מבוא לסטטיסטיקה והסתברות א

פרק 67 - סטטיסטיקה תיאורית - מדדי אסימטריה

תוכן העניינים

1. מדד צידוד המבוסס על המומנט השלישי של ציוני התקן 1
2. מדד אסימטריה המבוסס על המרחק בין השכיח לממוצע (מקדם פירסון הראשון לצידוד) .. 5
3. מדד אסימטריה המבוסס על המרחק בין החציון לממוצע (מקדם פירסון השני לצידוד) 9
4. מדד אסימטריה המבוסס על רבעונים 13

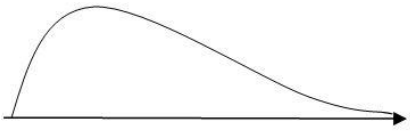
מדד אסימטריה (צידוד) המבוסס על המומנט השלישי של ציוני התקן:

רקע:

המטרה היא למדוד עד כמה ההתפלגות היא אסימטרית.
צידוד (או באנגלית skewness) הוא מידת האסימטריה של ההתפלגות.

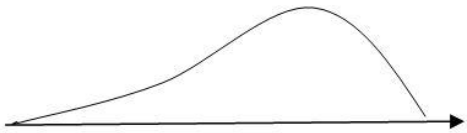
התפלגות אסימטרית חיובית/ימנית:

רוב התצפיות נמצאות בערכים הנמוכים וככל שהערכים גדלים יש פחות ופחות מקרים.



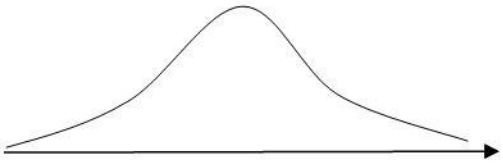
התפלגות אסימטרית שלילית/שמאלית:

רוב התצפיות נמצאות בערכים הגבוהים וככל שהערכים קטנים יש פחות ופחות מקרים.



התפלגות סימטרית פעמונית:

מתקיים שרוב התצפיות במרכז ההתפלגות וככל שהערכים מתרחקים מהמרכז יש פחות מקרים באופן סימטרי.



המדד הבא, נקרא **מדד פישר-פירסון** לאסימטריה.
הוא רלבנטי לבדיקת אסימטריה בהתפלגות חד-שיאית, כלומר עם שכיח אחד, והוא בעצם המומנט השלישי של ציוני התקן.

© כל הזכויות שמורות לברק קנדל

ציון תקן של תצפית מוגדר להיות לפי הנוסחה הבאה: $Z_i = \frac{X_i - \bar{X}}{S}$.

ציון תקן של תצפית ה- i נותנת בכמה סטיות תקן התצפית סוטה מהממוצע.
המומנט השלישי של ציוני התקן הוא בעצם הממוצע של ציוני התקן שהם מועלים בחזקה שלישית. כלומר, המדד הוא: $SKE = \frac{\sum_i Z_i^3}{n}$.

אם ההתפלגות היא סימטרית פעמונית יתקבל: $SKE = 0$.

אם ההתפלגות היא אסימטרית ימנית יתקבל: $SKE > 0$.

אם ההתפלגות היא אסימטרית שמאלית יתקבל: $SKE < 0$.

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 בכל 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 מדד האסימטריה על סמך המומנט השלישי של ציוני התקן. האם ההתפלגות היא אסימטרית ולאיזה כיוון הצידוד?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) במחקר שנערך על 300 נערים ונערות בדקו כמה 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. חשבו את הממוצע וסטיית התקן של ההתפלגות (הסתמכו על אמצע כל מחלקה בחישוב).

ב. חשבו לכל מחלקה את ציון התקן שלה.

ג. חשבו את מדד האסימטריה של פישר-פירסון.

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

הכיתה	1	2	3
SKE	0.7	0	-1

א. דרגו את הכיתות לפי מידת האסימטריה.

ב. באיזו כיתה רוב הסטודנטים קיבלו ציונים גבוהים יחסית לשאר הכיתה?

3) נתון שעבור נתונים מסוימים התקבל: $SKE = 1$. 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 skewness?

א. 0.

ב. 1.

ג. 0.5.

ד. המדד אינו מוגדר במקרה זה.

© כל הזכויות שמורות לברק קנדל

5) בהתפלגות מספר ימי האשפוז במחלקה מסוימת התקבל שממוצע ציוני התקן הוא אפס. מהי התשובה הנכונה בהכרח לגבי ההתפלגות?

א. אסימטרית ימנית.

ב. אסימטרית שמאלית.

ג. סימטרית פעמונית.

ד. אף אחת מהתשובות אינה נכונה בהכרח.

תשובות סופיות:

1) א. ממוצע: 395.3, סטיית תקן: 275.4.

ב. להלן טבלה:

ג. 0.375.

מספר המילים	מספר הנערים והנערות	ציון תקן למחלקה
0-200	90	-1.072
200-400	88	-1.060
400-600	50	0.380
600-800	40	1.106
800-1000	25	1.833
1000-1200	7	2.559

2) א. כיתה 3 > כיתה 2 > כיתה 1.

ב. כיתה 3.

3) ד'.

4) ד'.

5) ד'.

© כל הזכויות שמורות לברק קנדל

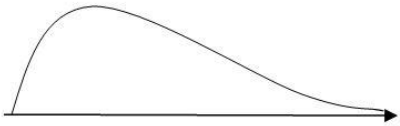
מדד אסימטריה המבוסס על המרחק בין השכיח לממוצע (מקדם פירסון הראשון לצידוד):

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.

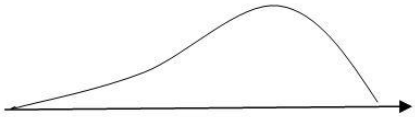
התפלגות אסימטרית חיובית/ימנית:

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.

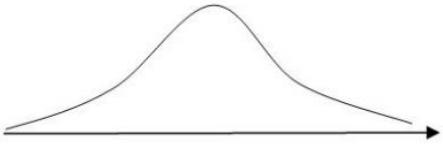


התפלגות סימטרית פעמונית:

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

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

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



© כל הזכויות שמורות לברק קנדל

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?

© כל הזכויות שמורות לברק קנדל

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	1	2	3
S_{K1}	0.7	0	-1

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.

© כל הזכויות שמורות לברק קנדל

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'.

© כל הזכויות שמורות לברק קנדל

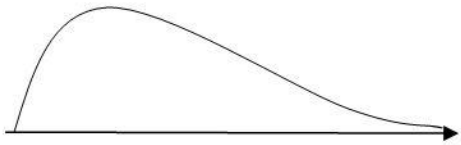
מדד אסימטריה המבוסס על המרחק בין החציון לממוצע (מקדם פירסון השני לצידוד):

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 מתבסס על המרחק בין החציון לממוצע של הנתונים.

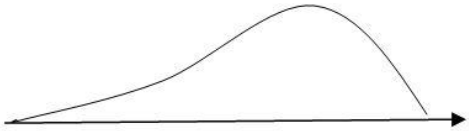
התפלגות אסימטרית חיובית/ימנית:

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



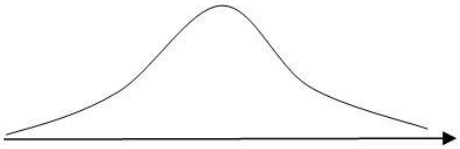
התפלגות אסימטרית שלילית/שמאלית:

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



התפלגות סימטרית:

בהתפלגות כזו הממוצע שווה לחציון. המדד מחושב באופן הבא: $S_{K2} = \frac{3\cdot\left(\bar{X} - Md\right)}{S}$.



© כל הזכויות שמורות לברק קנדל

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

אם ההתפלגות היא סימטרית יתקבל: $S_{K2}=0$.

אם ההתפלגות היא אסימטרית ימנית יתקבל: $S_{K2}>0$.

אם ההתפלגות היא אסימטרית שמאלית יתקבל: $S_{K2}<0$.

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 השני לצידוד.
האם ההתפלגות היא אסימטרית ולאיזה כיוון הצידוד?

© כל הזכויות שמורות לברק קנדל

שאלות:

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

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

א. חשבו את הממוצע והחציון של הנתונים.
ב. חשבו את סטיית התקן של הנתונים.
ג. חשבו את מדד האסימטריה S_{K2} ונתחו האם ההתפלגות היא סימטרית או אסימטרית ולאיזה כיוון ההטיה?

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

הכיתה	1	2	3
S_{K2}	0.7	0	-1

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

3) נתון שעבור נתונים מסוימים התקבל: $S_{K2} = -1$.
איזה מהמשפטים הבאים הכי נכון?
א. ההתפלגות היא סימטרית.
ב. ההתפלגות היא אסימטרית.
ג. ההתפלגות היא עם זנב שמאלי.
ד. ההתפלגות היא עם זנב ימני.

4) בהתפלגות מסוימת התקבל שהטווח הוא 0. מה ניתן להגיד על מדד skewness?
א. 0.
ב. 1.
ג. 0.5.
ד. המדד אינו מוגדר במקרה זה.

© כל הזכויות שמורות לברק קנדל

5) ברצוננו להשוות בין מדינה A למדינה B מבחינת אסימטריה של השכר.
באיזו מדינה קיים אסימטריה יותר גדולה 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 שהחציון קטן מהממוצע. מהי התשובה הנכונה בהכרח לגבי מקדם פירסון השני לצידוד?

א. 0.
ב. חיובי.
ג. שלילי.
ד. לא ניתן לדעת.

תשובות סופיות:

1) א. ממוצע: 1.333, חציון: 1. ב. סטיית תקן: 0.85.
ג. 118.

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

3) ג'.

4) ד'.

5) ג'.

6) ב'.

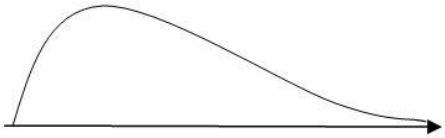
© כל הזכויות שמורות לברק קנדל

מדד אסימטריה המבוסס על רבעונים:

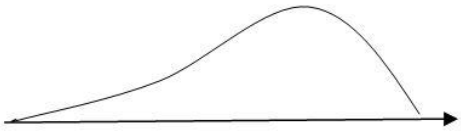
רקע:

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

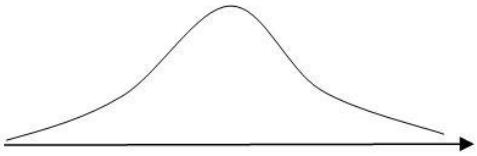
בהתפלגות אסימטרית חיובית/ימנית מתקיים: $(Q_3 - Q_2) > (Q_2 - Q_1)$.



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



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



נגדיר את המדד הבא לאסימטריה: $S_q = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$.

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

המדד מקבל ערכים: $-1 \le S_q \le 1$.

המדד בודק את עוצמת האסימטריה ואת כיוון האסימטריה.

העוצמה: באה לידי ביטוי ב-$|S_q|$.

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

© כל הזכויות שמורות לברק קנדל

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 שמאלי): $S_q < 0$.

דוגמה (פתרון בהקלטה):

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

© כל הזכויות שמורות לברק קנדל

שאלות:

1) במחקר שנערך נלקחו 300 נערים ונערות ובדקו 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 ומעלה	7

א. מצאו את הרבעון התחתון והעליון ואת החציון של מספר המילים שהנערים והנערות מקלידים ביום.

ב. חשבו את מדד האסימטריה. מה ניתן ללמוד ממנו על האסימטריה של הנתונים?

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

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

א. דרגו את הכיתות לפי מידת האסימטריה.

ב. בכיתה אחרת היה החציון כמו התפלגות כיתה מספר 3, הרבעון העליון כמו התפלגות כיתה מספר 1 ו-$S_q = 0.5$.
מהו הרבעון התחתון בכיתה זו?

3) נתון שעבור נתונים 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. $Q_3 = Q_1$.

© כל הזכויות שמורות לברק קנדל

4) בהתפלגות מסוימת התקבל שהטווח 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 במחלקה מסוימת 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'.

© כל הזכויות שמורות לברק קנדל