

# מבוא לסטטיסטיקה והסתברות

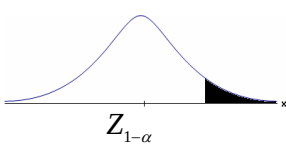
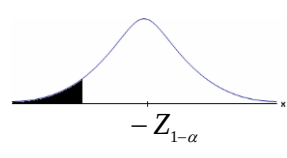
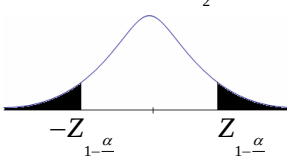
פרק 47 - בדיקת השערות על פרופורציה

תוכן העניינים

1. התהליך.....1
2. סיכוי לטעויות ועוצמה.....4

## התהליך:

רקע:

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

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

חלופה אחרת לכלל הכרעה:

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

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.

## שאלות:

- (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. בדקו 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  $\alpha_1$  וחוקר ב' ברמת מובהקות  $\alpha_2$  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.

## תשובות סופיות:

- (1) נדחה  $H_0$ .
- (2) נדחה  $H_0$ .
- (3) נקבל  $H_0$ .
- (4) א.  $H_0: p = 0.2$   
 ב.  $H_1: p \neq 0.2$   
 ג. המסקנה לא תשתנה.
- (5) א. נדחה  $H_0$ .  
 ב. המסקנה לא תשתנה.
- (6) ג'.

## סיכוי לטעויות ועוצמה:

**רקע:**

**הגדרת הסתברויות:**

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):

$$\alpha = P(H_0 \text{ נדחתה} \mid H_0 \text{ נכונה}) = P_{H_0}(H_0 \text{ נדחתה})$$

הסיכוי לבצע טעות מסוג 2:

$$\beta = P(H_0 \text{ נקבל} \mid H_1 \text{ נכונה}) = P_{H_1}(H_0 \text{ נקבל})$$

**רמת בטחון:**

$$(1-\alpha) = P(H_0 \text{ נקבל} \mid H_0 \text{ נכונה}) = P_{H_0}(H_0 \text{ נקבל})$$

**עוצמה:**

$$\pi = (1-\beta) = P(H_1 \text{ נדחתה} \mid H_1 \text{ נכונה}) = P_{H_1}(H_1 \text{ נדחתה})$$

| מציאות | הכרעה |             |          |
|--------|-------|-------------|----------|
|        |       | $H_0$       | $H_1$    |
|        | $H_0$ | טעות מסוג 1 | אין טעות |
|        | $H_1$ | טעות מסוג 2 | אין טעות |

**התהליך לחישוב סיכוי לטעות מסוג שני:**

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

| חישוב $\beta$ :                                                                                                                                                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $P_{H_1} \left( \hat{p} < p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} \right)$                                                                            |
| $P_{H_1} \left( p_0 - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} < \hat{p} < p_0 + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} \right)$ |
| $P_{H_1} \left( \hat{p} > p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}} \right)$                                                                            |

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

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.

- רשמו את ההשערות וכלל הכרעה ברמת מובהקות של 10%.
- מהי עוצמת המבחן אם מסתבר ש 60% מהאוכלוסייה אינם מבקרים אצל רופא שיניים באופן קבוע.

## שאלות:

- (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.
  - 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 (בחרו בתשובה הנכונה):
  - א. השערת האפס נכונה.
  - ב. השערת האפס נדחתה.
  - ג. השערת האפס לא נדחתה.
  - ד. אף אחת מהתשובות לא נכונה בהכרח.
- (6) קבעו אם הטענה הבאה נכונה: בבדיקת השערות לא ניתן לבצע בו זמנית טעות מסוג ראשון וטעות מסוג שני.

## תשובות סופיות:

- (1) א.  $H_0 : p = 0.1$   
 ב. 0.9015 ג. תגדל. ד. טענה לא נכונה.  
 $H_1 : p > 0.1$
- (2) 0.4404
- (3) א. 0.1446 ב. תקטן.
- (4) חוקר א'.
- (5) ג'.
- (6) נכונה.