

## סטטיסטיקה ב

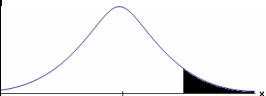
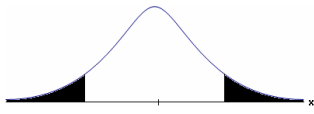
פרק 6 - בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

תוכן העניינים

1. כששונויות האוכלוסיה לא ידועות ומניחים שהן שוות.....1
2. ניתוח פלטים.....5

## בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

כששונות האוכלוסייה לא ידועות ומניחים שהן שוות – רקע

| השערת האפס:<br>השערת אלטרנטיבית:                                                                                                                                                               | תנאים:                                                                                                                                                                                          | אזור הדחייה של $H_0$ :                                                                                                                                                                                                                                                                                                          |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $H_0 \quad \mu_1 - \mu_2 = c$<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 $H_0$ |

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

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \quad \text{השונות המשוקללת:}$$

## 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 שנורות 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.

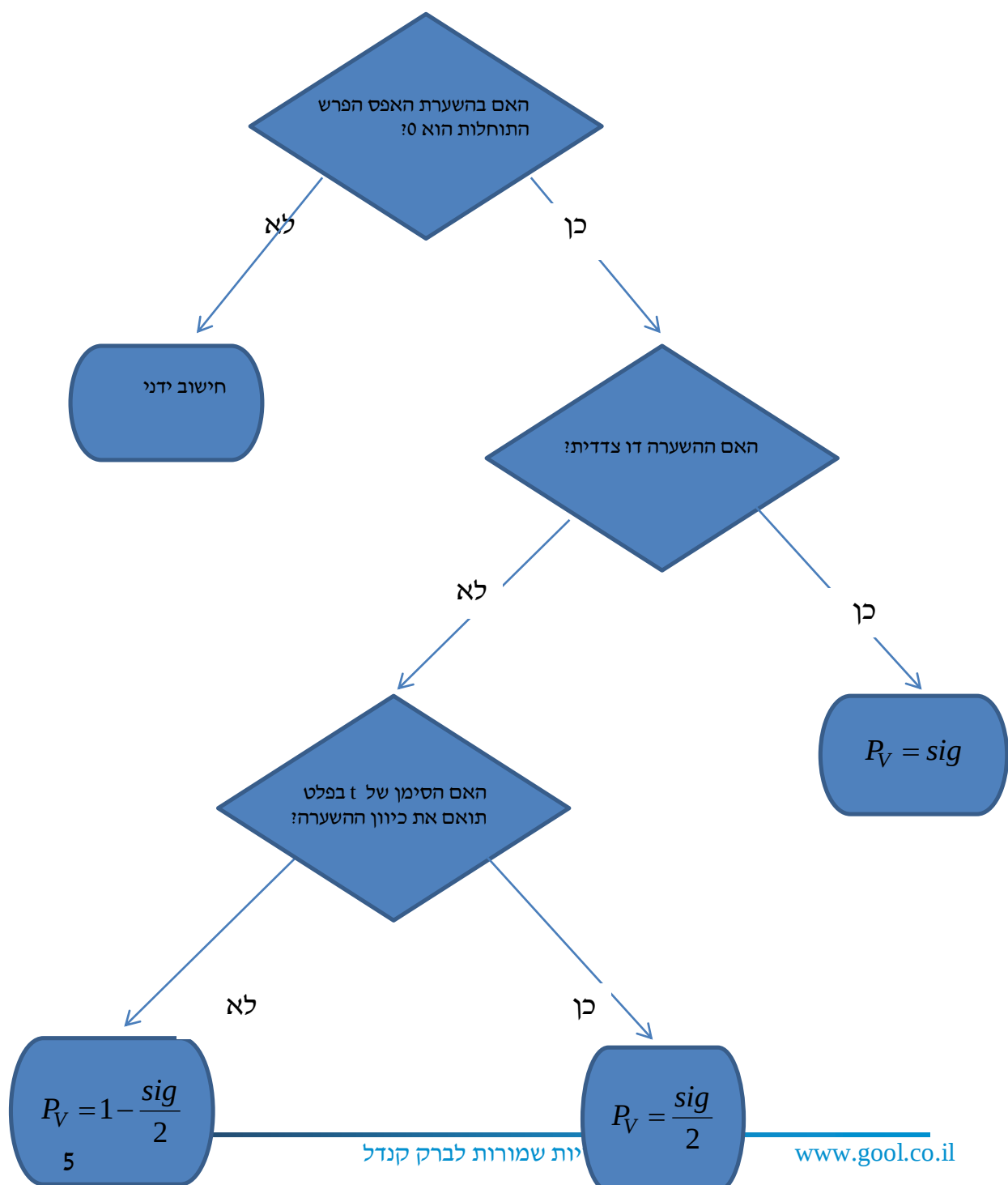
**תשובות סופיות**

- (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  $H_0$ .

## בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

### ניתוח פלטים – רקע

מובהקות התוצאה על סמך הפלט:



### דוגמה (פתרון בהקלטה) :

בסקר שנערך בארה"ב בשנת 1993 נשאלו נסקרים משני אזורים שונים במדינה על מס' האחים והאחיות שלהם. להלן הפלט שהתקבל :

Group Statistics

|                                | Region of the United States | N   | Mean | Std. Deviation | Std. Error Mean |
|--------------------------------|-----------------------------|-----|------|----------------|-----------------|
| Number of Brothers and Sisters | North East                  | 676 | 3.76 | 2.939          | .113            |
|                                | South East                  | 410 | 4.05 | 2.993          | .148            |

Independent Samples Test

|                                |                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |         |                 |                 |                       |                                           |       |
|--------------------------------|-----------------------------|-----------------------------------------|------|------------------------------|---------|-----------------|-----------------|-----------------------|-------------------------------------------|-------|
|                                |                             | F                                       | Sig. | t                            | df      | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |       |
|                                |                             |                                         |      |                              |         |                 |                 |                       | Lower                                     | Upper |
| Number of Brothers and Sisters | Equal variances assumed     | .173                                    | .677 | -1.583                       | 1084    | .114            | -.293           | .185                  | -.657                                     | .070  |
|                                | Equal variances not assumed |                                         |      | -1.576                       | 850.945 | .115            | -.293           | .186                  | -.658                                     | .072  |

- א. מהו המבחן הסטטיסטי שנעשה כאן?
- ב. בדוק ברמת מובהקות של 5% האם קיים שוויון שונויות בין שני האזורים?
- ג. בדוק האם קיים הבדל בין "South East" ל-"North East" ברמת מובהקות של 5% מבחינת מספר האחים והאחיות הממוצע.
- ד. מהי מובהקות התוצאה לבדיקת הטענה שההפרש הממוצע בין "South East" לבין "North East" חיובי?

## שאלות

- 1) להלן פלט מתוכנת SPSS מתוך מחקר שבחן את רמת האופטימיות של גברים ונשים. רמת האופטימיות נמדדה בסולם ציונים של 1 עד 5.

Group Statistics

|          | GENDER | N   | Mean   | Std. Deviation | Std. Error Mean |
|----------|--------|-----|--------|----------------|-----------------|
| optimizm | MALE   | 633 | 2.6053 | .49781         | .01979          |
|          | FEMALE | 568 | 2.5503 | .48483         | .02034          |

Independent Samples Test

|          |                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |          |                 |                 |                       |                                           |        |
|----------|-----------------------------|-----------------------------------------|------|------------------------------|----------|-----------------|-----------------|-----------------------|-------------------------------------------|--------|
|          |                             | F                                       | Sig. | t                            | df       | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |        |
|          |                             |                                         |      |                              |          |                 |                 |                       | Lower                                     | Upper  |
| optimizm | Equal variances assumed     | .383                                    | .536 | 1.935                        | 1199     | .053            | .05500          | .02842                | -.00076                                   | ?      |
|          | Equal variances not assumed |                                         |      | 1.938                        | 1190.977 | .053            | .05500          | .02838                | -.00068                                   | .11067 |

- א. האם ניתן להניח ששוונות האופטימיות של נשים וגברים שווה ברמת מובהקות של 5%?
- ב. ברמת מובהקות של 5% האם קיים הבדל בין הנשים לגברים ברמת האופטימיות הממוצעת שלהם?
- ג. מצא את הגבול העליון של רווח הסמך המסומן בסימן שאלה בפלט. דייק עד 5 ספרות אחרי הנקודה.
- ד. בנה רווח סמך לתוחלת רמת האופטימיות של הגברים ברמת סמך של 95%.

(2) פסיכולוגים טוענים שאנשים שניגשים למבחן אינטליגנציה יותר מפעם אחת נוטים לקבל ציונים גבוהים יותר. להלן הפלט שהתקבל:

### Group Statistics

|       |   | N  | Mean     | Std. Deviation | Std. Error Mean |
|-------|---|----|----------|----------------|-----------------|
| grade | A | 9  | 96.8889  | 9.40006        | 3.13335         |
|       | B | 11 | 108.4545 | 11.46616       | 3.45718         |

| Independent Samples Test |                             |                                         |      |                              |        |                 |                 |                       |                                           |          |
|--------------------------|-----------------------------|-----------------------------------------|------|------------------------------|--------|-----------------|-----------------|-----------------------|-------------------------------------------|----------|
|                          |                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |        |                 |                 |                       |                                           |          |
|                          |                             | F                                       | Sig. | t                            | df     | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |          |
|                          |                             |                                         |      |                              |        |                 |                 |                       | Lower                                     | Upper    |
| grade                    | Equal variances assumed     | .206                                    | .656 | -2.428                       | 18     | .026            | -11.56566       | 4.76333               | -21.57304                                 | -1.55828 |
|                          | Equal variances not assumed |                                         |      | -2.479                       | 17.997 | .023            | -11.56566       | 4.66583               | -21.36832                                 | -1.76299 |

T-Test

מקרא:

A = נגשו פעם אחת.

B = נגשו יותר מפעם אחת.

א. רשמו את השערות המחקר והסבירו מהו המבחן המתאים כאן.

ב. כיצד הייתה משתנה התשובה לסעיף הקודם אם היה מדובר על אותם אנשים שציונם נבדק פעם אחרי המבחן הראשון שעשו ופעם אחרי המבחן השני?

ג. האם ניתן לומר כי מידת הפיזור של ציוני אנשים הנבחנים בפעם הראשונה שונה ממידת הפיזור של ציוני האנשים אשר נבחנים בפעם השנייה. בדוק

ברמת מובהקות של  $\alpha = 0.05$ .

ד. האם נכונה טענת הפסיכולוגים ברמת מובהקות של  $\alpha = 0.01$ .

3) כחלק ממחקר בנושא הנישואין בישראל, אחד החוקרים העלה השערה שיש הבדל בממוצע גיל הנישואין (הראשונים), בין נשים הגרות בערים מרכזיות לבין נשים הגרות בערים מרוחקות מהמרכז. לשם כך נדגמו 50 כלות מכל אחת משתי ערים עיר א'-מרכזית ועיר ב'-מרוחקת ונרשם גילן. תוצאות עיבוד הנתונים מופיעות בטבלאות שלהלן:

## T-Test

Group Statistics

| מקום המגורים       | N  | Mean    | Std. Deviation | Std. Error Mean |
|--------------------|----|---------|----------------|-----------------|
| גיל הנישואין עיר א | 50 | 24.8072 | 1.38978        | .19654          |
| עיר ב              | 50 | 23.0131 | 1.62070        | .22920          |

Independent Samples Test

|              |                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |        |                 |                 |                       |                                           |         |
|--------------|-----------------------------|-----------------------------------------|------|------------------------------|--------|-----------------|-----------------|-----------------------|-------------------------------------------|---------|
|              |                             | F                                       | Sig. | t                            | df     | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |         |
|              |                             |                                         |      |                              |        |                 |                 |                       | Lower                                     | Upper   |
| גיל הנישואין | Equal variances assumed     | .330                                    | .567 | 5.942                        | 98     | .000            | 1.79415         | .30193                | 1.19497                                   | 2.39332 |
|              | Equal variances not assumed |                                         |      | 5.942                        | 95.772 | .000            | 1.79415         | .30193                | 1.19480                                   | 2.39350 |

- א. מהו המבחן הסטטיסטי שנעשה כאן?
- ב. מצא רווח סמך ברמת סמך של 95% להפרש בין עיר א לעיר ב מבחינת גיל הנשים הממוצע בנישואין הראשונים.
- ג. האם ניתן לומר ברמת מובהקות של 1% שנשים בערים מרכזיות מתחתנות בגיל מאוחר יותר מאשר נשים הגרות בערים מרוחקות?

## (4) להלן פלט של תוכנת SPSS :

## T-Test

|   | N  | Mean    | Std. Deviation | Std. Error Mean |
|---|----|---------|----------------|-----------------|
| X | 26 | 36.3077 | 13.23259       | 2.59513         |
| Y | 24 | 46.4583 | 20.96369       | 4.27920         |

## Independent Samples Test

|                             | Levene's Test for Equality of Variances |      | t-test for Equality of Means |        |                 |                 |                       |                                           |         |
|-----------------------------|-----------------------------------------|------|------------------------------|--------|-----------------|-----------------|-----------------------|-------------------------------------------|---------|
|                             | F                                       | Sig. | t                            | df     | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |         |
|                             |                                         |      |                              |        |                 |                 |                       | Lower                                     | Upper   |
| Equal variances assumed     | 4.446                                   | .040 | -2.164                       | ???    | .044            | -10.15064       | ???                   | -20.03781                                 | -.26347 |
| Equal variances not assumed |                                         |      | -2.038                       | 38.267 | .048            | ???             | 5.00462               | -20.27964                                 | -.02164 |

א. השלימו את סימני השאלה בטבלה.

ב. מהי מובהקות התוצאה לבדיקת הטענה שקיים הבדל בין השונות של  $X$  לזה של  $Y$ ?

ג. מהי מובהקות התוצאה לבדיקת הטענה שהתוחלת של  $X$  גדולה מהתוחלת של  $Y$ ?

ד. מהי מובהקות התוצאה לבדיקת הטענה שהתוחלת של  $X$  קטנה מהתוחלת של  $Y$ ?

## תשובות סופיות

- (1) א. נקבל את  $H_0$  ונכריע שיש שוויון שוניות.  
 ב. נקבע שלא קיים הבדל בין נשים לגברים מבחינת האופטימיות הממוצעת.  
 ג. 0.11076  
 ד.  $2.5665 \leq \mu \leq 2.6441$
- (2) א. מבחן T להפרש ממוצעים במדגמים בלתי תלויים.  
 ב. מבחן T למדגמים מזווגים.  
 ג. נקבל את  $H_0$ , נקבע לקיום שוויון שוניות.  
 ד. נקבל את  $H_0$ , לא נקבל את טענת הפסיכולוגים.
- (3) א. מבחן T להשוואת תוחלת במדגמים בלתי תלויים.  
 ב.  $1.19497 \leq \mu_1 - \mu_2 \leq 2.39332$  ג. כן.
- (4) א. 10.15, 4.69, -48  
 ב. 0.04  
 ג. 0.978  
 ד. 0.022