

# הסקה סטטיסטית לכלכלנים

פרק 18 - אומדן חסר הטיה וקרטריון mse

תוכן העניינים

- 1. אומדן חסר הטיה ..... 1
- 2. MSE ..... 7

## אומדן חסר הטיה:

רקע:

$\hat{\theta}$  יהיה אומדן חסר הטיה ל- $\theta$ , אם התוחלת של  $\hat{\theta}$  תהיה שווה ל- $\theta$ :  $E(\hat{\theta}) = \theta$ .

דוגמה (פתרון בהקלטה):

המשתנה  $X$  הוא בעל פונקציית ההסתברות הבאה:

3	2	1	$X$
$4\theta$	$1 - 60\theta$	$2\theta$	הסתברות

מעוניינים לאמוד את  $\theta$  על סמך שתי תצפיות מההתפלגות:  $X_1$  ו- $X_2$ .

א. הראו שהאומדן:  $T_1 = \frac{2X_1 + X_2}{2}$ , הוא אומדן מוטה ל- $\theta$ .

הטיה של אומדן היא:  $E(\hat{\theta}) - \theta$ . כמובן שלאומדן חסר הטיה אין הטיה.

ב. מהי ההטיה של האומדן  $T_1$ ?

ג. תקנו את  $T_1$ , כך שיהיה אומדן חסר הטיה.

אם יש שני אומדנים חסרי הטיה עדיף זה עם השונות היותר קטנה.

ד. מוצא האומדן הבא:  $T_3 = 1.5X_1 - X_2 - 1$ .

האם הוא עדיף על האומדן שהצעת בסעיף ג'?

אם  $\hat{\theta}$  אומדן חסר הטיה ל- $\theta$ , אז  $g(\hat{\theta})$  יהיה אומדן חסר הטיה עבור  $g(\theta)$ , רק אם  $g$  תהיה לינארית.

ה. מצאו אומדן חסר הטיה ל:  $P(X = 3)$ .

אומדן חסר הטיה לשונות האוכלוסייה  $\sigma^2$ :  $S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$ .

ו. מצאו אומדן חסר הטיה לשונות של  $X$ .

**תזכורות חשובות:**

אם:  $Y = aX + b$ , אזי:  $E(Y) = aE(X) + b$ ,  $V(Y) = a^2 \cdot V(X)$ ,  $\sigma_Y = |a| \sigma_X$ .

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

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

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

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

## שאלות:

- (1) הציון במבחן מסוים של תלמידי כתה ח' הנו משתנה מקרי בעל תוחלת  $\mu$  וסטיית תקן 10. כדי לאמוד את התוחלת  $\mu$ , נלקח מדגם של 5 ציונים:  $X_1, \dots, X_5$ . שלושה חוקרים הציעו אומדים לתוחלת על סמך מדגם זה:

$$T_1 = \frac{X_1 + \dots + X_5}{5} \quad \text{חוקר א' הציע:}$$

$$T_2 = \frac{2X_1 - X_3 + X_4}{2} \quad \text{חוקר ב' הציע:}$$

$$T_3 = \frac{2X_1 + X_3}{2} \quad \text{חוקר ג' הציע:}$$

- איזה מן האומדים הוא חסר הטיה?
- הציעו תיקון לאומד המוטה כך שיהיה חסר הטיה.
- במדגם התקבלו הציונים הבאים: 100, 82, 58, 78, 65. חשבו את האומדנים המתקבלים עבור האומדים חסרי הטיה.
- איזה מבין שני האומדים חסרי הטיה עדיף? נמקו.

- (2) כדי לאמוד את המשקל הממוצע של הנשים בארה"ב, נבחר מדגם של  $2n$  נשים. נסמן את שונות הגובה ב- $\sigma^2$ . הוצעו שני אומדים לממוצע המשקל על סמך מדגם

$$T_1 = \frac{1}{n} \sum_{i=1}^n X_i, T_2 = \frac{1}{2n} \sum_{i=1}^{2n} X_i \quad \text{זה:}$$

- בדקו לגבי כל אומד אם הוא בלתי מוטה.
- איזה אומד עדיף? נמקו.

- (3)  $X \sim B(n, p)$ . כלומר,  $X$  הינו משתנה מקרי המתפלג בינומית עם פרמטר  $P$  (סיכוי להצלחה בניסיון בודד) במדגם בגודל  $n$ .

- פתחו אומד חסר הטיה ל- $P$ .
- מהו אומד חסר הטיה לסיכוי לכישלון בניסיון בודד?
- מהו אומד חסר הטיה ל- $E(X)$ ?
- מצאו אומד חסר הטיה ל- $E(X^2)$ .

(4) בתיק מניות שתי מניות. מספר המניות שיעלו ביום מסוים הוא משתנה מקרי התלוי בפרמטר לא ידוע:  $\theta$ ,  $0 \leq \theta \leq 2$ .

פונקציית ההסתברות של  $X$  - מספר המניות שיעלו ביום מסוים:

$$P(X=0) = 1 - \frac{\theta}{2}, \quad P(X=1) = \frac{\theta}{3}, \quad P(X=2) = \frac{\theta}{6}$$

א. מצאו אומד בלתי מוטה ל- $\theta$ , שמתבסס על מספר המניות שיעלו ביום מסוים.

ב. מצאו אומד בלתי מוטה ל- $\theta$ , שמתבסס על מספר המניות שעלו ביום,

במשך שלושה ימים -  $X_1, X_2, X_3$  (לכל אחד מהם אותה התפלגות כנ"ל

והם בלתי תלויים).

(5) בקרב המטפלות בת"א, מספר התינוקות שבטיפולן הוא משתנה מקרי בעל

התפלגות התלוייה בפרמטר  $\theta$  באופן הבא:

הסיכוי שמטפלת תטפל בתינוק אחד בלבד הוא  $3\theta$ ,

הסיכוי שמטפלת תטפל ב-2 תינוקות הוא  $1 - 4\theta$ ,

הסיכוי שמטפלת תטפל ב-3 תינוקות הוא  $\theta$ .

במדגם מיקרי של 4 מטפלות מת"א, נמצא כי שתיים מהם מטפלות בתינוק

אחד בלבד, אחת מהן בשנים ואחת השלושה תינוקות.

א. מצאו אומד חסר הטיה לפרמטר  $\theta$  על סמך תצפית בודדת.

ב. מצאו אומד חסר הטיה לפרמטר  $\theta$  על סמך 4 תצפיות.

ג. מהו האומדן לפרמטר  $\theta$  על סמך תוצאות המדגם.

ד. מצאו אומד חסר הטיה לסיכוי שלמטפלת בת"א תטפל בתינוק בודד אחד.

ה. מצאו אומדים חסרי הטיה לתוחלת ולשונות של מספר התינוקות בטיפול

אצל מטפלת מת"א. חשבו אומדנים.

(6) קבעו אילו מהטענות הבאות נכונות:

א. אם  $T$  הוא אומד בלתי מוטה עבור פרמטר  $\theta$ , אז  $5T$  אומד בלתי מוטה

עבור הפרמטר  $5\theta$ .

ב. אם  $T$  הוא אומד בלתי מוטה עבור פרמטר  $\theta$ , אז  $T^2$  אומד בלתי מוטה

עבור הפרמטר  $\theta^2$ .

(7) במפעל שתי מכונות המייצרות מוצרים. במכונה הראשונה ההסתברות שמכשיר תקין היא  $p$ , ובמכונה השנייה ההסתברות שמכשיר תקין היא  $2p$ . דוגמים 20 מכשירים מהייצור של כל מכונה. נסמן ב-  $X$  את מספר המכשירים התקינים שיוצרו על ידי המכונה הראשונה, וב-  $Y$  את מספר המכשירים התקינים שיוצרו על ידי המכונה השנייה. איזה מבין האומדים הבאים אינו אומד חסר הטיה ל-  $p$ ?

א.  $\frac{X}{20}$ .

ב.  $\frac{Y}{20}$ .

ג.  $\frac{X+Y}{60}$ .

ד.  $\frac{2X+Y}{80}$ .

(8) יהיו  $T_1$  ו-  $T_2$  אומדים חסרי הטיה ובלתי תלויים לפרמטר  $\theta$ .  
 א. מצאו אומד חסר הטיה ל-  $\theta^2$ , המתבסס על  $T_1$  ו-  $T_2$ .  
 ב. מצאו אומד חסר הטיה ל-  $\theta(1-\theta)$ , המתבסס על  $T_1$  ו-  $T_2$ .

(9) נתון ש-  $X$  הינו משתנה מקרי עם תוחלת  $\mu$  ושוונות  $\sigma^2$ . נדגמו  $n$  תצפיות בלתי תלויים מאותה אוכלוסיה.

א. הראו ש-  $\sum_{i=1}^n p_i x_i$  אומד חסר הטיה ל-  $\mu$ , כאשר:  $\sum_{i=1}^n p_i = 1$ .

ב. נתבונן במכפלת שתי התצפיות הראשונות:  $X_1 \cdot X_2$ .

הראו שהוא אומד חסרי הטיה ל-  $\mu^2$ .

## תשובות סופיות:

$$(1) \quad \text{א. } T_1 \text{ ו- } T_2 \quad \text{ב. } \frac{2}{3} T_3 \quad \text{ג. } T_1 = 76.6, T_2 = 110 \quad \text{ד. } T_1$$

$$(2) \quad \text{א. ראו בוידאו.} \quad \text{ב. } T_2$$

$$(3) \quad \text{א. } \frac{x}{n} \quad \text{ב. } 1 - \frac{x}{n} \quad \text{ג. } X \quad \text{ד. } \theta$$

$$(4) \quad \text{א. } \frac{3x}{2} \quad \text{ב. } \frac{3\bar{x}}{2}$$

$$(5) \quad \text{א. } 1 - \frac{x}{2} \quad \text{ב. } 1 - \frac{1}{2} \bar{x} \quad \text{ג. } 0.125 \quad \text{ד. } 3 \left( 1 - \frac{1}{2} \bar{x} \right)$$

ה. לשונות 0.917.

$$(6) \quad \text{א. נכון.} \quad \text{ב. לא נכון.}$$

(7) ב'.

$$(8) \quad \text{א. } T_1 \cdot T_2 \quad \text{ב. } T_1 - T_1 \cdot T_2$$

$$(9) \quad \text{א. שאלת הוכחה.} \quad \text{ב. שאלת הוכחה.}$$

## קריטריון MSE – תוחלת ריבוע הטעות:

**רקע:**

הקריטריון הנפוץ ביותר כדי לבדוק את טיב האומד הוא קריטריון MSE: Mean Squared Error – תוחלת ריבוע טעות האמידה.

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = V(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2$$

כאשר:  $V(\hat{\theta})$  – הינה שונות האומד.

$E(\hat{\theta}) - \theta$  – הינה ההטויה של האומד.

אם  $T_1$  ו- $T_2$  הינם אומדים לפרמטר  $\theta$ , האומד העדיף יהיה זה עם MSE קטן יותר.

כלומר, אם:  $MSE(T_1) > MSE(T_2)$ , אז  $T_2$  עדיף על  $T_1$ .

**דוגמה (הפתרון בהקלטה):**

נתון משתנה  $X$  המתפלג אחיד רציף באופן הבא:  $X \sim U(3, \theta)$ .

מוצעים שני אומדים לפרמטר  $\theta$  על סמך תצפית בודדת:  $T_1 = 2X - 3$  ו- $T_2 = \frac{3X - 3}{2}$ .

איזה אומד עדיף לאמידת הפרמטר  $\theta$ ?



## שאלות:

(1) מעוניינים לאמוד את התוחלת של התפלגות מסוימת. מוצעים שני אומדים אפשריים ממוצע של שתי תצפיות וממוצע של שלוש תצפיות. לפי קריטריון תוחלת ריבוע הטעות (MSE), איזה אומדן עדיף? הסבירו.

(2) בעיר מסוימת בשוויץ בכל  $\theta$  דקות רכבת מגיעה לתחנה מסוימת. דוד מגיע לתחנה בזמן אקראי ומוודד את זמן ההמתנה לרכבת -  $X$ .  
 א. הצע אומדן חסר הטיה ל- $\theta$ , על סמך  $X$ .  
 ב. סטטיסטיקאי הציע לאמוד את  $\theta$  על סמך האומדן:  $1.5X$ . האם האומדן הנ"ל מוטה?  
 ג. איזה אומדן מבין האומדים בסעיפים א' ו-ב' עדיף?

(3) חוקר מעוניין לאמוד את הסיכוי לחלות במחלת השפעת בחורף (להלן: הפרמטר  $P$ ). הוא דוגם חמישה אנשים בריאים, ומתבונן בסטטיסטי  $X$  - מספר האנשים שחלו בשפעת בחורף. הוא מתלבט בין שני אומדים:  $T_1 = \frac{X}{5}$  ו-  $T_2 = \frac{X+1}{7}$ .  
 א. מי מבין האומדים הללו הוא חסר הטיה?  
 ב. מי מבין האומדים עדיף אם  $P = 0.5$ ?  
 ג. מי מבין האומדים עדיף אם  $P = 0.1$ ?

(4) מספר השריפות המתרחשות בארץ בחודש אוקטובר מתפלג פואסונית עם תוחלת  $\lambda$ . נלקח מדגם של 10 חודשי אוקטובר. להלן שני אומדים אפשריים:

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^{10} X_i}{10} \quad \text{ו-} \quad \hat{\lambda}_2 = \frac{2 \cdot \sum_{i=1}^5 X_i + 2 \cdot \sum_{i=6}^{10} X_i}{10}$$

כאשר:  $X_i$  = מספר השריפות בחודש אוקטובר ה- $i$ .  
 איזה מהאומדים עדיף, לצורך אמידת הפרמטר  $\lambda$ ?

(5) הוכח ש:  $E(\hat{\theta} - \theta)^2 = V(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2$ .

**תשובות סופיות:**

- (1) שלוש תצפיות.
- (2) א.  $2x$ . ב. אומד מוטה. ג. סעיף ב.
- (3) א.  $T_1$ . ב.  $T_2$ . ג.  $T_1$ .
- (4)  $\hat{\lambda}_1$ .
- (5) שאלת הוכחה.