

# סטטיסטיקה לתואר שני במדרשת פיינברג

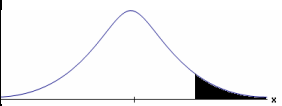
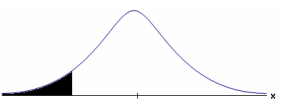
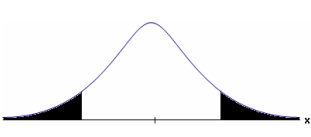
פרק 42 - בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

תוכן העניינים

1. כששונויות האוכלוסייה אינן ידועות והמדגמים גדולים.....1
2. כששונויות האוכלוסייה לא ידועות ומניחים שהן שוות.....4

## בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

כשהשונות של האוכלוסייה לא ידועות והמדגמים גדולים – רקע

| השערת האפס:<br>השערה אלטרנטיבית:                                                                                                                                     | השערת האפס:<br>השערה אלטרנטיבית:                                                                                                                                       | השערת האפס:<br>השערה אלטרנטיבית:                                                                                                                                                                                                                                                | השערת האפס:<br>השערה אלטרנטיבית:                                                               |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|
| $H_0 \quad \mu_1 - \mu_2 = c$<br>$H_1 \quad \mu_1 - \mu_2 > c$                                                                                                       | $H_0 \quad \mu_1 - \mu_2 = c$<br>$H_1 \quad \mu_1 - \mu_2 < c$                                                                                                         | $H_0 \quad \mu_1 - \mu_2 = c$<br>$H_1 \quad \mu_1 - \mu_2 \neq c$                                                                                                                                                                                                               | תנאים:<br>1. מדגמים בלתי תלויים<br>2. $\sigma_1, \sigma_2$ לא ידועות<br>3. מדגמים מספיק גדולים |
| $Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\alpha}$<br><br>$Z_{1-\alpha}$<br>דוחים את $H_0$ | $Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\alpha}$<br><br>$-Z_{1-\alpha}$<br>דוחים את $H_0$ | או $Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\frac{\alpha}{2}}$<br>$Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\frac{\alpha}{2}}$<br><br>$-Z_{1-\frac{\alpha}{2}}, Z_{1-\frac{\alpha}{2}}$<br>דוחים את $H_0$ | כלל ההכרעה:<br>אזור הדחייה של $H_0$ :                                                          |

סטטיסטי המבחן: 
$$Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - c}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

## חלופה אחרת לכלל הכרעה:

| נדחה $H_0$ אם מתקיים:                                                                         |                                                                                                                                                                                                          |
|-----------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ | $\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$<br>או<br>$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ |
| $\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ |                                                                                                                                                                                                          |

## דוגמה (פתרון בהקלטה):

נרצה לבדוק האם קיים הבדל בין ממוצע ציוני הפסיכומטרי של חיילים לממוצע ציוני הפסיכומטרי של תלמידי תיכון. במדגם של 46 נבחנים חיילים התקבל ממוצע 543 וסטיית תקן 123. במדגם של 50 תלמידי תיכון התקבל ממוצע 508 וסטיית תקן 178. מה המסקנה ברמת מובהקות 5%?

## שאלות

(1) חברה להנדסת בניין מעוניינת להשוות ברמת הקשיות של שני סוגי ברגים. במדגם של 35 ברגים מסוג א' התקבל רמת קשיות ממוצעת של 28 יחידות וסטיית תקן 4, ובמדגם של 45 ברגים מסוג ב' התקבל רמת קשיות ממוצעת של 25 וסטיית תקן 6. האם על סמך תוצאות המדגם יש הבדל בין סוגי הברגים מבחינת רמת הקשיות שלהם? בדקו ברמת מובהקות של 5%.

(2) כדי לבדוק האם נהגים השותים תחת השפעת אלכוהול נוהגים מהר יותר מאלו שאינם שותים בוצע מדגם שבו בדקו את המהירות המקסימאלית של כל נהג בקמ"ש. להלן התוצאות:

| S  | $\bar{X}$ | גדול מדגם |                           |
|----|-----------|-----------|---------------------------|
| 20 | 80        | 70        | נהגים השותים אלכוהול      |
| 15 | 60        | 100       | נהגים שאינם שותים אלכוהול |

א. מהי מובהקות התוצאה?

ב. מה המסקנה ברמת מובהקות של 5%?

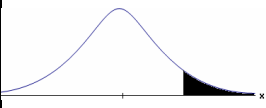
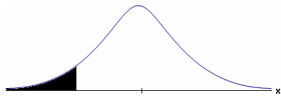
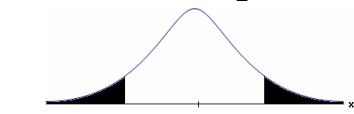
## תשובות סופיות

(1) נדחה את  $H_0$ .

(2) א. 0 ב. נדחה את  $H_0$ .

## בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

כששונות האוכלוסייה לא ידועות ומניחים שהן שוות – רקע

| השערת האפס:<br>השערה אלטרנטיבית:<br>תנאים:                                                                     | $H_0 \quad \mu_1 - \mu_2 = c$<br>$H_1 \quad \mu_1 - \mu_2 > c$                                                                                                                                 | $H_0 \quad \mu_1 - \mu_2 = c$<br>$H_1 \quad \mu_1 - \mu_2 < c$                                                                                                                                  | $H_0 \quad \mu_1 - \mu_2 = c$<br>$H_1 \quad \mu_1 - \mu_2 \neq c$                                                                                                                                                                                                                                                               |
|----------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. מדגמים בלתי תלויים<br>2. $\sigma_1, \sigma_2$ לא ידועות אך שוות<br>3. המשתנים בכל אוכלוסייה מתפלגים נורמלית | $t_{\bar{x}_1 - \bar{x}_2} > t_{1-\alpha}^{(n_1+n_2-2)}$<br><br>$t_{1-\alpha, n_1+n_2-2}$<br>דוחים את $H_0$ | $t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\alpha}^{(n_1+n_2-2)}$<br><br>$-t_{1-\alpha, n_1+n_2-2}$<br>דוחים את $H_0$ | $t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$ או $t_{\bar{x}_1 - \bar{x}_2} > t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$<br><br>$-t_{1-\frac{\alpha}{2}, n_1+n_2-2}$ $t_{1-\frac{\alpha}{2}, n_1+n_2-2}$<br>דוחים את $H_0$ |

$$t_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

סטטיסטי המבחן:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

השונות המשוקללת:

## חלופה אחרת לכלל הכרעה:

| נדחה $H_0$ אם מתקיים:                                                                                       |                                                                                                                                                                                                                                                      |
|-------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\bar{x}_1 - \bar{x}_2 < c - t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$ | $\bar{x}_1 - \bar{x}_2 > c + t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$<br>או<br>$\bar{x}_1 - \bar{x}_2 < c - t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$ |
| $\bar{x}_1 - \bar{x}_2 > c + t_{1-\alpha}^{(n_1+n_2-2)} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$ |                                                                                                                                                                                                                                                      |

## דוגמה (פתרון בהקלטה):

חברה המייצרת מוצרי בנייה טוענת שפיתחה סגסוגת (תערובת מתכות) שטמפרטורת ההתכה שלה גבוהה משמעותית מטמפרטורת ההתכה של הסגסוגת לבנייה שמשתמשים בה כיום לבניית בניינים. לצורך בדיקת טענת המחקר נדגמו 10 יחידות של מתכות מהסוג הישן ו-12 יחידות של מתכות מהסוג החדש. להלן תוצאות המדגם:

טמפרטורת ההתכה הממוצעת במתכת הישנה 1170 מעלות עם אומד חסר הטיה לשונות  $S^2 = 200$ .

טמפרטורת ההתכה הממוצעת במתכת החדשה 1317 מעלות עם אומד חסר הטיה לשונות  $S^2 = 260$ .

נניח לצורך פתרון שטמפרטורת ההתכה מתפלגת נורמאלית עם אותה שונות במתכות השונות. בדקו ברמת מובהקות של 5%.

## שאלות

- (1) להלן נתונים של שטחי דירות מתוך דירות שנבנו בשנת 2012 ובשנת 2013 (במ"ר):

|     |    |    |     |    |     |     |      |
|-----|----|----|-----|----|-----|-----|------|
| 120 | 94 | 90 | 130 | 95 | 112 | 120 | 2012 |
|     | 69 | 74 | 105 | 91 | 82  | 100 | 2013 |

בדקו שבשנת 2013 הייתה ירידה משמעותית בשטחי הדירות לעומת שנת 2012 עבור רמת מובהקות של 5%.  
הניחו ששטחי הדירות בכל שנה מתפלגים נורמלית עם אותה שונות.

- (2) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו למבחן IQ. להלן תוצאות המדגם:

| המדינה              | ישראל   | ארה"ב   |
|---------------------|---------|---------|
| גודל המדגם          | 15      | 15      |
| סכום הציונים        | 1560    | 1470    |
| סכום ריבועי הציונים | 165,390 | 147,560 |

בדקו ברמת מובהקות של 5% האם קיים הבדל של נקודה בין ישראלים לאמריקאים מבחינת ממוצע הציונים במבחן ה-IQ לטובת ישראל.  
רשמו את כל ההנחות הדרושות לצורך פתרון התרגיל.

- (3) להלן תוצאות מדגם הבדק אורך חיים של נורות מסוג W60 ומסוג W100. אורך החיים נמדד בשעות.

| הקבוצה    | 60W  | 100W |
|-----------|------|------|
| $\bar{X}$ | 1007 | 956  |
| $S$       | 80   | 72   |
| $n$       | 13   | 15   |

- א. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקות בממוצע יותר מאשר נורות מסוג W100. רשמו את כל ההנחות הדרושות לפתרון.  
ב. עבור איזו רמת מובהקות ניתן לקבוע שנוורות מסוג W60 דולקות בממוצע יותר מאשר נורות מסוג 100?  
ג. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקות יותר מ 1000 שעות. רשמו את כל ההנחות הדרושות.

**תשובות סופיות**

- (1) נדחה את  $H_0$ .
- (2) הנחות:
1. סטיות התקן שוות.
  2. המשתנים מתפלגים נורמלית.
- נקבל את  $H_0$ .
- (3) א. נדחה את  $H_0$ . ב. רמת מובהקות של לפחות 5%.
- ג. לא נדחה את  $H_0$ .