שיטות מחקר־ תרגיל 6

מגישות־ זוהר בוחניק 311142293 אלה דובדבן 305564866

14 במאי 2019

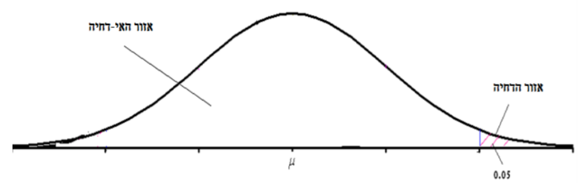
שאלה 1:

נתונים:
מנת המשכל באוכלוסיה הכללית־ מתפלגת $N\left(100, 15^2\right)$.
נתוני רמת המשכל של n חובבי מוזיקה קלאסית, ממוצע הנתונים־ 106.

חוקר א'
השערה־
H_1 לחובבי מוזיקה קלאסית יש רמת משכל גבוהה יותר מהאוכלוסיה הכללית.
H_0 לחובבי מוזיקה קלאסית אין רמת משכל גבוהה יותר מהאוכלוסיה הכללית.
כלומר הוא מבצע מבחן **חד** זנבי

חוקר ב'
השערה־
H_1 לחובבי מוזיקה קלאסית יש רמת משכל שונה מהאוכלוסיה הכללית.
H_0 לחובבי מוזיקה קלאסית יש רמת משכל דומה לאוכלוסיה הכללית.
כלומר הוא מבצע מבחן **דו** זנבי וזאת על ידי רווח בר סמך.
כדי לדחות את השערת האפס 100צריך להיות מחוץ לרווח הבר סמך.

מכך שממוצע נתוני רמת המשכל הוא 106, ברור כי רווח בר הסמך נמצא בטווח $x_1 \leq 106 \leq x_2$ (כאשר $x_1, x_2 \in \mathbb{R}$ ניתנים לחישוב בביצוע המבחן). אם 100לא נמצא ברווח בר הסמך, אזי ניתן לדחות את השערת האפס של חוקר ב'. נשים לב כי נובע מכך ש־ $100 > x_1$, אחרת 100 היה חייב להיות ברווח בר סמך. כיוון שהמבחן של חוקר א' הוא חד זנבי עם אותה α, הערך סף שלו לדחיית השערת האפס נמוך יותר ובגלל שראינו שבמבחן הדו זנבי התוחלת הייתה גדולה מ 100 (וקטנה מ x_2) בסיכוי של 95% נסיק שגם ניתן לדחות את השערת האפס של חוקר א'
אזור הדחייה של חד זנבי־



לכן התשובה היא: ב.

שאלה 2:

מספר נבדק	לוח טרמפים	480	רכב פרטי	מסוף שפירים	מונית שירות
1	123	112	111	126	107
2	113	117	115	123	135
3	127	105	117	130	126
4		130	113	89	131
5		119	99		123
6		115			

סעיף א'

ממוצעים:

1- לוח טרמפים	2- קו 480	3- רכב פרטי	4- מסוף שפירים	5- מונית שירות	ממוצע כללי
$\frac{363}{3}=121$	$\frac{698}{6}=116.33$	$\frac{555}{5}=111$	$\frac{468}{4}=117$	$\frac{622}{5}=124.4$	117.652

כאשר את הממוצע הכללי חישבנו על ידי סכימת כל הערכים בטבלה וחלוקה ב- 23 שזה מספר הערכים בטבלה.

סעיף ב'

נבצע מבחן ניתוח שונות חד גורמי עבור רמת מובהקות 5% ובדקו האם יש הבדל בין הקבוצות. ניתן להניח שיוויון שונויות והתפלגות נורמלית.

השערות:
H_0: $\mu_1=\mu_2=\mu_3=\mu_4=\mu_5=\mu$ כלומר מניחים שהם מאותה התפלגות.
H_1: המשלים- יש הבדלים בין הקבוצות.

הליך בדיקת ההשערות:
עבור כל קבוצה $i\in[5]$,

$$SS_i=\sum_{j=1}^{n_i}\left(y_{i,j}-\bar{Y}_i\right)^2$$

כך ש $SSW=\sum_{i=1}^{k}SS_i$.
נחשב:

$$SS_1=(123-121)^2+(113-121)^2+(127-121)^2=104$$

$$SS_2=(112-116.33)^2+(117-116.33)^2+(105-116.33)^2+(130-116.33)^2+(119-116.33)^2+(115-116.33)^2=343.33$$

$$SS_3=(111-111)^2+(115-111)^2+(117-111)^2+(113-111)^2+(99-111)^2=200$$

$$SS_4=(126-117)^2+(123-117)^2+(130-117)^2+(89-117)^2=1070$$

$$SS_5=(107-124.4)^2+(135-124.4)^2+(126-124.4)^2+(131-124.4)^2+(123-124.4)^2=463.2$$

ולכן:

$$SSW=104+343.33+200+1070+463.2=2180.53$$

סכום הריבועים בין הקבוצות

$$SSB=\sum_{i=1}^{k}n_i\left(\bar{Y}_i-\bar{\bar{Y}}\right)^2=$$

$$=3\left(121-117.652\right)^2+6\left(116.33-117.652\right)^2+5\left(111-117.652\right)^2+4\left(117-117.652\right)^2+5\left(124.4-117.652\right)^2=$$

$$=494.736872$$

נאמוד את השונות בתוך כל אוכלוסייה ובין אוכלוסיות באמצעות ה SSW וה־ SSB:

$$MSB = \frac{SSB}{df_{between}} = \frac{SSB}{k-1} = \frac{494.736872}{4} = 123.68$$

$$MSW = \frac{SSW}{df_{within}} = \frac{SSW}{N-k} = \frac{2180.53}{18} = 121.14$$

תחת השערת האפס, מתקיים: $\frac{MSB}{MSW} \sim F_{4,18}$
אם נקבל תוצאה קיצונית ביחס להתפלגות F נוכל לדחות את השערת האפס.

$$\frac{MSB}{MSW} = \frac{123.68}{121.14} = 1.02096$$

נבדוק בטבלת F עבור $\alpha = 0.05$: הערך הקריטי הוא 2.9277.

על כן **לא ניתן לדחות את השערת האפס** (כי $1.02096 < 2.9277$).

סעיף ג'

מה יקרה לערך F הנצפה אם יתברר שנפלה טעות בנתונים וזמן ההגעה של כל אדם בקבוצת "רכב פרטי" הוא למעשה גבוה יותר ב 4 דק' לזמן ההגעה?
לאחר השינוי המדובר כיוון שאנחנו מעלים את כל הדגימות של אותה הקבוצה בקבוע זהה אנחנו לא משפיעים על השונות ואז נקבל כי $SS_3 = \sum_{j=1}^{5}\left(y_{3,j} - \bar{Y_3}\right)^2$ לא ישתנה. כיוון ש־SSW מורכב מהסכום של ה SS_i עבור כל הקבוצות, נקבל כי גם SSW לא ישתנה כלל, ובעקבותיו גם $MSW = \frac{SSW}{df_{within}}$ לא ישתנה. עבור $SSB = \sum_{i=1}^{k} n_i\left(\bar{Y_i} - \bar{\bar{Y}}\right)^2$ נצפה כן לשינוי כיוון שהוא בודק רעש בין נבדקי והרי הוספנו לאחת מהקבוצות את הקבוע 4 ולכן שינינו את היחסים בין הקבוצות השונות. כיוון שספציפית ממוצע הנסיעה ברכב פרטי היה נמוך יותר מכל השאר, נצפה לכך שהוא עכשיו יקטין את הפערים, הממוצע של רכב פרטי יעלה וכן גם הממוצע הכללי אבל ההפרש בניהם יקטן, כלומר SSB יקטן בהתאמה. מכאן נובע כי $MSB = \frac{SSB}{df_{between}}$ יקטן גם הוא.
נשים לב כי הסטטיסטי מורכב מהמנה $\frac{MSB}{MSW}$, כך שראינו כי המונה יקטן והמכנה לא ישתנה, מכאן נובע כי המנה תקטן ביחס לנתונים הקודמים. על כן הערך F הנצפה יקטן (יהיה יותר קשה לדחות את השערת האפס).

סעיף ד'

מה יקרה לערך F הנצפה שנקבל אם מכל האנשים בכל הקבוצות נחסיר קבוע של 10 דק? כיוון שאנחנו מחסירים באופן אחיד קבוע מכל הקבוצות, אנו שומרים על היחס בניהן ולא משנים את השונות. נזכור כי המטרה שלנו במבחן $ANOVA$ היא לזהות הבדלים בין הקבוצות (האם הן מאותה התפלגות או לא) ואם מזיזים בקבוע את כל הנתונים עדיין ניתן באותה מידה לבדוק את היחסים והערך F הנצפה לא ישתנה. ניתן לראות גם ברמה המתמטית בחישוב SSW ו־ SSB שהשינוי בקבוע לא ישנה את המנה $\frac{MSB}{MSW}$.

שאלה 3:

סעיף א'

<u>ניסוח ההנחות:</u>

1. כמות החרקים מתפלגת נורמלית באיזורים שבהם השתמשו בסוגי הספריי השונים.
2. השונויות זהות בין הקבוצות.

<u>ניסוח ההשערות</u> ־ מירב משערת שלאחר השימוש בספריי A, כמות החרקים שנותרו גדולה מכמות החרקים שנותרו לאחר השימוש בספריי B.

השערת H_0: $\mu_A \leq \mu_B$
השערת H_1: $\mu_A > \mu_B$

<u>ביצוע המבחן וניתוח התוצאות:</u>

היות ואנו נדרשים להכריע בין 2 השערות תוך קביעה האם אחת טובה מהשניה ־ ניעזר במבחן t.

תוצאות ההרצה ב־R:

```
data:  A$count and B$count
t = -0.45352, df = 22, p-value = 0.6727
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -3.988519       Inf
sample estimates:
mean of x mean of y
 14.50000  15.33333
```

נשים לב שקיבלנו כי p_value, אשר משקף את ההסתברות לקבלת סט ה־data הנ"ל תחת הנחת H_0 מקיים $p\_value > \alpha$ עבור $\alpha = 5\%$.

כלומר, ההסתברות גבוהה מהסף, ולכן לא ניתן לדחות את השערת האפס H_0, ולא ניתן לקבוע כי $\mu_A > \mu_B$.
מעבר לכך, נשים לב לתוחלות $\text{mean of A} = 14.5$, $\text{mean of B} = 15.3$, אשר למעשה נוגדות את H_1.
<u>מסקנה</u> ־ לא ניתן לומר כי תכשיר הספריי B מביא לידי יותר חרקים מתים.

סעיף ב'

ניסוח ההנחות:

1. כמות החרקים מתפלגת נורמלית באיזורים שבהם השתמשו בסוגי הספריי השונים.

2. השונויות זהות בין הקבוצות.

ניסוח השערות ־ מירב משערת שלאחר השימוש בספריי A, כמות החרקים שנותרו גדולה מכמות החרקים שנותרו לאחר השימוש בספריי B.
השערת H_0: $\mu_A = \mu_B = \mu_C = \mu_D = \mu_E = \mu_F = \mu$
השערת H_1: $\sim \mu$
<u>ביצוע המבחן וניתוח התוצאות:</u>
היות וגדעון משווה את כלל הקבוצות, נעזר במבחן F על מנת לקבוע האם ניתן לדחות את השערת האפס.
תוצאות ההרצה ב־R:

```
            Df Sum Sq Mean Sq F value Pr(>F)
spray        5   2669   533.8    34.7 <2e-16 ***
Residuals   66   1015    15.4
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

נשים לב שקיבלנו כי ערך ה־p_value (ערך הדחייה) מקיים $***$־ כלומר הינו קטן מ־ 0.001.
היות ומדובר בהסתברות לקבל את ה־data הנ"ל תחת השערת האפס ־ נקבל כי ההסתברות נמוכה מערך הסף $\alpha = 5\%$.
ולכן נוכל לדחות את השערת האפס.
מסקנה ־ לא כל תכשירי הספריי עובדים באותה היעילות.

סעיף ג'

השוואה בין שני הסעיפים:

- השערת האפס ־ ראשית נציין כי בשני המקרים, השערת האפס כללה את האפשרות שהתכשירים המשתתפים עובדים באותה היעילות.
- השוואת המבחן אשר נבחר להשתמש בו:
  - בסעיף א', היות ורצינו להכריע ביעילות בין תוצאות של 2 קבוצות בלבד, יכולנו להשתמש במבחן t חד זנבי עם השערת H_1 ספציפית לקביעת היעילות הטובה יותר בין תוצאות 2 הקבוצות.
  - בסעיף ב', היות ורצינו לבדוק את 6 הקבוצות, היה עלינו להשתמש במבחן $anova$, אשר מאפשר לנו לבדוק האם קיים הבדל ברמת היעילות של כל 6 התכשירים, ללא יכולת להכריע בין תכשירים ספציפיים.
- תוצאות המבחנים:
  - בסעיף א' לא יכולנו לדחות את השערת האפס, כלומר לא נוכל לטעון כי התכשיר A פחות יעיל מתכשיר B, וזאת בשל ערך הדחייה שקיבלנו.
    מעבר לכך, הדפסת הממוצעים מראה כי אכן מהמדגם הנתון, השערת H_1 אינה תקפה.
  - בסעיף ב', דחינו את השערת האפס עבור $\alpha = 5\%$ ועפ"י התוצאות יכולנו לדחות את השערת האפס עבור $\alpha = 1\%$ גם כן. כלומר, נוכל לטעון כי ששת התכשירים לא עובדים ביעילות דומה.
    כמו כן, הדפסת הממוצעים מראה לדוגמה כי עבור תכשיר C , ממוצע כמות החרקים הנותרים הוא נמוך באופן משמעותי ביחס לתכשיר A.

שאלה 4:

מוטי חשב שישנה טעות בחישוביה של דפנה בשל הערך $F = 0.79$ שנחשב מאוד נמוך ולא סביר. זאת בגלל שהרי הערך F מחושב על ידי המנה $\frac{MSB}{MSW}$ כאשר ערך 0.79 מסמל כי ה־ MSW היה גדול יותר מה־ MSB בצורה משמעותית. כלומר לפי הערך המתקבל נראה כי השונות בין הקבוצות קטנה מהשונות בתוך הקבוצות. זאת תוצאה פחות סבירה כיוון שכאשר יש שוני בין הקבוצות יתקבל ערך גבוה של F (גבוה משמעותית מ־ 1) וכאשר אין שינוי בין הקבוצות וכולן בעצם מאותה התפלגות אנחנו נקבל ערך מאוד קרוב ל־ 1. בכל מצב לא נצפה לערך קטן משמעותית מ־ 1 ולכן סביר שזאת טעות חישוב.

שאלה 5:

נשים לב כי הטעות שהוא עשה העלתה את פרמטר דרגות החופש המחושב על ידי $n_x + n_y - 2$. במקום שהוא יהיה 28 הוא הפך להיות 58.

- א' נכון כי־ רווח בר סמך מחושב על ידי

$$P((\overline{x}-\overline{y}) - t^{1-a/2}_{n_x+n_y-2} * \hat{S}_{pooled}\sqrt{\frac{1}{n_x}+\frac{1}{n_y}} \leq \boldsymbol{\mu_x} - \boldsymbol{\mu_y} \leq (\overline{x}-\overline{y}) + t^{1-a/2}_{n_x+n_y-2} * \hat{S}_{pooled}\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}) = 1-\alpha$$

  נשים לב כי $\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}$ עם הנתונים החדשים **קטן** יותר מהערך האמיתי שלו (כי אנחנו מגדילים את המכנה אז השבר קטן), ואז סה"כ המכפלה עם $\hat{S}_{pooled}$ קטנה ורווח הבר סמך קטן.

- ב' נכון כי־ ככל שמספר התצפיות גדול יותר, כך האומדן שלנו מדויק יותר והתפלגות t צרה יותר ושואפת לנורמלית. מכאן נובע כי העלת דרגות החופש מקטינה את הערך הקריטי.

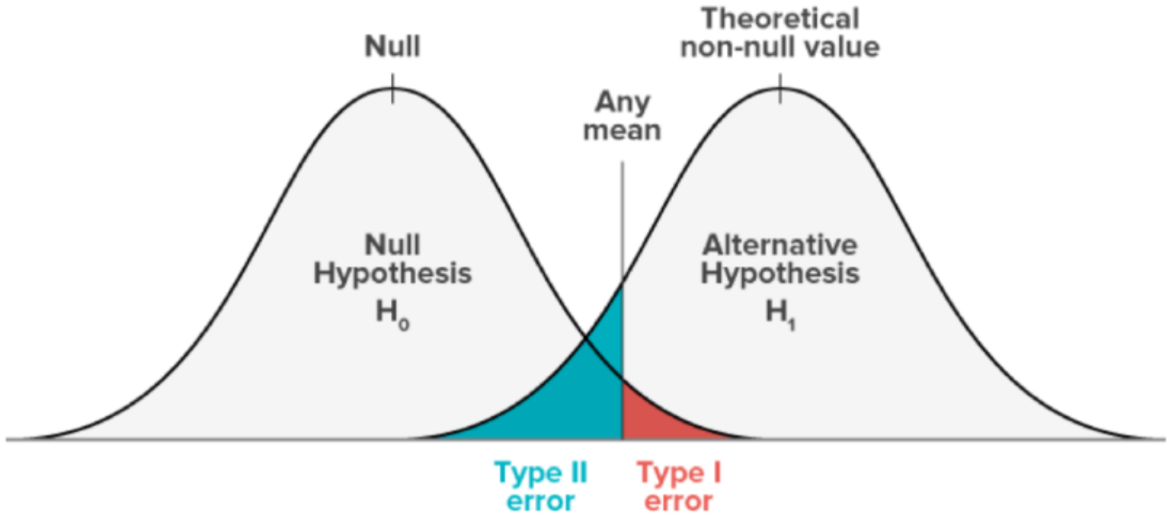
- ג' נכון כי־ הערך ה־t מחושב על ידי הנוסחה $\frac{(\bar{x}-\bar{y})}{\hat{S}_{pooled}\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}} \sim t_{n_x+n_y-2}$.

  כמו שאמרנו קודם הביטוי בשורש קטן ועל כן המכנה של הביטוי קטן־ והמנה הכוללת גדלה. על כן הערך של t גדל.

  לכן **התשובה הנכונה היא ד** כל התשובות נכונות.

שאלה 6:

נכון. בטעות מסוג שני לא דחינו את השערת האפס למרות שהשערת המחקר היא נכונה. כלומר H_1 נכונה אבל לא הצלחנו לדחות את השערת ה־ 0. טעות זו תלויה בשטח החפיפה בין ההתפלגויות:



ככל ששטח החפיפה בין ההתפלגויות גדול יותר, כך יש סיכוי גבוה יותר לטעות מסוג 2, כלומר למצב שבו הדגימות שלנו יפלו באיזור החפיפה (משמאל לאיזור הדחייה) ולא נוכל לדחות את השערת ה־0. לכן ככל שאחוז החפיפה בין ההתפלגויות יהיה קטן יותר כך עוצמת המבחן תהיה גדולה יותר.

* עוצמת המבחן היא בעצם $1-\beta$ כאשר β זה הסיכוי לטעות מסוג 2.

כיצב ניתן להקטין את אחוז החפיפה בין ההתפלגויות? להגדיל את המדגם. ככל שהמדגם גדל כך העוצמה גדלה, זאת משום שסטיית התקן של התפלגות הדגימה קטנה ככל שגודל המדגם גדל, ולכן החפיפה בין ההתפלגויות קטנה.