# ביוסטייסטיקה ב

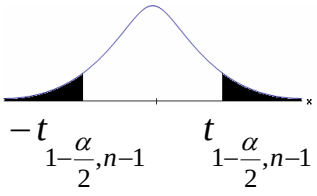
## פרק 4 - בדיקת השערות לתוחלת ההפרש במדגמים מזווגים

### תוכן העניינים

1. בדיקת השערות למדגמים מזווגים ..... 1
2. ניתוח פלטים ..... 5

# בדיקת השערות על תוחלת ההפרשים במדגמים מזווגים (תלויים)

## בדיקת השערות למדגמים מזווגים – רקע

| | | | |
|---|---|---|---|
| **השערת האפס:** $H_0$<br>**השערה אלטרנטיבית:** | $H_0: \mu_D = C$<br>$H_1: \mu_D \neq C$ | $H_0: \mu_D = C$<br>$H_1: \mu_D < C$ | $H_0: \mu_D = C$<br>$H_1: \mu_D > C$ |
| **תנאים:** | 1. $\sigma_D$ אינה ידועה<br>2. $D \sim N$ או מדגם מספיק גדול | | |
| **כלל ההכרעה:**<br>אזור הדחייה של $H_0$: | $t_{\bar{D}} > t^{(n-1)}_{1-\frac{\alpha}{2}}$ או $t_{\bar{D}} < -t^{(n-1)}_{1-\frac{\alpha}{2}}$<br>  <br><br>■ - דוחים את $H_0$ | $t_{\bar{D}} < -t^{(n-1)}_{1-\alpha}$<br> <br><br>■ - דוחים את $H_0$ | $t_{\bar{D}} > t^{(n-1)}_{1-\alpha}$<br> <br><br>■ - דוחים את $H_0$ |
| **חלופה לכלל הכרעה:**<br>נדחה $H_0$ אם מתקיים: | $\bar{D} > C + t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{S_D}{\sqrt{n}}$<br>**או**<br>$\bar{D} < C - t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{S_D}{\sqrt{n}}$ | $\bar{D} < C - t^{n-1}_{1-\alpha} \cdot \frac{S_D}{\sqrt{n}}$ | $\bar{D} > C + t^{n-1}_{1-\alpha} \cdot \frac{S_D}{\sqrt{n}}$ |

**סטטיסטי המבחן:** $S_D{}^2 = \frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1} = \frac{\sum_{i=1}^{n} D^2{}_i - n\bar{D}^2}{n-1}$ , $t_{\bar{D}} = \frac{\bar{D} - \mu_D}{S_D / \sqrt{n}}$

**דוגמה (פתרון בהקלטה):**


© כל הזכויות שמורות לברק קנדל

חברה שיווקית מעוניינת לבדוק את טענת רשת השיווק "מגה בעיר" הטוענת שמחיריה נמוכים מהמחירים מרשת השיווק "שופרסל".
לצורך הבדיקה נבחרו באקראי 4 מוצרים שונים. המחירים נבדקו בשתי הרשתות.
להלן המחירים:

| המוצר / רשת | מגה בעיר | שופרסל |
|---|---|---|
| שמפו | 17 | 18 |
| ג'ל כביסה | 48 | 57 |
| עוגת גבינה | 35 | 35 |
| לחם | 12 | 10 |
| קפה נמס | 49 | 47 |
| בקבוק יין | 113 | 142 |
| גבינה בולגרית | 20 | 26 |

בהנחה והמחירים מתפלגים נורמאלית, בדקו ברמת מובהקות של 5% את טענת רשת "מגה בעיר".


© כל הזכויות שמורות לברק קנדל

## שאלות

**1)** במטרה לבדוק האם קיים הבדל בין חברת $X$ לחברת $Y$ מבחינת המחירים לשיחות בינ"ל. נדגמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה. להלן התוצאות:

| חברה/ מדינה | ארה"ב | קנדה | הולנד | פולין | מצרים | סין | יפן |
|---|---|---|---|---|---|---|---|
| $X$ | 1.5 | 2.1 | 2.2 | 3 | 3.5 | 3.2 | 4.2 |
| $Y$ | 1.4 | 2 | 1.9 | 3.1 | 3.2 | 3.2 | 4.2 |

בהנחה והמחירים מתפלגים נורמלית בכל חברה, בדקו ברמת מובהקות של 5% האם קיים הבדל בין החברות מבחינת המחירים בממוצע?

**2)** מכון המכין לפסיכומטרי טוען שהוא מעלה את ממוצע הציונים ביותר מ-30 נקודות. 8 נבחנים נבדקו לפני ואחרי שהם למדו במכון. להלן התוצאות שהתקבלו:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| לפני | 506 | 470 | 420 | 640 | 670 | 390 | 500 | 590 |
| אחרי | 570 | 540 | 430 | 610 | 680 | 510 | 520 | 580 |

מה מסקנתכם ברמת מובהקות 5%? הניחו שציוני פסיכומטרי מתפלגים נורמלית.

**3)** נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'. להלן הציונים שלהם בסמסטר א' ו- ב':

| | | | | | |
|---|---|---|---|---|---|
| סטטיסטיקה א' | 74 | 68 | 90 | 75 | 82 |
| סטטיסטיקה ב' | 80 | 84 | 87 | 76 | 100 |

פורסם שתלמידים שמסיימים את סמסטר ב משפרים בממוצע את הציונים ב-5 נקודות לעומת סמסטר א'. הניחו שהציונים מתפלגים נורמלית.

א. מהי מובהקות התוצאה לבדיקת הטענה שהשיפור הוא יותר מ 5 נקודות?

ב. על סמך הסעיף הקודם, מהי רמת המובהקות המינימלית להכרעה שהשיפור הוא יותר מ- 5 נקודות?

ג. לאור זאת, מה המסקנה ברמת מובהקות של 10%?

**4)** לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני ללא היפנוזה. לאחר מכן נערך לשניהם מבחן באנגלית. נניח שציוני המבחן מתפלגים נורמאלית ללא ידיעת השונות האמתית. המבחן שיש לבצע כאן הוא:

א. מבחן $Z$ למדגם יחיד.

ב. מבחן $T$ למדגם יחיד.

ג. מבחן $T$ למדגמים בלתי תלויים.

ד. מבחן $T$ למדגמים מזווגים.


© כל הזכויות שמורות לברק קנדל

**5)** בתחנת טיפת חלב מסוימת יש שני מכשירי שקילה. על מנת להשוות בין שני המשקלים נדגמו 4 תינוקות. כל תינוק בן חודשיים נשקל בכל אחד מהמשקלים. להלן תוצאות השקילה (בק"ג):

| | | | | |
|---|---|---|---|---|
| משקל במכשיר 1 | 4.5 | 9.6 | 0.7 | 2.5 |
| משקל במכשיר 2 | 3.5 | 6.9 | 1.7 | 0.5 |

נניח שהמשקלים מתפלגים נורמלית, המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.

ב. מבחן T למדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

**6)** כדי להשוות בין שני אצנים נדגמו 5 תוצאות מריצת 100 מטר של כל אצן. זמני הריצה נרשמו ויש להניח שמתפלגים נורמלית. המטרה להשוות בין האצנים. המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.

ב. מבחן T למדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

## תשובות סופיות

**1)** לא נדחה $H_0$.

**2)** לא נדחה $H_0$.

**3)** א. $0.25 \le pv \le 0.5$ ב. 0.5 ג. לא נדחה $H_0$.

**4)** ד'.

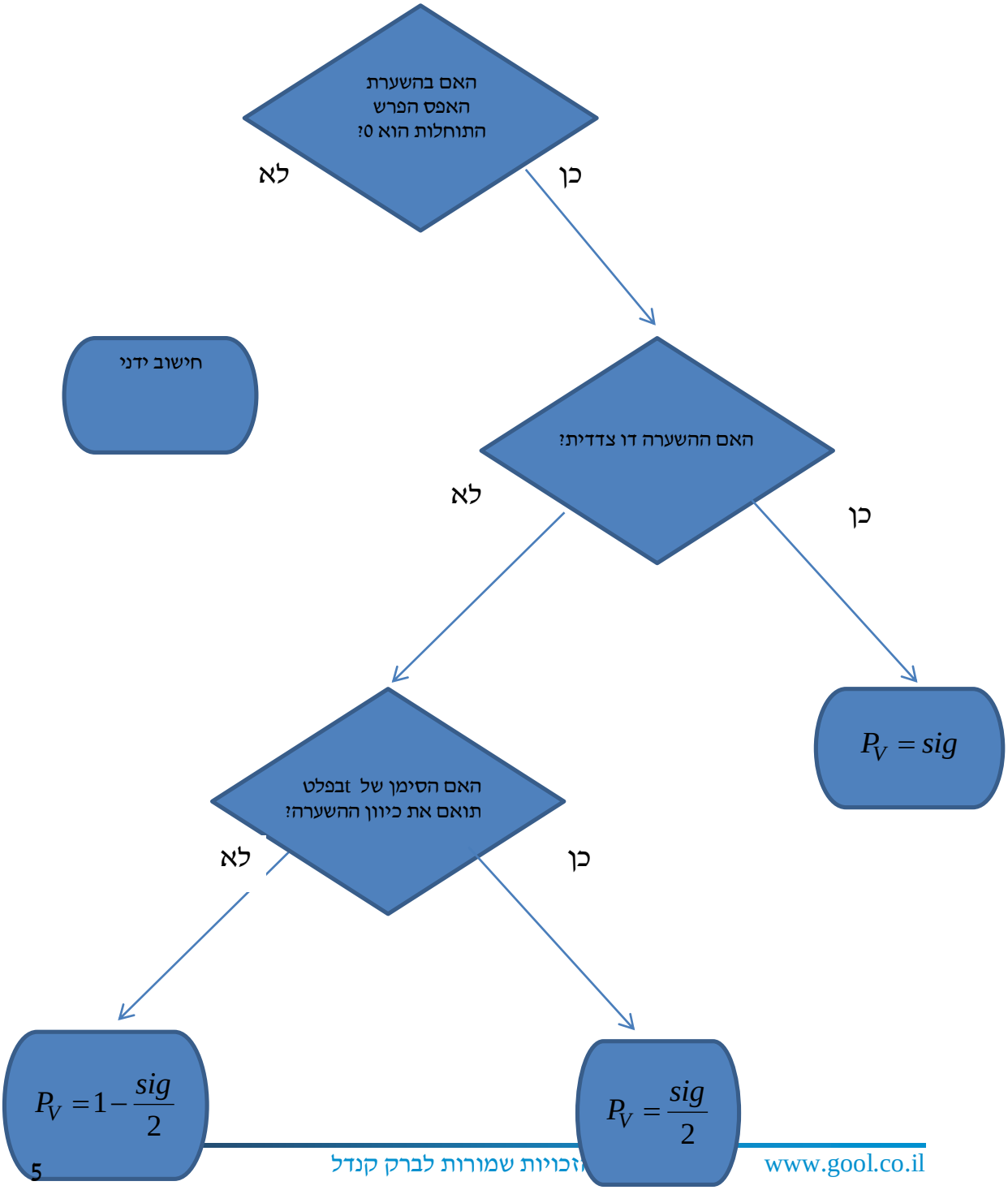
**5)** ד'.

**6)** ג'.


© כל הזכויות שמורות לברק קנדל

# בדיקת השערות על תוחלת ההפרשים במדגמים מזווגים (תלויים)

## מדגמים מזווגים – ניתוח פלטים – רקע





כל הזכויות שמורות לברק קנדל

**דוגמה (פתרון בהקלטה) :**

כדי לבדוק את ההשפעה של קורס לגמילה מעישון נלקח מדגם מקרי של 5 נבדקים. עבור כל אחד מהם נמדדה צריכת הסיגריות היומית לפני הקורס וחודשיים אחריו. הניחו שצריכת הסיגריות מתפלגת נורמלית. להלן התוצאות:

| נבדק | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| לפני | 40 | 22 | 25 | 28 | 30 |
| אחרי | 30 | 24 | 13 | 10 | 12 |

**Paired Samples Statistics**

| | | Mean | N | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Pair 1 | BEFORE | 29.0000 | 5 | 6.85565 | 3.06594 |
| | AFTER | 17.8000 | 5 | 8.72926 | 3.90384 |

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 90% Confidence Interval of the Difference Lower | Upper | | | |
| Pair 1 | BEFORE - AFTER | 11.20000 | 8.19756 | 3.66606 | 3.38452 | 19.01548 | 3.055 | 4 | .038 |

בדקו ברמת מובהקות של 5% האם הקורס יעיל.


© כל הזכויות שמורות לברק קנדל

## שאלות

**1)** בסקר שנערך בארה"ב בשנת 1993 נשאלו נסקרים על השכלת הוריהם, להלן הפלט שהתקבל:

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference Lower | Upper | | | |
| Pair 1 | Highest Year School Completed, Father - Highest Year School Completed, Mother | -.007 | 3.115 | .100 | -.203 | .189 | -.072 | 973 | .943 |

א. תנו אומדן להפרש הממוצעים.

ב. תנו אומדן לטעות התקן של הפרש הממוצעים.

ג. האם קיים הבדל מובהק בין השכלת האבות להשכלת האימהות ברמת מובהקות של 5%?

**2)** בתחרות קפיצה למים שופטים באופן קבוע שופט איטלקי ושופט דרום קוריאני. להלן פלט המנתח את הציונים ששופטים אלה נתנו בתחרויות השונות:

**Paired Samples Statistics**

| | | Mean | N | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Pair 1 | Italy | ??? | 300 | .86742 | .05008 |
| | South Korea | 8.9183 | ??? | .81992 | .04734 |

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference Lower | Upper | | | |
| Pair 1 | Italy - South Korea | -.42233 | .36153 | .02087 | -.46341 | -.38126 | -20.234 | ??? | ??? |

א. השלימו את החלקים החסרים בפלט (מסומנים בסימני שאלה).

ב. בדקו את הטענה שהשופט הדרום קוריאני נותן בממוצע 0.2 נקודות יותר מאשר השופט האיטלקי ברמת מובהקות של 5%.

ג. מהו רווח הסמך ברמת סמך של 95% לתוחלת פער הציונים בין השופטים?

ד. בנו את הרווח כעת ברמת סמך של 98% לתוחלת פער בציונים בין השופטים.


© כל הזכויות שמורות לברק קנדל

**3)** בדקו את ציוניהם של 44 נבדקים אקראיים במבחן הפסיכומטרי. פעם אחת לפני הכנה (Before) ופעם אחת אחרי הכנה (After).

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference Lower | Upper | | | |
| Pair 1 | Before - After | -7.45455 | 19.28303 | 2.90703 | -13.31712 | -1.59197 | -2.564 | 43 | .014 |

א. רשמו מהו המבחן הסטטיסטי ונסח את ההשערות אליהם מתייחס הפלט.

ב. בדקו את ההשערה שממוצע ציונים משתפרים לאחר ההכנה ברמת מובהקות של 5%.

ג. בדקו את ההשערה שממוצע ציונים משתפרים לאחר ההכנה ביותר מ-5 נקודות ברמת מובהקות של 5%.

ד. מצאו רווח סמך לתוחלת שיפור ממוצע הציונים לאחר ההכנה ברמת ביטחון של 95%.


© כל הזכויות שמורות לברק קנדל

**4)** להלן פלט של תכנת SPSS:

**T-Test**

**Paired Samples Statistics**

| | | Mean | N | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Pair 1 | x | 54.0000 | 6 | 5.86515 | 2.39444 |
| | y | 46.5000 | 6 | 10.72847 | 4.37988 |

**Paired Samples Test**

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | | | |
| | | | | | Lower | Upper | | | |
| Pair 1 | x - y | 7.50000 | ?? | 4.72405 | -4.64356 | 19.64356 | ?? | 5 | .173 |

א. מלא את החלקים החסרים בטבלה.

ב. מהי רמת המובהקות המינימלית לקבלת הטענה שיש הבדל בין $X$ ל- $Y$ בממוצע?

ג. האם התשובה לסעיף הקודם הייתה משתנה, ואם כן גדלה או קטנה,
אם הינו מוסיפים עוד תצפית שההפרש בין $X$ ל- $Y$ הוא 0.

ד. מהי מובהקות התוצאה לבדיקת הטענה ש $X$ גדול מ- $Y$ בממוצע?

ה. מהי מובהקות התוצאה לבדיקת הטענה ש $X$ קטן מ- $Y$ בממוצע?

ו. בנו רווח סמך לתוחלת של $X$ ברמת סמך של 90%.


© כל הזכויות שמורות לברק קנדל

## תשובות סופיות

**1)** א. -0.007 ב. 0.1 ג. אין הבדל מובהק.

**2)** א. $d.f = 299$, $n = 300$, $\bar{X} = 8.496$, $\text{Sig} = 0$. ב. נדחה את $H_0$. ג. $(0.46341, 0.38126)$ ד. $(0.4708, 0.3738)$.

**3)** א. ראה וידאו. ב. נדחה את $H_0$. ג. לא נדחה את $H_0$. ד. (1.592,13.317).

**4)** א. 1.5876, 11.5715 ב. 0.173 ג. יגדל. ד. 0.0865 ה. 0.9135 ו. $49.18 < \mu < 58.82$.


© כל הזכויות שמורות לברק קנדל