11	10	8

14 מהו החציון של מספר הפגומים בקופסא:

- א. 1.
- ב. 2.
- ג. 4.
- ד. לא ניתן לדעת.

15 מהו הרבעון התחתון של מספר הפגומים בקופסא?

- א. 1.
- ב. 2.
- ג. 3.
- ד. 4.
- ה. לא ניתן לדעת.

16) מספר הפגומים בקופסא הוא משתנה :

- א. סדר.
- ב. שמי.
- ג. כמותי בדיד.
- ד. כמותי רציף.

17) השכיח של מספר הפגומים בקופסא :

- א. 63.
- ב. 1.
- ג. 200.
- ד. לא ניתן לדעת.

18) ביחס לציר המספרים, רוב הערכים בהתפלגות א-סימטרית ימנית נמצאים :

- א. בערכים הגבוהים.
- ב. בחלוקה זהה בין הערכים הגבוהים והנמוכים.
- ג. בערכים הנמוכים.
- ד. לא ניתן לדעת.
- ה. אף לא תשובה מהני"ל נכונה.

19) בוצע מחקר על מספר העובדים בחברות מזון לעומת חברות תקשורת. החציון והממוצע בשתיהן שווה 8.

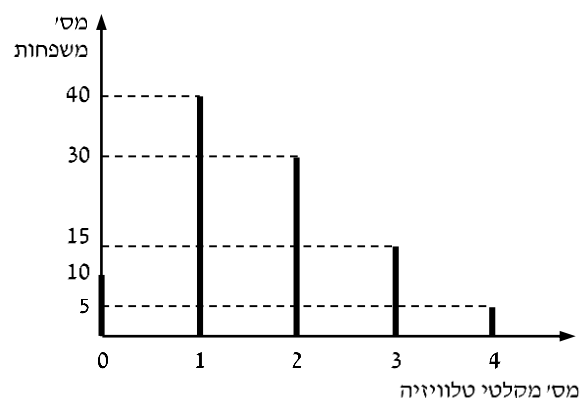
איזה מהטענות הבאות היא הנכונה והמלאה ביותר :

- א. השכיחות ב-2 החברות זהה אך שונה מ-8.
- ב. השכיח ב-2 החברות זהה אך לא ניתן לדעת מהו.
- ג. השכיח בשתי חברות הינו בהכרח 8.
- ד. שכיח בחברה אחת שונה מ-8 ובשנייה הוא 8.
- ה. אף תשובה אינה נכונה.

הנתונים הבאים מתייחסים לשאלות 20 עד 24 :

נערך סקר על מספר מקלטי הטלוויזיה הנמצאים בבית.

תוצאות הסקר נתונות בדיאגרמת מקלות הבאה :



(20) המשתנה הנחקר כאן הוא :

- א. משתנה שמי.
- ב. משתנה מסולם סדר.
- ג. משתנה כמותי בדיד.
- ד. משתנה כמותי רציף.

(21) הטווח של ההתפלגות הוא :

- א. 35.
- ב. 4.
- ג. 3.
- ד. 2.

(22) ממוצע מספר מקלטי הטלוויזיה למשפחה הוא :

- א. 1.65.
- ב. 1.5.
- ג. 1.
- ד. 2.

(23) השכיח של התפלגות זו היא :

- א. 40.
- ב. 1.5.
- ג. 1.
- ד. 2.

(24) מסתבר שיש בין 2 ל-5 משפחות נוספות שאין להם מקלטי טלוויזיה ויש לצרף את המשפחות הללו להתפלגות. כיצד הנתון זה ישפיע על סטיית התקן?

- א. יקטין אותו.
- ב. יגדיל אותו.
- ג. לא ישנה אותו.
- ד. אין לדעת.

תשובות סופיות:

(5 ב'	(4 ג'	(3 ג'	(2 ג'	(1 א'
(10 ב'	(9 ב'	(8 ג'	(7 א'	(6 ג'
(15 א'	(14 ב'	(13 ד'	(12 ה'	(11 ג'
(20 ג'	(19 ה'	(18 ג'	(17 ב'	(16 ג'
	(24 ב'	(23 ג'	(22 א'	(21 ב'

סטטיסטיקה תיאורית מסלול ניצבים א

פרק 13 - התפלגויות רציפות מיוחדות - התפלגות נורמלית

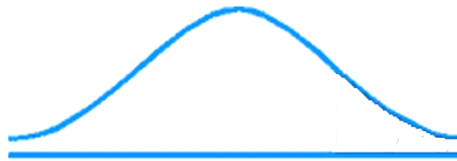
תוכן העניינים

1. כללי.....43

התפלגויות רציפות מיוחדות – התפלגות נורמלית:

רקע:

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



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

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

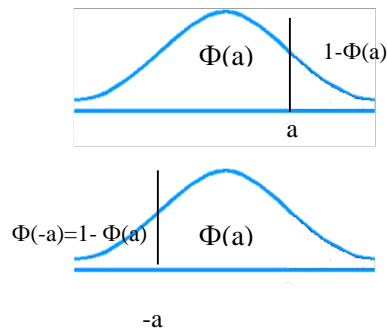
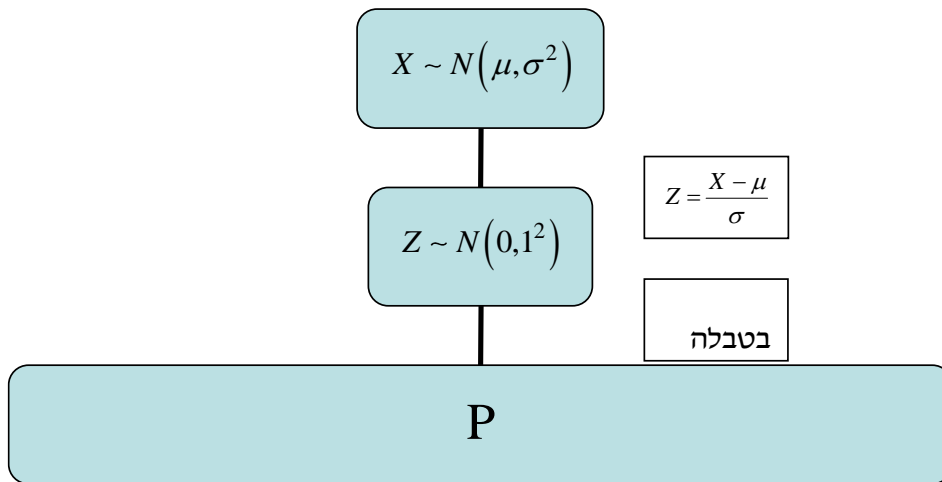
$$\text{נוסחת פונקציית הצפיפות: } f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

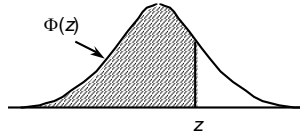
כדי לחשב הסתברויות בהתפלגות נורמלית יש לחשב את השטחים הרלוונטיים שמתחת לעקומה. כדי לחשב שטחים אלה נמיר כל התפלגות נורמלית להתפלגות נורמלית סטנדרטית על ידי תהליך הנקרא תקנון. התפלגות נורמלית סטנדרטית היא התפלגות נורמלית שהממוצע שלה הוא אפס וסטיית התקן היא אחת, והיא תסומן באות Z : $Z \sim N(0, 1^2)$.

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

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

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



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


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

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

דוגמה (הפתרון בהקלטה):

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

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

שאלות:

- (1) הגובה של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 170 ס"מ וסטית תקן של 10 ס"מ.
- מה אחוז האנשים שגובהם מתחת ל-182.4 ס"מ?
 - מה אחוז האנשים שגובהם מעל 190 ס"מ?
 - מה אחוז האנשים שגובהם בדיוק 173.6 ס"מ?
 - מה אחוז האנשים שגובהם מתחת ל-170 ס"מ?
 - 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.
- מהי פרופורציית המקרים בהן התרופה תעזור אחרי יותר משעה?
 - מה אחוז מהמקרים שבהן התרופה תעזור בין 35 ל-37 דקות?
 - מה הסיכוי שהתרופה תעזור בדיוק תוך 36 דקות?
 - 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.
- מה אחוז האנשים שמשקלם נמוך מ-55 ק"ג?
 - מהי פרופורציית האנשים באוכלוסייה שמשקלם לפחות 50 ק"ג?
 - מהי השכיחות היחסית של האנשים באוכלוסייה שמשקלם בין 60 ל-70 ק"ג?
 - לאיזה חלק מהאוכלוסייה משקל הסוטה מהמשקל הממוצע בלא יותר מ-4 ק"ג?
 - 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.
- מצאו את העשירון העליון.
 - מצאו את האחוזון ה-95.
 - מצאו את העשירון התחתון.

- 5) ציוני מבחן אינטליגנציה מתפלגים 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, ונתון ש-33% מהבקבוקים בעלי 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 בקבוק 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 המאיון 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.
- לאיזו התפלגות הממוצע הגבוה ביותר?
 - במה מבין המדדים הבאים התפלגות 1 ו-2 זהות?
 - בעשירון העליון.
 - בממוצע.
 - בשונויות.
 - לאיזו התפלגות סטיית התקן הקטנה ביותר?
 - 1.
 - 2.
 - 3.
 - 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?

תשובות סופיות:

1	א. 89.25%	ב. 2.28%	ג. 0	ד. 50%	ה. 50%
2	א. 0%	ב. 3.76%	ג. 0%	ד. 68.26%	
3	א. 26.43%	ב. 89.44%	ג. 39.44%	ד. 0.383	
	ה. 100%				
4	א. 3812.8	ב. 3958	ג. 2787.2		
5	א. 119.2	ב. 80.8	ג. 112.6	ד. 87.4	ה. 75.3
6	א. 500	ב. 532.9	ג. 453.48		
7	א. 500	ב. 100	ג. 0.3446	ד. 733	
	ה. 267				
8	א. 3	ב. בממוצע.	ג. 1		
9	א. 0.1587	ב. 0.0228	ג. 0.8563	ד. 0.3975	