

מבוא לסטטיסטיקה א

פרק 12 - סטטיסטיקה תיאורית - ממד אסימטריה

תוכן העניינים

1. ממד אסימטריה המבוסס על המרחק בין השכיח לממוצע (מקדם פירסון הראשון לצידוד) ... 1
2. ממד אסימטריה המבוסס על המרחק בין החציוון לממוצע (מקדם פירסון השני לצידוד) ... 5
3. ממד אסימטריה המבוסס על רביעונים 9

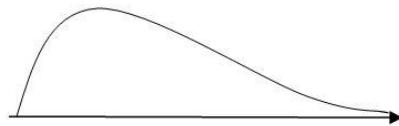
מדד אסימטריה המבוסס על המרחק בין השכיח למומוצע (מקדם פירסון הראשון לצידוד):

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). מזד זה רלבנטי רק במדידת אסימטריה בהתפלגות חד-שיאית (שכיח אחד) והוא מtabס על המרחק בין השכיח למומוצע של הנתונים.

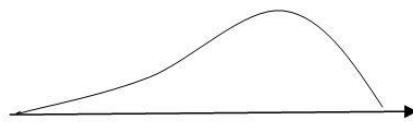
התפלגות אסימטרית חיובית/ימנית:

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



התפלגות אסימטרית שלילית/שמאלית:

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

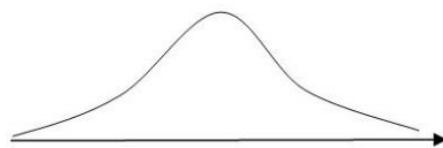


התפלגות סימטרית פעומונית:

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

$$\text{המדד מחושב באופן הבא : } S_{K1} = \frac{\bar{X} - MO}{S}$$

חלוקת בסטיית התקן מטרתה לנטרל את היחידות ולהשווות בין ההתפלגותות שונות.



אם ההתפלגות היא סימטרית פעומנית יתקבל: $S_{K_1} = 0$.

אם ההתפלגות היא אסימטרית ימנית יתקבל: $S_{K_1} > 0$.

אם ההתפלגות היא אסימטרית שמאלית יתקבל: $S_{K_1} < 0$.

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

דוגמה (פתרון בהקלטה):

בבנייה 10 דירות. ספרו לכל דירה את מספר המחשבים שיש בה.

להלן התוצאות שהתקבלו:

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

חשבו את מקדם פירסון הראשון לצידוד.
אם ההתפלגות היא אסימטרית ולאיזה כיוון הצדוד?

שאלות:

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

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

- א. מצאו את השכיח והמומוצע של הנתונים.
- ב. חשבו את סטיית התקן של הנתונים (השתמשו באמצעות מחלוקת).
- ג. חשבו את מדד האסימטריה, S_{K1} , ונתחו האם ההתפלגות היא סימטרית או אסימטרית ולאיזה כיוון ההטייה?

2) בשכבה שלוש כיתות לימוד. להלן נתונים לגבי ההתפלגות הציונית בכל כיתה:

			הכיתה
3	2	1	
-1	0	0.7	S_{K1}

- א. דרגו את הנקודות לפי מידת האסימטריה.
- ב. באיזו כיתה רוב הסטודנטים קיבלו ציונים גבוהים יחסית לשאר הנקודות?

3) נתון שבעבור נתונים מסוימים התקבל: $S_{K1} = 1$. איזה מהמשפטים הבאים נכון?

- א. ההתפלגות היא סימטרית.
- ב. ההתפלגות היא אסימטרית שלילית.
- ג. ההתפלגות היא עם זנב שמאלית.
- ד. ההתפלגות היא עם זנב ימני.

4) בהתפלגות מסוימת התקבל שהטוווח 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 גודלה יותר בשכר?

- א. במדינה שבה מدد הצידוד יותר גדול.
 - ב. במדינה שבה ממד הצידוד הוא חיובי.
 - ג. במדינה שבה ערכו של ממד הצידוד יותר
 - ד. במדינה שבה ממד הצידוד יותר קרוב

6) בהתפלגות מספר ימי האשפוז במחלקה מסוימת התקבל שהשכיח גדול מהමוצע.
מהי התשובה הנכונה בהכרח לגבי מועד פירסון הראשון לצידוד?

- . א. לא ניתן לדעת.
 - . ב. חיובי.
 - . ג. שלילי.

תשובות סופיות:

- .275.4 א. מומוצע : 395.3, שכיח : 100.
ב. סטיית תקן : 1.072.

.3. ב. כיתה 3 < כיתה 2 < כיתה 1.
א. כיתה 3 > כיתה 2 > כיתה 1.

.4. ד'.
ד'.
ג'.
ג'.
ג'.
ג'.

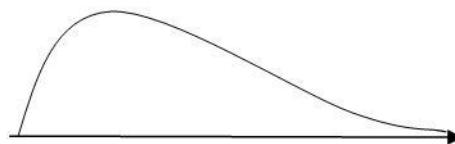
מדד אסימטריה המבוסס על המרחק בין החציון לממוצע (מקדם פירסון השני לצידוד):

רקע:

המטרה היא למדוד עד כמה ההתפלגות היא אסימטרית. צידוד (או באנגלית skewness) הוא מידת האסימטריה של ההתפלגות. המדד שנלמד כאן נקרא מקדם פירסון השני לצידוד (Pearson's second coefficient of skewness). מדד זה מtabסס על המרחק בין החציון לממוצע של הנתונים.

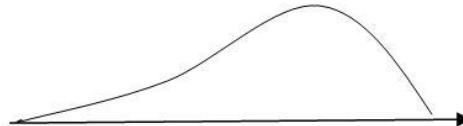
התפלגות אסימטרית חיובית/ימנית:

רוב התצפויות נמצאות בערכים הנמוכים וככל שהערכים גדולים יש פחות ופחות מקרים בההתפלגות כזו הממוצע גדול מהחציון.



התפלגות אסימטרית שלילית/שמאלית:

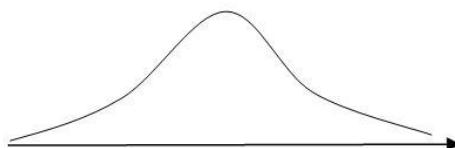
רוב התצפויות נמצאות בערכים הגבוהים וככל שהערכים קטנים יש פחות ופחות מקרים בההתפלגות כזו הממוצע קטן מהחציון.



התפלגות סימטרית:

בההתפלגות כזו הממוצע שווה לחציון. המדד מחושב באמצעות הבא :

$$S_{K2} = \frac{3 \cdot (\bar{X} - Md)}{S}$$



החלוקת בסטיית התקן מטרתה לנטרל את הייחדות ולהשווות בין התפלגויות שונות.

אם ההתפלגות היא סימטרית יתקבל: $S_{K_2} = 0$.

אם ההתפלגות היא אסימטרית ימנית יתקבל: $S_{K_2} > 0$.

אם ההתפלגות היא אסימטרית שמאלית יתקבל: $S_{K_2} < 0$.

כל 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
5	1
7	2
5	3
3	4
2	5
6	6
0	7
5	8
1	9
4	10

חשבו את מקדם פירסון השני לצידוד.

אם ההתפלגות היא אסימטרית ולאיזה כיוון הצדוד?

שאלות:

1) במשרד התיירות מעוניינים לעודד את תיירות הפנים במדינה, ובמיוחד לעוזד יציאה של משפחות לנופשונים קצרים לציררים. על מנת לקבל מושג ראשוני על הרגלי הנופש של משפחות בארץ החליטו, במשרד התיירות, לדגום משפחות ברחבי הארץ ולשאול אותן כמה נופשונים יצאו בשנה שubra. התפלגות מספר הנופשונים למשפחה בשנה נתונה בטבלה הבאה :

מספר נופשונים	שכיחות מצטברת
50	0
90	1
120	2
140	3
150	4

- א. חשבו את הממוצע והחיצון של הנתונים.
- ב. חשבו את סטיית התקן של הנתונים.
- ג. חשבו את מדד האסימטריה S_{K2} ונתנו האם ההתפלגות היא סימטרית או אסימטרית ולאיזה כיוון ההטייה?

2) בשכבה שלוש כיתות לימוד. להלן נתונים לגבי התפלגות הציונים בכל כיתה :

הכיתה			S_{K2}
3	2	1	
-1	0	0.7	

- א. דרגו את הנקודות לפי מידת האסימטריה.
- ב. באיזו כיתה רוב הסטודנטים קיבלו ציונים גבוהים יחסית לשאר הנקודות?

3) נתון שבעבור נתונים מסוימים התקבל: $S_{K2} = -1$.

איזה מהמשפטים הבאים הכí נכון?

- א. ההתפלגות היא סימטרית.
- ב. ההתפלגות היא אסימטרית.
- ג. ההתפלגות היא עם זנב שמאלית.
- ד. ההתפלגות היא עם זנב ימנית.

4) בהתפלגות מסוימת התקבל שהטוווח הוא 0. מה ניתן להגיד על מדד ?skewness

א. 0.

ב. 1.

ג. 0.5

ד. המדד אינו מוגדר במקרה זה.

- 5) ברכוננו להשוות בין מדינה A למדינה B מבחינת אסימטריה של השכר. באיזו מדינה קיים אסימטריה יותר גדולה בשכר?
- במדינה שבה מדד הצידוד יותר גדול.
 - במדינה שבה מדד הצידוד הוא חיובי.
 - במדינה שבה ערכו של מדד הצידוד יותר רחוק מהאפס.
 - במדינה שבה מדד הצידוד יותר קרוב לערך 0.5.
- 6) בהתפלגות מספר ימי האשפוז במחלקה מסוימת התקבל שהחציוון קטן מה ממוצע. מהי התשובה הנכונה בהכרח לגבי מקדם פירסון השני לצידוד?
- 0.
 - חיובי.
 - שלילי.
 - לא ניתן לדעת.

תשובות סופיות:

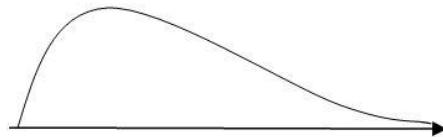
- (1) א. ממוצע : 1.333 , חציוון : 1.085 .
ב. סטיית תקן : 1.118 .
- (2) א. כיתה 3 < כיתה 2 < כיתה 1 .
ב. כיתה 3 .
- (3) ג'.
(4) ד'.
(5) ג'.
(6) ב'.

מדד אסימטריה המבוסס על רביעונים:

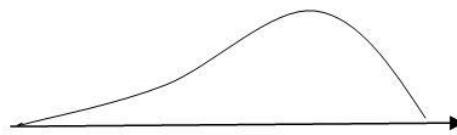
רקע:

המטרה היא למדוד עד כמה ההתפלגות היא אסימטרית על ידי שימוש ברביעונים של ההתפלגות.

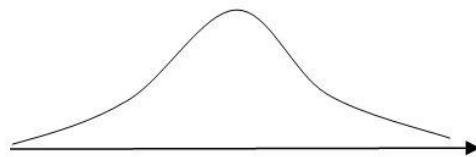
. $(Q_3 - Q_2) > (Q_2 - Q_1)$ בהתפלגות אסימטרית חיובית/ימנית מתקיים :



. $(Q_3 - Q_2) < (Q_2 - Q_1)$ בהתפלגות אסימטרית שלילית/שמאלית מתקיים :



. $(Q_3 - Q_2) = (Q_2 - Q_1)$ בהתפלגות סימטרית מתקיים :



$$\text{מדד אסימטריה}: S_q = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

מדד זה נקרא גם צידוד בולי (Bowley's skewness) או צידוד גלטון (Galton skewness).

המדד מקבל ערכים : $-1 \leq S_q \leq 1$.

המדד בודק את עוצמת האסימטריה ואת כיוון האסימטריה.

העוצמה : באה לידי ביטוי ב- $|S_q|$.

בהתפלגות סימטרית המדד הוא 0 וככל ש- $|S_q|$ קרובה ל-1 ההתפלגות יותר אסימטרית.

כיוון האסימטריה בא לידי ביטוי בסימן של המדד :

בהתפלגות סימטרית : $S_q = 0$.

בהתפלגות היא אסימטרית חיובית (זנב ימני) : $S_q > 0$.

בהתפלגות היא אסימטרית שלילית (זנב שמאל) : $S_q < 0$.

דוגמה (פתרון בהקלטה):

בהתפלגות הציונים בכיתה התקבל : הציון החציוני הוא 75, הרביעון התחתון הוא 65 והרביעון העליון הוא 81.
חשבו את מדד האסימטריה וקבעו את כיוון האסימטריה ועוצמתו.

שאלות:

1) במחקר שנערך נלקחו 300 נערים ונערות ובדקו את מספר המילים שהם מקלידים ביום. להלן התוצאות שהתקבלו :

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

- א. מצאו את הרבעון התחתון והעליון ואת החציון של מספר המילים שהנערים והנערות מקלידים ביום.
- ב. חשבו את מדד האסימטריה. מה ניתן למוד ממנה על האסימטריה של הנתונים?

2) בשכבה שלוש כיתות לימוד. להלן נתוניים לגבי התפלגות הציונים בכל כיתה :

3	2	1	הכיתה רבעונים
77	85	82	עליון
75	80	80	שני
71	75	70	תחתון

- א. דרגו את הנקודות לפי מידת האסימטריה.
- ב. בכיתה אחרת היה החציון כמו התפלגות כיתה מס' 3, הרבעון העליון כמו התפלגות כיתה מס' 1 והוא $S_q = 0.5$. מהו הרבעון התחתון בכיתה זו?

3) נתון שעבור נתוניים מסוימים התקבל: $S_q = 1$.

איזה מהמשפטים הבאים נכון בהכרח?

א. ההתפלגות היא סימטרית.

ב. $Q_3 = Q_2$.

ג. $Q_2 = Q_1$.

ד. $Q_3 = Q_1$.

(4) בהתפלגות מסויימת התקבל שהטוווח הוא 0.
מה ניתן להגיד על מדד ?skewness

- א. 0.
- ב. 1.
- ג. 0.5
- ד. המדד אינו מוגדר במקרה זה.

(5) בהתפלגות מספר ימי האשפוז בחלוקת מסויימת התקבל:
 $Q_2 = Q_3$. מהי התשובה הנכונה לגבי ההתפלגות?

- א. $S_q = 1$
- ב. $S_q = -1$
- ג. $S_q = 0$
- ד. $S_q = 0.5$

תשובות סופיות:

(1) א. $Q_1 = 166 \frac{2}{3}$, $Q_2 = 336.36$, $Q_3 = 588$
ב. 0.195

(2) א. כיתה 2 < כיתה 3 > כיתה 1.
ב. $Q_1 = 72 \frac{2}{3}$

(3) ג'.

(4) ד'.

(5) ב'.