

סטטיסטיקה ב

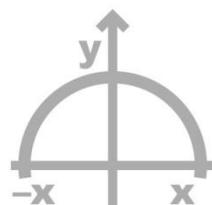


$$\begin{matrix} 1 & \sqrt{2} \\ \diagdown & \diagup \\ 1 & 1 \end{matrix}$$



$$\begin{matrix} + & - & 0 \\ \diagup & \diagdown & \diagdown \\ 0 & 0 & 0 \end{matrix}$$

$$\{\sqrt{x}\}^2$$



תוכן העניינים

1.	הסקה סטטיסטית - הקדמה	1
4.	התפלגות הדגימה ומשפט הגבול המרכזי	4
14.	מושגי יסוד באמידה	14
19.	רוחת סמך לתוחלת (ממוצע)	19
29.	מבוא לבדיקה השערות על פרמטרים	29
35.	בדיקה השערות על תוחלת (ממוצע)	35
53.	רוחת סמך להפרש תוחלות (ממוצעים) במדגים בלתי תלויים	53
57.	בדיקה השערות על הפרש תוחלות במדגים בלתי תלויים	57
65.	רוחת סמך לתוחלת (ממוצע) ההפרשים במדגים מזוגגים	65
67.	בדיקה השערות לתוחלת ההפרש במדגים מזוגגים	67
71.	הקשר בין רוחת סמך לבדיקה השערות להפרש תוחלות	71
74.	רוחת סמך לפרוורציה	74
80.	בדיקה השערות על פרוורציה	80
87.	רוחת סמך להפרש פרוורציות	87
89.	בדיקה השערות על הפרש פרוורציות	89
93.	שאלות מסכימות על רוחתי סמך	93
96.	שאלות מסכימות בבדיקה השערות	96
109.	מבחני חי ברייבור	109
123.	ניתוח שונות חד כיוונית	123

سطויסטיקה ב

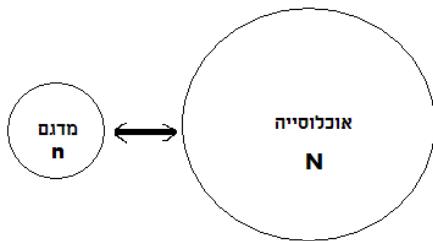
פרק 1 - הסקה סטטיסטית - הקדמה

תוכן העניינים

1. כללי

הסקה סטטיסטית – הקדמה:

רקע:



אוכלוסייה: קבוצה שאליה מפנים שאלת מחקרית. למשל, חברת תרופות שמעוניינת לפתח תרופה למחלת הסוכרת מתעניינת באוכלוסייה חולית הסוכרת בעולם.

מבחן:

חלק מתוך האוכלוסייה.
למשל, אם נדגום באקראי 10 אנשים מתוך חולית הסוכרת אז זהו מבחן מתוך אוכלוסיית חולית הסוכרת.

במקרים רבים אין אפשרות לחקור את כל האוכלוסייה כיון שאין גישה לכללה, היא גדולה מדי, אנו מוגבלים בזמן ובאמצעים טכניים ולכן מבצעים מבחן במטרה לבצע הסקה סטטיסטית מהמבחן לאוכלוסייה.
הדגימה בקורס תהיה דגימה מקראית - הכוונה לדוגמה שבה לכל תצפית באוכלוסייה יש את אותו סיכוי להיכל במדגם.

סטטיסטי:

גודל המוחושב על המבחן.

פרמטר:

גודל המתאר את האוכלוסייה.

הסימונים לפרמטר וסטטיטיסטי הם שונים:

סטטיטיסטי (מדגם)	פרמטר (אוכלוסייה)	
μ	\bar{X}	משמעות
P	\hat{p}	פרופורציה (שכיחות יחסית)

פרמטר הוא גודל קבוע גם אם אנו לא יודעים אותו סטטיטיסטי הוא משתנה ממוגן למדגם ולכן יש לו התפלגות הנקראות התפלגות הדגימה.

דוגמה (פתרון בהקלטה):

62.5% מאזרחי המדינה תומכים בהצעת החוק של חבר הכנסת מסוים.

הוחלט לדגום 200 אזרחים ומתוכם לבדוק מהו אחוז התומכים בהצעת החוק.

א. מי האוכלוסייה?

ב. מה המשתנה?

ג. מה הפרמטרים?

ד. מהו גודל המדגם?

ה. מהו הסטטיטיסטי שמתכוונים להוציא ממדגם?

ו. האם הפרמטר או הסטטיטיסטי הוא משתנה מקרי?

שאלות:

- 1)** מתווך כלל הסטודנטים במכללה שסיוומו סטטיסטיקה א' נדגמו שני סטודנטים. נתון שסכום הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מי האוכלוסייה?
 - מה המשטנה?
 - מהם הפרמטרים?
 - מהו גודל המדגם?
- 2)** להלן התפלגות מספר מקלט טלויזיה למשפחה בישוב "העוגן".
נגידר את X להיות מספר המקלטים של משפחה אקראית. מתכנים לדגום מאוכלוסייה זו 4 משפחות ולהתבונן במספר מקלט טלויזיה במדגם.
- מיי האוכלוסייה ומהו המשטנה הנחקר?
 - מהו הסטטיסטי שיילקח מהדגם ומה סימונו?

מספר משפחות	מספר מקלטים
0	50
1	250
2	350
3	300
4	50
	סך הכל $N = 1000$

- 3)** נתון כי 20% מהשכירים במדינה הם אקדמיים. נבחרו באקראי 10 שכירים באותו אוכלוסייה ומתכנים לפרש את מספר האקדמאים שנדגמו.
- מיי האוכלוסייה?
 - מה המשטנה באוכלוסייה?
 - מהם הפרמטרים?
 - מהו הסטטיסטי?

תשובות סופיות:

- 1)** א. כלל הסטודנטים במכללה שסיימו סטטיסטיקה א'. ב. ציון.
ג. ממוצע: 78, סטיית תקן: 15. ד. 2.
- 2)** א. האוכלוסייה: 1000 משפחות בישוב העוגן, המשטנה הנחקר: מס' מקלטים.
ב. $\bar{X} = \text{ממוצע מדגם}$.
- 3)** א. השכירים במדינה.
ב. השכלה: אקדמי, לא אקדמי.
ג. מס' האקדמאים באוכלוסייה: 0.2.

סטטיסטיקה ב

פרק 2 - התפלגות הדגימה ומשפט הגבול המרכזי

תוכן העניינים

1. התפלגות ממוצע המדגם ומשפט הגבול המרכזי	4
2. התפלגות פרופורציית ההצלחות במדגם	10

התפלגות ממוצע המדגם ומשפט הגבול המרכזי:

רקע:

בפרק זה נדון בהתפלגות של ממוצע המדגם : $\bar{x} = \frac{\sum x_i}{n}$.
 מכיוון שמדובר במדגם אנו יכולים לקבל ממוצע מדגם שונה, אזי ממוצע המדגם הוא משתנה מקרי ויש לו התפלגות.
 גורמים המتأרים בהתפלגות כלשהי או אוכלוסייה כלשהי נקראים פרמטרים.
 להלן רישימה של פרמטרים החשובים לפרק זה:
 ממוצע האוכלוסייה נסמן ב- μ (נקרא גם תוחלת).
 שונות אוכלוסייה נסמן ב- σ^2 .
 סטיית תקן של אוכלוסייה: σ .

תכונות ההתפלגות:

ממוצע כל ממוצעי המדגם האפשריים שווה לממוצע האוכלוסייה: $E(\bar{x}) = \mu_{\bar{x}} = \mu$.
 שונות כל ממוצעי המדגם האפשריים שווה לשונות האוכלוסייה מחולק ב- n .

$$\text{תמונה זו נcona רק במדד מקרי: } V(\bar{x}) = \frac{\sigma^2}{n}.$$

ישיחס הפוך בין גודל המדגם לבין שונות ממוצעי המדגם.
 אם נוציא שורש לשונות נקבל סטיית תקן של ממוצע המדגם שנקרה גם

$$\text{טעות תקן: } \sigma(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}.$$

דוגמה (פתרון בהקלטה) :

השכר הממוצע במשק הינו 9000 לפ עם סטיית תקן של 4000. דגמו באקראי 25 עובדים.

א. מייהי אוכלוסיית המחקר? מהו המשתנה הנחקר?

ב. מהם הפרמטרים של האוכלוסייה?

ג. מה התוחלת ומהי סטיית התקן של ממוצע המדגם?

דוגמה מההתפלגות נורמללית:

אם נדגם מתוך אוכלוסייה שהמשתנה בה מתפלג נורמלית עם ממוצע μ ושונות σ^2 .

$$\text{ממוצע המדגם גם יתפלג נורמלית: } \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right), Z_{\bar{x}} = \frac{\bar{x} - \mu}{\sigma}$$

דוגמה (פתרון בהקלטה):

משקל תינוק ביום הiolדו מתפלג נורמלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.

מה ההסתברות שבמדגם של 4 תינוקות אקראיים בעת הולדתם המשקל הממוצע של התינוקות יהיה מתחת ל-3.5 ק"ג?

משפט הגבול המרכזי:

אם אוכלוסייה מתפלגת כלשהו עם ממוצע μ ושונות σ^2 אז עבור מדגם מספיק

$$\text{גדול } (n \geq 30) \text{ ממוצע המדגם מתפלג בקרוב לנורמל}: \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

דוגמה (פתרון בהקלטה):

משקל חפיסת שוקולד בכו ייצור מתפלג עם ממוצע 100 גרם וסטיית תקן של 4 גרם.

דגמו מכו הייצור 36 חפיסות שוקולד אקראיות.

מה ההסתברות שהמשקל הממוצע של חפיסות השוקולד שנדגוño יהיה מתחת ל-102 גרם?

שאלות:

- 1)** מתווך כל הסטודנטים במכלה שסימנו סטטיסטיקה א' נדגמו שני סטודנטים.
 נתון שסכום הציונים של כל הסטודנטים היה 78 עם סטיית תקן של 15.
- מיהי האוכלוסייה?
 - מה המשנה?
 - מהם הפרמטרים?
 - מהו גודל המדגם?
 - מהו תוחלת ממוצע המדגים?
 - מהי טעות התקן?
- 2)** משקל תינוק ביום היולדו מתפלג נורמללית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם.
 א. מה ההסתברות שתינוק אكري בעת הלידה ישקל פחות מ-3800 גרם?
 נתון כי ביום מסוים נולדו 4 תינוקות.
 ב. מה ההסתברות שהמשקל הממוצע שלהם עלה על 4 ק"ג?
 ג. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-2.5 ק"ג?
 ד. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת
 ללא יותר מ-50 גרם?
 ה. הסבירו לא חישוב כיצד התשובה לשיער הקודם הייתה משתנה אם היה
 מדובר על יותר מ-4 תינוקות?
- 3)** הגובה של המתגיסים לצה"ל מתפלג נורמללית עם תוחלת של 175 ס"מ
 וסטיית תקן של 10 ס"מ. ביום מסוים התגיסו 16 חיילים.
- מה ההסתברות שהגובה הממוצע שלהם יהיה לפחות 190 ס"מ?
 - מה ההסתברות שהגובה הממוצע שלהם יהיה בדיק 180 ס"מ?
 - מה ההסתברות שהגובה הממוצע שלהם יסטה מהתוחלת הגבוהים בפחות
 מ-5 ס"מ?
 - מהו הגובה שבהסתברות של 90% הגובה הממוצע של המדגים יהיה
 נמוך ממנו?

- 4) הזמן הממוצע שלוקח לאדם להגיע לעבודתו 30 דקות עם שונות של 16 דקות רבעות. האדם נושא לעובדה במשך שבוע 5 פעמים. לצורך הפתרון הניחו שזמן הנסעה לעובדה מתפלג נורמאלית.
- א. מה ההסתברות שבמשך שבוע משך הנסעה הממוצע יהיה מעל 33 דקות?
- ב. מהו הזמן שהסתברות של 90% ממוצע משך הנסעה השבועי יהיה גבוה ממוני?
- ג. מה ההסתברות שממוצע משך הנסעה השבועי יהיה מרוחק מ-30 דקות לפחות 2 דקות?
- ד. כיצד התשובה לסעיף הקודם הייתה משתנה אם האדם היה נושא לעובדה 6 פעמים בשבוע?
- 5) נפח היין בבקבוק מתפלג נורמאלית עם תוחלת של 750 סמ"ק וסטיית תקן של 10 סמ"ק.
- א. בארכז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארכז יהיה בדיק 755 סמ"ק?
- ב. בארכז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארכז יהיה יותר מ-755 סמ"ק?
- ג. בארכז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארכז יהיה לפחות 755 סמ"ק?
- ד. בקבוקי היין שבארצנו נציגים לקורה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקורה?
- 6) משתנה מתפלג נורמאלית עם תוחלת 80 וסטיית תקן 4.
- א. מה ההסתברות שממוצע המדגם יסטה מהתוחלתו ללא יותר מichiיה כאשר גודל המדגם הוא 9?
- ב. מה ההסתברות שממוצע המדגם יסטה מהתוחלתו ללא יותר מichiיה כאשר גודל המדגם הוא 16?
- הסביר את ההבדל בתשובות של שני הטעיפים.
- 7) לפי הערכות הלשכה המרכזית לסטטיסטיקה השכר הממוצע במשק הוא 8000 ₪ עם סטיית תקן של 3000 ₪. מה ההסתברות שבמדגם מקרי של 100 עובדים השכר הממוצע יהיה יותר מ-8500 ₪?

8) אורך צינור שמבצע מייצר הינו עם ממוצע של 70 ס"מ וסטיית תקן של 10 ס"מ.

- א. נלקחו באקריאי 100 מוטות, מה ההסתברות שסכום אורך המוטות יהיה בין 68 ל 78 ס"מ?

ב. יש לחבר 2 בניינים באמצעות מוטות. המרחק בין שני הבניינים הינו 7200 ס"מ. מה ההסתברות ש 100 המוטות יספיקו למלאה?

- ג. מה צריך להיות גודל המדגם המינימאלי, כדי שהסתברות של 5% ממוצע המדגם יהיה קטן מ-69 ס"מ. הייערו במשפט הגבול המרכזי.

9) נתון $\bar{X} \sim N(\mu, \sigma^2)$. דגמו 5 תצפיות מאותה ההתפלגות והתבוננו בממוצע המדגם \bar{X} . לכן: $P(\bar{X} > \mu)$ יהיה (בחרו בתשובה הנכונה):

- א. 0.
- ב. 0.5.
- ג. 1.
- ד. לא ניתן לדעת.

10) נתון \bar{X} מתפלג כלשהו עם תוחלת μ ושונות σ^2 .

החליטו לבצע מדגם בגודל 200 מתוך ההתפלגות הנתונה לפי משפט הגבול המרכזי מתקיים (בחרו בתשובה הנכונה):

$$\text{א. } X \sim N\left(\mu, \frac{\sigma^2}{200}\right)$$

$$\text{ב. } \mu \sim N\left(\mu, \frac{\sigma^2}{200}\right)$$

$$\text{ג. } \bar{X} \sim N(\mu, \sigma^2)$$

$$\text{ד. } \bar{X} \sim N\left(\mu, \frac{\sigma^2}{200}\right)$$

11) נתון $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$. אם נדgos n תצפיות מתוך ההתפלגות ונגיד: איזי (בחרו בתשובה הנכונה):

- א. μ ו- \bar{X} יהיו משתנים מקרים.
- ב. μ יהיה משתנה מקרי ו- \bar{X} קבוע.
- ג. \bar{X} יהיה משתנה מקרי ו- μ קבוע.
- ד. μ ו- \bar{X} יהיו קבועים.

תשובות סופיות:

- 1) א. כלל הסטודנטים במללה שסימנו סטטיסטיקה א. ב. ציון. ד. 2. ג. ממוצע : 78, סטיית תקן : 15.
- .10.6 .1 .0.1974. 0.2 .0.0013. ב. 0.8413. (2)
- .178.205. ד. 0.9544. ג. 0. 0. א. 0. (3)
- ד. התשובה הייתה קטנה. 0.2628. ג. 27.71. ב. 0.0465. (4)
- .0.5. ד. 0.1587. ג. 0.1587. ב. 0. א. 0. (5)
- .0.6826. ב. 0.5468. (6)
- .0.0475. (7)
- .271. ג. 0.0228. ב. 0.9772. (8)
- . (9) ב'. (10) ד'. (11) ג'.

התפלגות פרופורציית ההצלחות במדגם:

רקע:

בפרק זה נדון בהתפלגות הדגימה של פרופורציית המדגם.
Y - מספר ההצלחות במדגם (למשל, מספר המובטלים במדגם).

$$\hat{p} = \frac{y}{n} - \text{פרופורציית ההצלחות במדגם.}$$

למשל, שיעור המובטלים במדגם - $n = 200$ -
מספר המובטלים : $Y = 20$.

$$\text{פרופורציית המובטלים במדגם : } \hat{p} = \frac{20}{200} = 0.1$$

נסמן ב- p את שיעור ההצלחה באוכלוסייה וב- q את שיעור הכשלונות באוכלוסייה.
נבע מדגם מקרי (הנחה שהתצפויות בלתי תלויות זו בזו) ונتابון בהתפלגות של
פרופורציית המדגם.

התוחלת, השונות וסטיית התקן של פרופורציית המדגם:

$$E(\hat{p}) = p , V(\hat{P}) = \frac{pq}{n}$$

משפט הגבול המרכזי עבור הפרופורציה המדגמית :

$$\text{אם : } Z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} . \hat{p} \sim N\left(p, \frac{pq}{n}\right) , np \geq 5 \& nq \geq 5 , \text{ אז :}$$

הערות:

- התנאים לקרוב הנורמלי הם נזילים, כלומר משתנים ממרכה אחד לשני.
התנאי שהציגי כאן הוא הפופולרי ביותר :

$$1. n \cdot p \geq 5$$

$$2. n \cdot (1-p) \geq 5$$

- ישנו מרצים שנוטנים את התנאי המחייב הבא :

$$1. n \cdot p \geq 10$$

$$2. n \cdot (1-p) \geq 10$$

- וישנו מרצים המשמשים בתנאי: $(n \geq 30)$.

- תאלצו לבדוק מהו התנאי שנותנו לכם בכיתה כדי לעبور לנורמלית.

- כיוון שפרופורציה אינה חייבות להיות מספרשלם בהכרח לא נהוג לבצע כאן תיקון רציפות.

דוגמה (פתרון בהקלטה):

לפי נתוני משרד החינוך בעיר ירושלים ל- 60% מתלמידי התיכון זכאים לתעודת בגרות. נדגו 200 תלמידי תיכון.

- מה ההסתברות שהשכיחות היחסית (\hat{p}) של הזכאים לבגרות במדגם עלה על 60% ?
- מה ההסתברות שפרופורציות הזכאים לבגרות במדגם עלה על 70% ?

שאלות:

- 1)** במדינה מסוימת 10% מכלל האוכלוסייה הינם מובטלים. נדגוו באקראי 140 אנשים מהמדינה.
- מה התוחלת ומהי השונות של פרופורציות המובטלים שנדגוו?
 - מה ההסתברות שבמדגם לפחות 10% יהיו מובטלים?
 - מה ההסתברות שלכל היוטר 9% מהמדגם יהיו מובטלים?
- 2)** נניח כי 30% מהאוכלוסייה תומכים בהצעת חוק מסוימת. אם נדגוו מהאוכלוסייה 200 איש. חשבו את ההסתברויות הבאות :
- פחות 35% יתמכו בהצעת החוק במדגם.
 - לכל היהת 25% יתמכו בהצעת החוק במדגם.
 - יוטר מ-27% יתמכו בהצעת החוק במדגם.
- 3)** לפי נתוני משרד התקשרות 40% מהאוכלוסייה מחזיקים בטלפון נייד מסווג "סמארטפון". נדגוו 400 אנשים מהאוכלוסייה.
- מה ההסתברות שבמדגם לכל היוטר ל-40% יש סמארטפון?
 - מה ההסתברות שבמדגם לרוב יש סמארטפון?
 - מה ההסתברות שפרופורציית בעלי הסמארטפון במדגם תסטה מהפרופורציה באוכלוסייה ללא יותר מ-4%?
 - כיצד התשובה לשיעיף הקודם הייתה משתנה אם היינו מגדילים את גודל המדגם?
- 4)** נתון כי 80% מבתי האב מחוברים לאינטרנט. נדגוו 400 בתים אב אקרים.
- מה ההסתברות שלפחות 340 מהם מחוברים לאינטרנט?
 - מה ההסתברות שפרופורציית המחוברים לאינטרנט במדגם תסטה מהפרופורציה האמיתית ביוטר מ-4%?
 - כמה בתים אב יש לדוגם כדי שהסתטיה בין הפרופורציה המדגמית לפרופורציה האמיתית לא תעלה על 3% בהסתברות של 90%?
 - מהו העשירון התיכון של התפלגות פרופורציית המדגם?
- 5)** נתון שציוני פסיקומטרי מתפלגים נורמלית עם תוחלת 500 וסטיית תקן 100. ב"מועדון ה-700" נכללים נבחנים שמקבלים ציון מעל 700 בפסיכומטרי. מה הסיכוי שבמועדון בו נבחנו 2000 נבחנים אקרים יהיו לפחות 3% המשתייכים למועדון?

6) נתון ש- $\hat{P} = \frac{X}{n}$, ונגידיר את המשתנה הבא :

$$\text{א. הוכיחו ש } E(\hat{P}) = p, V(\hat{P}) = \frac{p(1-p)}{n}$$

ב. מה p המביא את $V(\hat{P})$ להיות מקסימום?

תשובות סופיות:

1) א. התוחלת : 0.1, השונות : 0.00064. ב. 0.5.

2) א. 0.0618. ב. 0.0618. ג. 0.8238.

3) א. 0.5. ב. 0.8968. ג. 0.8.

4) א. 0.0062. ב. 0.0456. ג. 0.481. ד. 0.77436.

5) 0.0154

6) א. שאלת הוכחה. ב. 0.5.

סטטיסטיקה ב

פרק 3 - מושגי יסוד באמידה

תוכן העניינים

1. כללי

14

מושגי יסוד באמידה:

רקע:

כזכור מהפגש הקודם, פרמטר הוא גודל המתאר את האוכלוסייה או התפלגות מסויימת. כמו ממוצע הגבאים בקרוב מתגisiים לצה"ל - μ .

כמו פרופורצית התומכים במשלה בקרוב אזרחי המדינה - p .

בדרכ כל הפרמטרים הם גדלים שאינם ידועים באמת, ולכן מוצאים מוגדים במטרה לאמוד אותם. אין אפשרות לחשב אותם הניסיון הוא בהערכתו כמה הם שווים ככל שניתן.

- נסמן באופן כללי פרמטר באות θ ואומד ב- $\hat{\theta}$. $\hat{\theta}$ הוא סטטיסטי המוחשב על המוגדים ובאמצעותו נאמוד את θ .
- שגיאת אמידה: $|\hat{\theta} - \theta|$ - ההפרש בין האומד לאמת (הפרמטר).

דוגמה (פתרון בהקלטה):

בכנסת ה-19 קיבלת מפלגת העבודה 15 מנדטים. בערוץ 10 ברגע סגירת הקלפיות הערכו את מספר המנדטים של המפלגה להיות 17 מנדטים וזאת על סמך תוצאות מוגדים של העורץ.

- מה הפרמטר בדוגמה זו?
 - מהי טעות האמידה של ערוץ 10?
- $\hat{\theta}$ יהיה אומד חסר הטיה ל- θ אם התוחלת של $\hat{\theta}$ תהיה שווה ל- θ : $E(\hat{\theta}) = \theta$.
 - טעות התקן של אומד היא סטיית התקן שלו, כלומר: $\sigma(\hat{\theta}) = S.E.$

פרמטרים מרכזיים ואומדיים שלחה:**ממוצע האוכלוסייה μ :**

$$\text{האומד הנקודתי שלו יהיה: ממוצע המדגמים: } \bar{x} = \frac{\sum x}{n}$$

$$\text{. } \sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}} = SE \text{ . } E(\bar{x}) = \mu \text{ . } \text{כמו כן, טעות התקן: } \mu$$

פרופורציה באוכלוסייה p :

$$\text{האומד הנקודתי שלו יהיה: פרופורציה במדגם: } \hat{p} = \frac{y}{n}$$

$$\text{. } \sigma(\hat{p}) = \sqrt{\frac{p \cdot (1-p)}{n}} \text{ . } E(\hat{p}) = p \text{ . } \text{כמו כן טעות התקן: } p$$

שונות האוכלוסייה σ^2 :

$$\text{האומד הנקודתי שלו יהיה: } S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$\text{. } S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} \text{ . } \sigma^2 \text{ . } E(S^2) = \sigma^2 \text{ . } \text{ולכן } S^2 \text{ הינו אומד חסר הטיה ל- } \sigma^2$$

הערה: אומד הוא הנוסחה הכללית לאמידת הפרמטר ואומדן הוא הערך הספציפי שהתקבל במדגם מסוים.

דוגמה (פתרון בהקלטה):

נדגמו 10 משפחות בתל אביב ונבדק עבור כל משפחה מספר הילדים שלה.

להלן התוצאות שהתקבלו: 1, 3, 2, 1, 4, 5, 2, 1, 3, 2, 1.

אמדו באמצעות אומדיים חסרי הטיה את הפרמטרים הבאים:

1. ממוצע מספר הילדים למשפחה בתל אביב.
2. שונות מספר הילדים למשפחה בתל אביב.
3. פרופורציית המשפחות בנות שני ילדים.

שאלות:

- 1)** מתוך 500 טירונים, נמצאו 120 בעלי שברי הליכה. נתנו שהטיסוי שטירון יהיה עם שבר הליכה הוא 0.25.
- מהי האוכלוסייה המוצגת בשאלת? מהם הפרמטרים שלה?
 - מהי טעות התקן של האומד כשהמדגם בגודל 500?
 - מהו האומדן לפרמטר?
 - מהי טעות האמידה?
- 2)** לפי נתונים היכרנו, מקרר צורך ממוצע 2400 וואט לשעה עם סטיית התקן של 500 וואט לשעה.
- במדגם של 25 מקרים של היכרן התקבל ממוצע של 2342 וואט לשעה.
- מהי האוכלוסייה המוצגת בשאלת? מהם הפרמטרים שלה?
 - מהי טעות התקן של האומד?
 - מהו האומדן לפרמטר?
 - מהי טעות האמידה?
- 3)** נדגו עשרה מתגיים לzech"l. גובהם נמדד בס"מ. להלן התוצאות שהתקבלו: 168, 184, 192, 180, 171, 177, 187, 168, 177 ו-175.
- מצאו אומדן חסר הטיה לגובה הממוצע של מתגייסי zech"l.
 - מצאו אומדן חסר הטיה לשונות הגבהים של מתגייסי zech"l.
 - מצאו אומדן חסר הטיה לפ羅פּוֹרְצִיוֹת המתגיים בגובה של לפחות 180 ס"מ.
- 4)** נדגו 20 שכירים באקראי. עברו כל שכיר נמדד השכר באלפי שקלים.
- להלן התוצאות שהתקבלו: $\sum_{i=1}^{20} X_i^2 = 1502.2$, $\sum_{i=1}^{20} X_i = 162$
- AMDו את השכר הממוצע של השכירים במשק.
 - AMDו את סטיית התקן של שכר השכירים במשק.
- 5)** במטרה לאמוד את ממוצע האוכלוסייה, נדגו תכיפות בלתי תלויות מהאוכלוסייה וחישבו את הממוצע שלהם. מהי טעות התקן?
- סטיית התקן של האוכלוסייה.
 - סטיית התקן של ממוצע האוכלוסייה.
 - סטיית התקן של המדגם.
 - סטיית התקן של ממוצע המדגם.

6) משקל הממוצע של אוכלוסייה מסוימת הוא 75 ק"ג עם שונות של 25 .
 אם יבחרו כל המדגמים האפשריים בגודל 10 מאוכלוסייה זו סטיית התקן של ממוצעי המדגמים תהיה :

- .א. 3.
- .ב. 2.5
- .ג. 1.581
- .ד. אין מספיק נתונים לדעת.

7) במדגם מקרי, متى סכום ריבועי הסטיות מהממוצע, $\sum_{i=1}^n (x_i - \bar{x})^2$, מחולק ב- $n-1$?
 א. כאשר n קטן.
 ב. כאשר תצפיות המדגם אינן בלתי תלויות.
 ג. כאשר האוכלוסייה אינה מתפלגת נורמללית.
 ד. כאשר מעוניינים באומד חסר הטיה לשונות האוכלוסייה ממנה הוצאה המדגם.
 ה. כאשר מעוניינים לחשב את שונות התפלגות הדגימה של ממוצע המדגם.

8) מדגם מקרי מתוך אוכלוסייה בעלת ממוצע μ לא ידוע ושונות : $64 = \sigma^2$. טעות התקן של האומד ל- μ היא :

- .א. 16.
- .ב. 8.
- .ג. 4.
- .ד. 2.

9) מהו אומד חסר הטיה?
 א. אומד שערכו שווה לממוצע התפלגות הדגימה שלו.
 ב. אומד שערכו שווה לערך הפרטר באוכלוסייה.
 ג. אומד שממוצע התפלגות הדגימה שלו שווה לערך הפרטר באוכלוסייה.
 ד. אומד שהסיכוי שערכו יהיה גבוה מערך הפרטר באוכלוסייה שווה לשינוי שיהיה נזוק ממנו.

תשובות סופיות:

- א. אוכלוסייה: מקרים של יczron, תוחלת: 2400, סטיית תקן: 500.
 ב. 100.
 ג. 2342.
 ד. .58.

א. 0.25.
 ב. 0.019.
 ג. 0.24.
 ד. 0.01.

א. 177.9.
 ב. 64.1.
 ג. 0.4.
 ד. 3.16.

א. 8.1.
 ב. 3.16.
 ג. 1.6.
 ד. 0.4.

א. 5.
 ב. 6.
 ג. 7.
 ד. 8.
 ז. 9.

סטטיסטיקה ב

פרק 4 - רוח סמרק לתוחלת (ממוצע)

תוכן העניינים

19	1. רוח סמרק כשינוי האוכלוסייה ידועה
24	2. קביעת גודל מוגם
26	3. רוח סמרק כשינוי האוכלוסייה לא ידועה

רוח סמך כשינויות האוכלוסייה ידועה:

רקע:

ממוצע המדגם הוא אומד לממוצע האוכלוסייה, אך לא באמת ניתנו להבין ממנה על גודלו של ממוצע האוכלוסייה. ההסתברות שממוצע המדגם יהיה בדיקות כמו הממוצע האמתי הוא אפסי.

מה שנחוג לעשות כדי לאמוד את ממוצע האוכלוסייה, זה לבנות רוח סמך.

בנייה מרוחה בטחון שהסיכוי שהפרט μ ייכל בתוכו הוא: $1 - \alpha$.

$\alpha - 1$: נקרא רמת בטחון או רמת סמך. כך ש: $\alpha - 1 = P(A \leq \mu \leq B)$.

A - גבול תחתון של רוח הסמך.

B - הגבול העליון של רוח הסמך.

$L = B - A$ - אורך רוח הסמך.

דוגמה (פתרון בהקלטה):

חווק דגם 25 חיילים שנבחנו ב מבחון הפסיכומטרי. הוא בנה רוח סמך לממוצע הציונים ב מבחון הפסיכומטרי ב קרב אוכלוסיית החיילים ו קיבל בין 510 ל-590. רוח הסמך בונה ברמת סמך של 95%.

1. מהי אוכלוסיית המחקר?
2. מה המשתנה באוכלוסייה?
3. מה הפרט שהחווק רצה לאמוד?
4. מהו רוח הסמך?
5. מה אורך רוח הסמך?
6. מהי רמת הביטחון של רוח הסמך?

בפרק זה נרצה לבנות רוח סמך לתוחלת (μ) במקהה ש- σ^2 (שונות האוכלוסייה) ידועה. פרמטרו אותו נרצה לאמוד: μ .

אומד נקודתי: \bar{x} .

תנאים לבניית רוח הסמך: $N \sim X$ או $n \geq 30$.

σ^2 (שונות האוכלוסייה) ידועה.

$$\text{נוסחה לרוח הסמך: } \bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

דוגמה (פתרון בהקלטה):

על פי נתוני היצרן אורך חיי סוללה מתפלג נורמלית עם סטיית תקן של 1 שעה. מעוניינים לאמוד את תוחלת חיי סוללה. נציגו באקראי 4 סוללות, אורך החיים הממוצע שהתקבל הוא 13.5 שעות. בנו רוח סמך ברמת סמך של 95% לתוחלת אורך חיי סוללה.

$$\text{שגיאת האמידה המקסימלית: } \varepsilon = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

ע - נותן את שגיאת האמידה המקסימלית, דבר שנקרה גם טעות סטטיסטית, טעות דגימה.

דוגמה (פתרון בהקלטה):

בשימוש לשאלת עם הסוללות. מה ניתן להגיד בביטחון של 95% על שגיאת האמידה?

קשרים מתמטיים ברוח הסמך:

- אורך רוח הסמך הוא פערם שגיאת האמידה המקסימלית: $L = 2\varepsilon$.
- ממוצע המדגים נופל תמיד באמצע רוח הסמך: $\bar{X} = \frac{A+B}{2}$.
- ככל שמספר התצפיות (n) גבוהה יותר, כך יש יותר אינפורמציה ולכן האומד יותר מדויק, ולכן מקבל רוח סמך יותר קצר.
- ככל שרמת הביטחון ($\alpha-1$) גבוהה יותר, כך: $\varepsilon = z_{1-\frac{\alpha}{2}}$ יותר גבוהה, ורוח הסמך יותר ארוך.

שאלות:

- 1)** חוקר התענין למד את השכר המומוצע במשק. על סמך מוגם הוא קבע שבביטחון של 95% כי השכר המומוצע במשק נع בין 9200 ל-9800 ₪.
- מי האוכלוסייה במחקר?
 - מה המשנה הנחקר?
 - מה הפרמטר שאותו רוצים למד?
 - מה רוח הסמך לפרמטר?
 - מה רמת הסמך לפרמטר?
 - מה אורך רוח הסמך?
 - מה הסיכוי שטעות הדגימה תעלה על 300 ₪?
- 2)** מעוניינים למד את התפוקה היומית המומוצעת של מפעל מסוים ברמת סמך של 95%. בדוגמאות אקראי של 100 ימים התקבלה תפוקה ממוצעת 4950 מוצרים ביום. לצורך פתרון הנח שטית התקן האמצעית ידועה ושויה 150 מוצרים ביום. בנו את רוח הסמך.
- 3)** מעוניינים למד את ממוצע אורך החיים של מכשיר. מנתוני היצרך ידוע שאורך החיים מתפלג נורמלית עם סטיית תקן של 20 שעות. נגמו 25 מכשירים ונמצא כי ממוצע אורך החיים שלהם היה 230 שעות.
- בנו רוח סמך ברמת סמך של 90% לאורך החיים המומוצע של מכשיר.
 - בנו רוח סמך ברמת סמך של 95% לאורך החיים המומוצע של מכשיר.
 - הסבירו כיצד ומדוע השתנה רוח הסמך.
- 4)** נגמו 200 עובדים מהמשק הישראלי. השכר המומוצע שלהם היה 9700 ₪. נניח שטית התקן של השכר במשק היא 3000 ₪.
- בנו רוח סמך ברמת סמך של 95% לתוכלת השכר במשק.
 - מה ניתן לומר בביטחון של 95% על הסטייה המרבית בין ממוצע המוגם לתוכלת השכר?
 - מה היה צריך להיות גודל המוגם אם היו רוצחים להקטין את רוח הסמך ב-50%?
 - אם היינו מגדילים את גודל המוגם ובונים רוח סמך באותה רמת סמך האם היה ניתן לטעון בביטחון רב יותר שרוח הסמך מכיל את הפרמטר?

- 5) בנו רוח סמך לממוצע הציוניים של מבחן אינטלייגנציה. ידוע שסטיטית התקן היא 15 והמדד מtabסס על 100 תוצאות. רוח הסמך שהתקבל הוא (105,99). שזרו את :
- ממוצע המדגמים.
 - שגיאת האמידה המקסימאלית.
 - רמת הסמך.
- 6) זמן החלמה מאנגינה מתפלג עם סטיטית התקן של יומיים. חברת תרופות מעוניינת לחקור אנטיביוטיקה חדשה שהיא פיתחה. במחקר השתתפו 60 אנשים שחלו באנגינה וקיבלו את האנטיביוטיקה החדשה. בממוצע הם החלימו לאחר 4 ימים.
- בנו רוח סמך לתוחלת זמן ההחלמה תחת האנטיביוטיקה החדשה ברמת סמך של 90%.
 - מה הייתה קורה לאורך רוח הסמך אם היה תקציב להגדלת גודל המדגמים פי 4? הסבירו.
 - מה הייתה קורה לאורך רוח הסמך אם היו בונים את רוח הסמך ברמת סמך גדולה יותר? הסבירו.
- 7) חוקר בנה רוח סמך לממוצע וקיבל את רוח הסמך הבא : $\mu = 82$. נתון שסטיטית התקן בהתפלגות שווה ל-10 ושהמדד מtabסס על 16 תוצאות. התפלגות המשתנה היא נורמללית.
- מהו ממוצע המדגמים?
 - מהי רמת הסמך של רוח הסמך שנבנה?
 - מה הסיכוי ששגיאת האמידה באמידת ממוצע האוכלוסייה תעלה על 5?
- 8) חוקר בנה רוח סמך לתוחלת כאשר השונות בהתפלגות ידועה ברמת סמך של 95%. אם החוקר כעת יבנה על סמך אותו נתונים רוח סמך ברמת סמך קטנה מ-95%, איזה מהמשפטים הבאים לא יהיה נכון.
- אורך רוח הסמך החדש יהיה קטן יותר.
 - גודל המדגמים יהיה כעת קטן יותר.
 - הMargin בין ממוצע המדגם לקצota רוח הסמך יהיה קטנים יותר ברוח הסמך החדש.
 - רמת הביטחון לבנות רוח הסמך החדש תהיה קטנה יותר.

9) חוקר בנה רוח סמך ל- μ וקיבל: $48 < \mu < 54$. מה נכון בהכרח:

- א. $\mu = 51$.
- ב. $\bar{X} = 6$.
- ג. $\bar{X} = 51$.
- ד. אורך רוח הסמך הינו 3.

10) אייזה מהגורמים הבאים אינם משפיע על גודלו של רוח בר סמך, כאשר שונות האוכולוסייה ידועה (בחרו בתשובה הנכונה):

- א. רמת הבטיחון.
- ב. סטיית התקן באוכולוסייה.
- ג. מספר המשתתפים.
- ד. סטיית התקן במדגם.

תשובות סופיות:

- 1) א. העובדים במשק. ב. שכר ב-₪.
 ד. $\mu < 9800$. ג. $\mu > 9200$.
 ה. 0.95. ו. 0.600.
- 2) $4979.4 < \mu < 4920.6$
- 3) א. $236.58 < \mu < 237.84$. ב. $223.42 < \mu < 222.16$.
 ג. ראה סרטוון.
- 4) א. $10,116 < \mu < 9284$. ב. הסטייה המרבית בין \bar{x} ל- μ היא 416 נס בבטיחון של 95%.
- 5) א. 0.102. ב. 0.3. ג. 0.9544.
 ד. לא. ו. 0.800.
- 6) א. 4.42 < $\mu < 4.83$. ב. יקטן פי 2. ג. גדול.
- 7) א. 0.87. ב. 0.5. ג. 0.9544.
- 8) א. ב'.
 ג. ז'.
- 9) א'.
 ז'.
- 10) ד'.

קביעת גודל מוגן:

רקע:

אם מעוניינים לאמוד את ממוצע האוכלוסייה כאשר סטיטית התקן של האוכלוסייה ידועה: σ ברמת סמך של $\alpha=1$ ושיגיאת אמידה שלא עלתה על ϵ מסויים, נציב

$$n \geq \left(\frac{z_{1-\frac{\alpha}{2}} \cdot \sigma}{\epsilon} \right)^2$$

בנוסחה הבאה:

כדי להציג בנוסחה צריך שהמשתנה הנחקר يتפלג נורמלית או שהמוגן ייצא בגודל של לפחות 30 תצפיות.

דוגמה (פתרון בהקלטה):

חברת תעופה מעוניינת לאמוד את תוחלת משקל המטען של נוסע. נניח שמשקל מטען של נוסע מתפלג נורמללית עם סטיטית התקן של 2 ק"ג. כמה נוסעים יש לדוגם אם מעוניינים שבבביחוון של 98% הסטייה המרבית בין ממוצע המוגן לממוצע האמתי לא עליה על 0.5 ק"ג? (תשובה: 87).

שאלות:

- (1)** משתנה מקרי מתפלג נורמללית עם סטיטית תקן ידועה 12. מה צריך להיות גודל המדגם כדי לבנות רוח סמך ברמת סמך של 98% שאורכו לא עולה על 2?
- (2)** מעוניינים לאמוד את הדופק הממוצע של מתגייסים לצבאי. מעוניינים שבביטחון של 95% שגיאת האמידה המרבית תהיה 0.5. נניח שהדופק מתפלג נורמלית על סטיטית תקן של 3 פעימות לדקה.
- כמה מתגייסים יש לדוגום?
 - אם ניקח מדגם הגדל פי 4 מהמדד של סעיף א' ונאמוד את הממוצע באותה רמת סמך כיצד הדבר ישפייע על שגיאת האמידה?
- (3)** יהיו X משתנה מקרי עם ממוצע μ וסטיטית תקן σ . חוקר רוצה לבנות רוח בר סמך ל- μ ברמת ביטחון של 0.95, כך שהאורך של הרוח יהיה $\sigma = 0.5\sigma$. מהו גודל המדגם הנדרש?

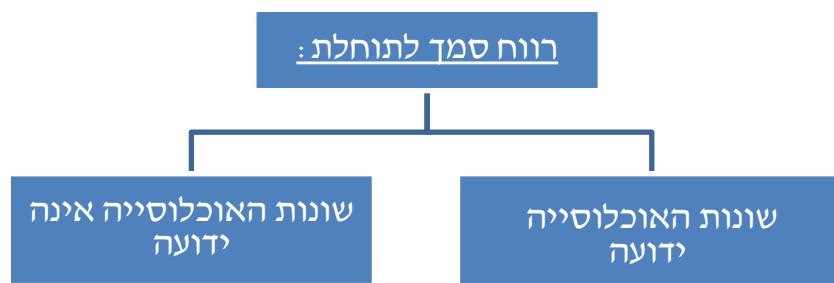
תשובות סופיות:

- (1) .780
 (2) א. 139
 (3) . $n = 62$
- ב. הדבר יקטין את σ פי 2.

רוח סמך כשינויות האוכלוסייה לא ידועה:

רקע:

בבואהנו לבנות רוח סמך לתוחלת אנו צריכים להתמקד בשני המצבים הבאים:



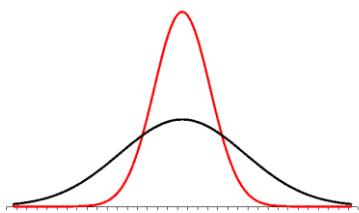
בפרק זה עוסק במקרה **שינויות האוכלוסייה** (σ^2) אינה ידועה לנו.

מקרה יותר פרקטי.

התנאי: $N \sim X$ או שהמדגם גדול.

$$\text{רוח סמך: } \bar{X} \pm t_{\frac{n-\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$$

$$\text{האומד לשונות: } S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$$



התפלגות T:

הינה התפלגות סימטרית בעומוניות שהתוחלת שלה היא 0. ההתפלגות דומה

להתפלגות Z רק שהיא רחבה ולכון הערכים שלה יהיו יותר גבוהים.

התפלגות T תלויות במושג שנקרוא דרגות החופש. דרגות החופש הן: $df = n - 1$.

כל שדרגות החופש עלות התפלגות הופכת להיות יותר גבוהה וצרה.

שדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה (פתרון בהקלטה):

הזמן שלוקח לפתרון שאלה מסוימת ביחסו מתפלג אצל תלמידי כיתות ח' נורמלית.

במטרה לאמוד את תוחלת זמן הפתרון נדגומו 4 תלמידים בכיתה ח'. להלן התוצאות

שהתקבלו בבדיקות: 4.7, 5.2, 4.6, 5.3.

בנו רוח סמך ברמת סמך של 95% למומוצע זמן הפתרון לשאלת קרב תלמידי כיתה ח'.

שאלות:

- 1)** מחקר מעוניין לדעת כיצד תרופה מסוימת משפיעה על קצב פעימות הלב.
ל-5 אנשים שנטלו את התרופה מדדו את הדופק והתקבל מספר פעימות לדקה : 84, 84, 88, 79, 89.
הערה : לצורך פתרון הנח שקצב פעימות הלב מתפלג נורמלית בקירוב.
א. בנו רוח סמך ברמת סמך של 95% לתוחלת הדופק של נוטלי התרופה הניל.
ב. נתון שהדופק הממוצע ללא לקיחת התרופה הינו 70. לאור זאת, האם בביטחון של 95% התרופה משפיעה על הדופק?
ג. בהמשך לסעיף א', אם היינו בונים את רוח הסמך ברמת ביטחון של 99%, כיצד הדבר היה משפיע על רוח הסמך?
- 2)** במדגם שנעשה על 25 מתגייסים לצבא האמריקאי התקבל כי גובה ממוצע של חיל הינו 178 ס"מ עם סטיית תקן : $S = 13$ ס"מ.
בנו רוח סמך ברמת סמך של 90% לתוחלת גובה המתגייסים לצבא האמריקאי. מה יש להניח לצורך פתרון?
- 3)** אדם מעוניין לאמוד את זמן הנסיעה הממוצע שלו לעבודה. לצורך כך הוא דוגם 5 ימים שזמן הנסעה בהם בבדיקות הוא : 27, 34, 32, 40, 30.
א. ברמת ביטחון של 95% אמוד את זמן הנסעה הממוצע. מהי ההנחה הדורשahn לorzuch פתרון?
ב. איך גודל רוח הסמך היה משתנה אם היו דוגמים עוד ימים?
- 4)** ציוני מבחן אינטיליגנציה מתפלגים נורמלית. נדגו 25 מבחנים והתקבל ממוצע ציונים 102 וסטיית תקן מדגמית 13.
א. בנו רוח סמך לממוצע הציונים באוכלוסייה ברמת ביטחון של 95%.
ב. חזרו על סעיף א' אם סטיית התקן הינה סטיית התקן האמתית של כלל הנבחנים.
ג. הסבירו את ההבדלים בין שני השעיפים הניל.
- 5)** נשקלו 60 תינוקות אשר נולדו בשבועות-40 של ההריון. המשקל נמדד בKİLOGRAMS. להלן התוצאות שהתקבלו : $\sum_{i=1}^{60} X_i^2 = 643.19$, $\sum_{i=1}^{60} X_i = 195$.
בנו רוח סמך ברמת סמך של 95% לתוחלת משקל תינוק ביום היולדו.

- 6) נדגו 120 אנשים אקראים מעל גיל 50. עבור כל אדם נבדק מספר שנות השכלתו. להלן התוצאות שהתקבלו : $S = 2$, $\bar{x} = 13.8$.
בנו רוח סמך ברמת סמך של 96% למומוצע ההשכלה של אזרחים מעל גיל 50.
- 7) שני סטטיסטיקים בנו רוח בר-סמך לאותו פרמטר μ .
 לכל אחד מהסטטיסטיקים מוגם אחר, אך באותו גודל 10.
 שניהם קבעו אותה רמת סמך.
 סטטיסטיκי א : הניח $20 = \sigma$.
 סטטיסטיκי ב : חישב לפי המוגם וקיבל $20 = S$.
 למי משני הסטטיסטיקים יהיה רוח סמך ארוך יותר?
 א. סטטיסטיκי א.
 ב. סטטיסטיκי ב.
 ג. אותו ארוך רוח סמך לשני הסטטיסטיקים.
 ד. תלוי בתוצאות המוגם של כל סטטיסטיκי.
- 8) נתון ש : $N(\mu, \sigma^2)$ ביצעו מוגם בגודל 16 וקיבלו סטיית תקן מוגמית 10.
 אורך רוח הסמך שהתקבל הוא : 8.765. מהי רמת הביטחון של רוח הסמך?

תשובות סופיות:

- (1) א. $\mu < 89.72$ ב. $\mu > 79.88$
ג. הוא היה גדול. ב. כן.
ראה בסרטון.
- (2) א. $\mu < 107.37$ ב. $\mu > 96.63$
ב. לא ניתן לדעת. ג. ראה בסרטון.
א. צרייך להניח שהמשתנה מתפלג נורמלית.
- (3) א. $\mu < 107.37$ ב. $\mu > 96.63$
ג. ראה בסרטון.
- (4) $.3.149 < \mu < 3.351$ (5)
 $.13.42 < \mu < 14.18$ (6)
- (6) ב'.
(.90%)

סטטיסטיקה ב

פרק 5 - מבוא לבדיקת השערות על פרמטרים

תוכן העניינים

29	1. הקדמה
33	2. סוגי טעויות

הקדמה:

רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטי. בבדיקה השערות על פרמטרים עוסcid לפיה שלבים הבאים:

שלב א: נזהה את הפרמטר הנחקר.

שלב ב: נרשום את השערות המחקר. השערת האפס המסומנת ב- H_0 .

בדרך כלל השערת האפס מסמלת את אשר היה מקובל עד עכשו, את השגרה הנורמה.

השערה אלטרנטיבית (השערת המחקר) המסומנת ב- H_1 .

ההשערה האלטרנטיבית מסמלת את החדשנות בעצם ההשערה האלטרנטיבית בדברות על הסיבה שהמחקר נעשה היא שאלת המחקר.

שלב ג: נבדוק האם התנאים לביצוע התהליך מתקיימים ונניח הנחות במידת הצורך.

שלב ד: נרשום את כל ההכרעה. בתהליך של בדיקת השערות יוצרים כלל שנקרה כלל הכרעה. הכליל יוצר אзорים שנקרים:

1. **אזור דחיה:**

דחיה של השערת האפס כלומר קבלה של האלטרנטיבית.

2. **אזור קבלה:**

קיבלה של השערת האפס ודחיה של האלטרנטיבית. כלל ההכרעה מתבסס על איזושהו סטטיסטי. אזור הדחיה מוכתב על ידי סיכון שלוקח החוקר מראש שנקרה רמת מובהקות ומסומן ב- α .

שלב ה: בתהליך יש ל选取 תוצאות המדגם וליחס את הסטטיסטי המתאים ולבדוק האם התוצאות נופלות באזור הדחיה או הקבלה.

שלב ו: להסיק מסקנה בהתאם לתוצאות המדגם.

דוגמה (פתרון בהקלטה):

משרד הבריאות פרסם משקל ממוצע של תינוקות ביום לידתם בישראל 3300 גרם. משרד הבריאות רוצה לחזור את הטענה שנשים מעשנות בזמן ההריון יולדות תינוקות במשקל נמוך מהממוצע. במחקר השתתפו 20 נשים מעשנות בהריון. להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:

$$\bar{X} = 3120, S = 280, n = 20.$$

- א. מהי אוכלוסיות המחקר?
- ב. מה המשתנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

שאלות:**בשאלות הבאות, ענו על הטעיפים הבאים:**

- א. מהי אוכלוסיית המחקר?
- ב. מה המשנה הנחקר?
- ג. מה הפרמטר הנחקר?
- ד. מהן השערות המחקר?

- 1)** ממוצע הציונים בבחינת הבגרות באנגלית הנו 72 עם סטיטית תקן 15 נקודות. מורה טוען שפיתחה שיטה לימוד חדשה שתעלה את ממוצע הציונים. משרד החינוך החליט לתת למורה 36 תלמידים אקראים. ממוצע הציונים של אותם תלמידים לאחר לימודו בשיטתו היה 75.5.
- 2)** לפי הצהרת היিירן של חברת משקאות מסויימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיטית תקן 20 סמ"ק. אגודה הרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצחרת. במדוג שעשתה אגודה הרכנים התקבל נפח ממוצע של 492 סמ"ק במדוג בגודל 25.
- 3)** במשך שנים אחדו המועמדים שהתקבלו לפיקולטה למשפטים היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. מחקר מעוניין לבדוק האם השנה מקשים על הקבלה לפיקולטה למשפטים.
- 4)** בחודש ינואר השנה פורסם שאחדו האבטלה במשק הוא 8% במדוג עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים. רוצחים לבדוק ברמת מובהקות של 5% האם אחדו האבטלה הוא כמו בתחילת השנה.

תשובות סופיות:

- ב. ציון.
- 1) א. נבחנים בברירות באנגלית.
 $H_0: \mu = 72$ ג. ממוצע הציונים בשיטת לימוד חדשה.
 $H_1: \mu > 72$
- ב. נפח משקה בקבוק של חברת מסויימת.
- 2) א. משקאות בקבוק של חברת מסויימת.
 $H_0: \mu = 500$ ג. ממוצע נפח המשקה בקבוק.
 $H_1: \mu < 500$
- ב. משתנה דיכוטומי (התקבל, לא התקבל).
- 3) א. מועמדים לפיקולטה למשפטים.
 $H_0: p = 0.25$ ג. אחוז הקבלה.
 $H_1: p < 0.25$
- ב. משתנה דיכוטומי (מובטל, עובד).
- 4) א. אזרחים בוגרים במשק.
 $H_0: p = 0.08$ ג. אחוז האבטלה ביום.
 $H_1: p \neq 0.08$

סוגי טעויות:

רکע:

בתחילת בדיקת השערות יוצרים כלל שנקרא כלל הכרעה. הכלל יוצר אзорים שנקראים:

1. אзор דחיה – דחיה של השערת האפס כלומר קבלה של האלטרנטיבה.
2. אзор קבלה – קבלה של השערת האפס ודחיה של האלטרנטיבה.

כל הכרעה מתבסס על איזשהו סטטיסטי. בתחילת יש ל选取 תוצאות המדגם ולבזוק האם התוצאות נופלות באזרור הדחיה או הקבלה וכן להגיע למסקנה – המסקנה היא עירובן מוגבל כיוון שהיא תלויה בכל הכרעה ובתוצאות המדגם. אם נשנה את כלל הכרעה אז אנחנו יכולים לקבל מסקנה אחרת. אם נבצע מדגם חדש אז אנחנו עלולים לקבל תוצאה אחרת. לכן יתכונו טעויות במסקנות שלנו:

		הכרעה	
מציאות		H_0	H_1
	H_0	טעות מסוג 1	טעות מסוג 2
	H_1	אין טעות	אין טעות

הגדרת הטעויות:

טעות מסוג ראשון: להכריע לדוחות את H_0 למראות שבמציאות H_0 נכונה.
טעות מסוג שני: להכריע לקבל את H_0 למראות שבמציאות H_1 נכונה.

דוגמה (פתרון בהקלטה):

אדם חשוד בביוץ עבירה ונتابע בבית המשפט.
אילו סוגי טעויות אפשריות בהכרעת הדין?

שאלות:

- 1)** לפי הצהרת היכרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הcrcנים מTELONNATE על הפחתת נפח המשקה בבקבוק מהכמות המוצחרת. בדוגמא שעשתה אגודת הcrcנים התקבל נפח ממוצע של 492 סמ"ק במדוגים בגודל 25. בסופו של דבר הוחלט להזכיר לטובת חברת המשקאות.
- רשמו את השערות המחקר.
 - מה מסקנת המחקר?
 - אייזו סוג טעות יתכן וביצעו במחקר?
- 2)** במחקר על פרמטר מסוים הוחלט בסופו של דבר לדוחות את השערת האפס.
- אם ניתן לדעת אם בוצע טעות במחקר?
 - מה סוג הטעות האפשרית?
- 3)** לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטיית תקן 0.4. ישנה טענה שכיוום ממוצע מספר הילדים במשפחה קטן יותר. לצורך כך הוחלט לדוגם 121 משפחות. בדוגמא התקבל ממוצע 2.17 ילדים למשפחה. על סמך תוצאות המדוגם נקבע שלא ניתן לקבוע שבאופנו מובהק תוחלת מספר הילדים למשפחה קטנה כיום.
- מהי אוכלוסיות המחקר?
 - מה המשנה הנחקרה?
 - מה הפרמטר הנחקר?
 - מה השערות המחקר?
 - מה מסקנת המחקר?
 - מי סוג הטעות האפשרית במחקר?

תשובות סופיות:

- 1)** א. $\mu = 500$.
ב. $\mu < 500$.
- 2)** א. לא ניתן לדעת.
ב. טעות מסווג ראשונה.
- 3)** א. משפחות כיום.
ב. מס' הילדים.
- ג. תוחלת מספר הילדים למשפחה כיום.
ה. לא לדוחות את H_0 . ו. טעות מסווג שנייה.
- $H_0 : \mu = 2.3$.
 $H_1 : \mu < 2.3$.

סטטיסטיקה ב

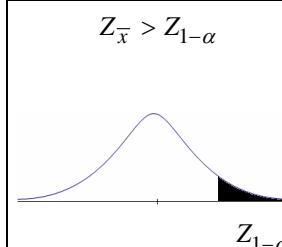
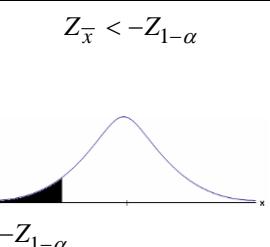
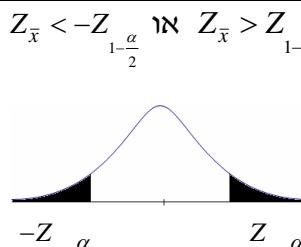
פרק 6 - בדיקת השערות על תוחלת (mmoatzu)

תוכן העניינים

1. בדיקת השערות על תוחלת (mmoatzu) כשינויים האוכלוסייה ידועה.	35
2. מובהקותות תוצאה - אלף מינימלית (שינויים האוכלוסייה ידועה)	39
3. בדיקת השערות על תוחלת (mmoatzu) כשינויים האוכלוסייה לא ידועה.	44
4. מובהקותות תוצאה - אלף מינימלית (שינויים האוכלוסייה לא ידועה)	48
5. הקשר בין רוח סמך לבדיקה השערות על תוחלת (mmoatzu)	51

בדיקות השערות על תוחלת (ממוצע) כשבונות האוכלוסייה ידועה:

רקע:

$H_0 : \mu \leq \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu = \mu_0$	השערת האפס: השערת אלטרנטיבית:
$H_1 : \mu > \mu_0$	$H_1 : \mu < \mu_0$	1. σ ידועה או מוגן מספיק גדול $X \sim N$.2	
$Z_{\bar{x}} > Z_{1-\alpha}$  -דוחים את H_0	$Z_{\bar{x}} < -Z_{1-\alpha}$  -דוחים את H_0	$Z_{\bar{x}} < -Z_{\frac{1-\alpha}{2}}$ או $Z_{\bar{x}} > Z_{\frac{1-\alpha}{2}}$  -דוחים את H_0	כל הכרעה: אזור הדחיה של H_0

סטטיסטי המבחן: $Z_{\bar{x}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

חלופה אחרת לכל הכרעה:

$\bar{X} > \mu_0 + Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} < \mu_0 - Z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ או $\bar{X} < \mu_0 - Z_{\frac{1-\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$	נתקה אם מתקיים: H_0
--	--	--	-------------------------------------

דוגמה:

יבול העגבנייהות מתפלג נורמלית עם תוחלת של 10 טון לדונם וסטיית תקן של 2.5 טון לדונם בעונה. משערים ששיטת זיוב חדש תעלה את תוחלת היבול לעונה מבלי לשנות את סטיית התקן. נדגמו 4 חלוקות שזובלו בשיטה החדש. היבול הממוצע שהתקבל היה 12.5 טון לדונם. בדקו את ההשערה ברמת מובהקות של 1%.

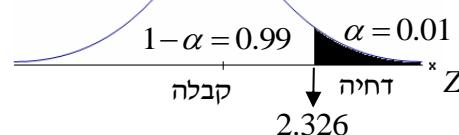
פתרונות:אוכלוסייה: עגבנייהות.המשתנה: X = יבול העגבנייהות בטון לעונה.הפרמטר: μ = תוחלת היבול בשיטה החדש.

$$\begin{aligned} H_0 : \mu &= 10 \\ H_1 : \mu &> 10 \end{aligned}$$

תנאים:

1. $X \sim N$.

2. $\sigma = 2.5$.

כל הכלעה:נדחה את H_0 אם $Z_{\bar{x}} > 2.326$ תוצאות: $n = 4$, $\bar{x} = 12.5$

$$\text{סטטיסטי המבחן} : Z_{\bar{x}} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

$$\text{נzieb} : Z_{\bar{x}} = \frac{12.5 - 10}{\frac{2.5}{\sqrt{4}}} = 2 < 2.326$$

מסקנה:לא נדחה H_0 (נקבל H_0).

ברמת מובהקות של 1% לא נוכל לקבל את הטענה ששיטה החדש עולה את תוחלת היבול של העגבנייהות.

שאלות:

- 1)** ממוצע הציונים בבחינות הבגרות באנגלית הנו 72 עם סטיית תקן 15 נקודות. מורה טוען שפיתח שיטת לימוד חדשה שתעלה את ממוצע הציונים. משרד החינוך החליט לתת למורה 36 תלמידים אקראיים. ממוצע הציונים של אותם תלמידים לאחר לימודו בשיטתו היה 75.5. בהנחה שגם בשיטתו סטיית התקן תהיה 15 מה מסקנתכם ברמת מובהקות של 5%?
- 2)** לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודות היצרנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המומוצרת. במדוגם שעשתה אגודות היצרנים התקבל נפח ממוצע של 492 סמ"ק במדוגם בגודל 25.
- מה מסקנתכם ברמת מובהקות של 2.5%?
 - האם ניתן לדעת מה תהיה המסקנה עבור רמת מובהקות גבוהה מ-5%?
- 3)** מהנדס האיכות מעוניין לבדוק אם מכונה מכילה (מאופסת). המכונה כוננה לחתווך מוטות באורך 50 ס"מ. לפי נתוני היצרן סטיית התקן בחיתוך המוטות היא 0.5 ס"מ. במדוגם של 50 מוטות התקבל ממוצע אורך המוט 50.93 ס"מ. מה מסקנתכם ברמת מובהקות של 5%?
- 4)** המשקל המומוצע של הספורטאים בתחום ספורט מסויים הוא 90 ק"ג, עם סטיית תקן 8 ק"ג. לפי דעת מומחים בתחום יש צורך בהורדת המשקל ובשימוש בדיאטה מסוימת לצריכה להביא להורדת המשקל. לשם בדיקתיעילות הדיאטה נלקח מדגם מקורי של 50 ספורטאים ובתום שנה של שימוש בדיאטה התברר שהמשקל המומוצע במדוגם זה היה 84 ק"ג. יש לבדוק בר"מ של 10%, האם הדיאטה גורמת להורדת המשקל.
- 5)** לפי מפרט נתון, על עובי בורג להיות 4 מ"מ עם סטיית תקן של 0.2 מ"מ. במדוגם של 25 ברגים העובי המומוצע היה 4.07 מ"מ. קבעו ברמת מובהקות 0.05, האם עובי הברגים מתאים למפרט. הניחו כי עובי של בורג מתפלג נורמלית וסטיית התקן של עובי בורג היא אכן 0.2 מ"מ.
- 6)** במחקר נמצא שתוצאה היא מובהקת ברמת מובהקות של 5% מה תמיד נכון? בחרו בתשובה הנכונה.
- הגדלת רמת המובהקות לא תנסה את מסקנת המחקר.
 - הגדלת רמת המובהקות תנסה את מסקנת המחקר.
 - הקטנת רמת המובהקות לא תנסה את מסקנת המחקר.
 - הקטנת רמת המובהקות תנסה את מסקנת המחקר.

- 7) חוקר ערך מבחן דו צדי ברמת מובהקות של α והחליט לדחות את השערת האפס. אם החוקר היה עורך מבחן דו צדי ברמת מובהקות של $\frac{\alpha}{2}$ אזי בהכרח:
- השערת האפס הייתה נדחתה.
 - השערת האפס הייתה לא נדחתה.
 - לא ניתן לדעת מה תהיה מסקנתו במקרה זה.
- 8) שני סטטיסטיקים בדקו השערות: $H_1: \mu > \mu_0$, $H_0: \mu = \mu_0$ נגדן עברו שנות ידועה ובאותה רמת מובהקות. שני החוקרים קיבלו אותו ממוצע במדגם אך לחוקר א' היה מדגם בגודל 100 ולחוקר ב' מדגם בגודל 200.
- אם בחוקר א' החלטת לדחות את H_0 , מה יהיה החלטת בחוקר ב'? נמקו.
 - אם בחוקר א' יחליט לא לדחות את H_0 , מה יהיה החלטת בחוקר ב'? נמקו.

תשובות סופיות:

- נקבל H_0 , בר"מ של 5% לא נקבל את הטענה של המורה ששיטת הלימוד שלו מעלה את ממוצע הציונים.
- א. נדחה H_0 , בר"מ של 2.5% נקבל את תלונת אגודות הרכנים בדבר הפחחת נפח המשקה בבקבוק.
ב. הגדלנו את רמת המובהקות בכך אנחנו נשארים בדוחיה של H_0 והמסקנה לא תשתנה.
- נדחה H_0 , בר"מ של 5% נקבע שהמכונה לא מאופסת.
- נדחה H_0 , בר"מ של 0.1 נקבל את הטענה שהדיאטה עיליה ומפחיתה את המשקל הממוצע.
- נקבל H_0 , בר"מ של 0.05 נכירע שתוחלת עובי הבורג מתיים למפרט.
- א'.
ב'. ג'.
- א. לדחות.
ב. לא ניתן לדעת.

mobekot_tozacha - alfa_minimalit (shevona) האוכלוסייה ידועה:

רקע:

דרך נוספת להגעה להכרעות שלא דרך כלל הכרעה, היא דרך חישוב מובהקות התוצאות :

באמצעות תוצאות המדגם מחשבים את מובהקות התוצאה שמסומן ב- p_v .
את רמת המובהקות החוקר קובע מראש לעומת זאת, את מובהקות התוצאה החוקר יוכל לחשב רק אחרי שייהיו לו את התוצאות.

המסקנה של המחקר תקבע לפי העיקרונו הבא : אם $\alpha \leq p_v$, דוחים את H_0 .
mobekot_tozacha זה הסיכוי לקבלת תוצאות המדגם וקייזוני מהתוצאות אלה בהנחה השערת האפס.

(לקבל את תוצאות המדגם וקייזוני) $\cdot p_v = P_{H_0}$

אם ההשערה היא דו צדדיות :

(לקבל את תוצאות המדגם וקייזוני) $\cdot p_v = 2P_{H_0}$

mobekot_tozacha היא גם האלפא המינימלית לדחיתת השערת האפס.

$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_1: \mu > \mu_0$	$H_1: \mu < \mu_0$	$H_1: \mu \neq \mu_0$	השערת האפס : השערה אלטרנטיבית :		
$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$. σ ידועה					
$P_{H_0}(\bar{X} \geq \bar{x})$	$P_{H_0}(\bar{X} \leq \bar{x})$	$2 \cdot P_{H_0}(\bar{X} \geq \bar{x}) \iff \bar{x} > \mu_0$ אם $2 \cdot P_{H_0}(\bar{X} \leq \bar{x}) \iff \bar{x} < \mu_0$ אם						

כאשר בהנחה השערת האפס :
 $Z_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} , \bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$

דוגמה:

המשקל הממוצע של מתגייסים לצבע לפני 20 שנה היה 65 ק"ג. מחקר מעוניין לבדוק האם כיום המשקל הממוצע של מתגייסים גבוה יותר. נניח שהמשקל המתגייסים מתפלג נורמלית עם סטטיסטיקה של 12 ק"ג. במדגם של 16 מתגייסים התקבל משקל ממוצע של 71 ק"ג.

- מהי מובהקות התוצאה?
- מה המסקנה אם רמת המובהקות היא 5% ואם רמת המובהקות היא ?!

פתרון:

a. אוכלוסייה: המתגייסים לצבע ביום.

משתנה: X = משקל בק"ג.

פרמטר: μ .

השערות:
 $H_0: \mu = 65$
 $H_1: \mu > 65$

תנאים:

. $X \sim N$. 1

. $\sigma = 12$. 2

תוצאות מדגם:

$$n = 16$$

$$\bar{X} = 71$$

$$P_V = P_{H_0} \left(\text{لتוצאות המזגם וקיצוני} \right) = P_{H_0} (\bar{X} \geq 71) = 1 - \phi(2) = 1 - 0.9772 = 0.0228$$

$$Z_{\bar{x}} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{71 - 65}{12 / \sqrt{16}} = 2$$

$$\alpha_{\min} = 0.0228$$

שאלות:

- 1)** להלן השערות של מחקר: $H_0: \mu = 70$, $H_1: \mu > 70$. המשתנה הנחקר מתפלג נורמלית עם סטיטית תקן 20. במדגם מאותה אוכלוסייה התקבלו התוצאות הבאות: $\bar{x} = 74$, $n = 100$. מהי מובהקות התוצאה?
- 2)** השכר הממוצע במשק בשנת 2012 היה 8800 נס' עם סטיטית תקן 2000. במדגם שנעשה אטמול על 100 עובדים התקבל שכר ממוצע 9500 נס'. מטרת המחקר היא לבדוק האם כיים חלה עלייה בשכר. עבור אילו רמות מובהקות שיבחר החוקר יוחלט שחלла עלייה בשכר הממוצע במשק?
- 3)** אדם חושד שהברת ממתקים לא עומדת בהתחביבוותה, ומשקלו של חטייף מסוים אותו הוא קונה מדי בוקר נזוק מ-100 גרם. חברות הממתקים טוענת מצידה שהיא אכן עומדת בהתחביבוותה. ידוע כי סטיטית התקן של משקל החטייף היא 12 גרם. האדם מתכוון לשקלול 100 חפיפות חטייפים ולאחר מכן מכון להגיע להחלטה.
לאחר הבדיקה הוא קיבל משקל הממוצע של 98.5 גרם.
א. רשמו את השערות המחקר.
ב. מהי רמת המובהקות המינימלית עבורה דוחים את השערת האפס?
ג. מהי רמת המובהקות המקסימלית עבורה קיבל את השערת האפס?
ד. מה המסקנה ברמת מובהקות של 5%?
- 4)** מכונה לחישוק מוטות בפעול חותכת מוטות באורך שמתפלג נורמלית עם תוחלת אליה כוונה המכונה וסטיטית תקן 2 ס"מ. ביום מסוים כוונה המכונה לחישוך מוטות באורך 80 ס"מ. אחרי האיכות מעוניין לבדוק האם המכונה מכילה. לצורך כך נדרגו מקו הייזור 16 מוטות שנחתכו אורכו הממוצע היה 81.7 ס"מ.
א. מהי רמת המובהקות המינימלית עבורה נכירע שהמכונה לא מכילה?
ב. אם נסיף עוד ציפוי שערכה יהיה 82 ס"מ, כיצד הדבר ישפיע על התשובה של הסעיף הקודם?
ג. הכרע ברמת מובהקות של 5% האם המכונה מכילה.
- 5)** אם מקבלים בחישובים לפחות מינימלית (value P) קטנה מאד, סביר להניח כי החוקר ידחה את השערת האפס בקלות. נכון/לא נכון? נמק.

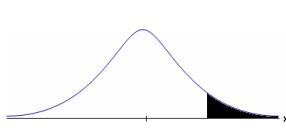
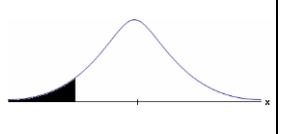
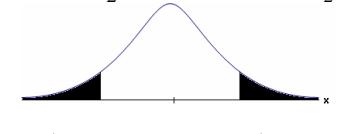
- 6) בבדיקה השערות התקבל שה- $p-value = 0.02$. מה תהיה מסקנת חוקר המשמש ברמת מובהקות 1%? בחרו בתשובה הנכונה.
- יקבל את השערת האפס בכל מקרה.
 - ידחה את השערת האפס מקרה.
 - ידחה את השערת האפס רק אם המבחן הנו דו צדדי.
 - לא ניתן לדעת כי אין מספיק נתונים.
- 7) מובהקות התוצאה (PV) היא גם (בחרו בתשובה הנכונה):
- רמת המובהקות המינימאלית לדחינת השערת האפס.
 - רמת המובהקות המקסימאלית לדחינת השערת האפס.
 - רמת המובהקות שנקבעה מראש על ידי החוקר שטרם קיבל את תוצאות המחקר.
 - רמת המובהקות המינימאלית לאי דחינת השערת האפס.
- 8) בבדיקה השערות מסוימת התקבל: $p value = 0.0254$ לכן (בחרו בתשובה הנכונה):
- ברמת מובהקות של 0.01 אך לא של 0.05 נדחה את H_0 .
 - ברמת מובהקות של 0.01 ושל 0.05 לא נדחה את H_0 .
 - ברמת מובהקות של 0.05 אך לא של 0.01 נדחה את H_0 .
 - ברמת מובהקות של 0.01 ושל 0.05 נדחה את H_0 .

תשובות סופיות:

- (1) 0.0228.
- (2) עבר כל רמת מובהקות סבירה.
- (3) $H_0: \mu = 100$.
 $H_1: \mu < 100$.
- א. 0.1056. ב. 0.1056. ג. נכון.
- ד. נכريع שיש עמידה בהתחייבות של החברה.
- (4) א. 0.0006. ב. יקטן. ג. נכريع שאין כיול.
- (5) נכון.
- (6) א'.
- (7) א'.
- (8) ג'.

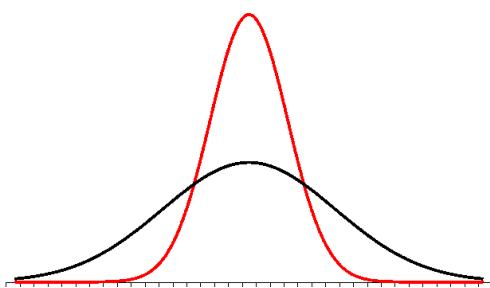
בדיקות השערות על תוחלת (ממוצע) כשבונות האוכלוסייה לא ידועה:

רקע:

$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	השערת האפס: השערה אלטרנטיבית:
$H_1 : \mu > \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu \neq \mu_0$	
.1. σ אינה ידועה או מוגן מספיק גדול $X \sim N$.2			תנאים:
$t_{\bar{x}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha,n-1}$	$t_{\bar{x}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha,n-1}$	$t_{\bar{x}} < -t_{\frac{1-\alpha}{2},n-1}^{(n-1)}$ או $t_{\bar{x}} > t_{\frac{1-\alpha}{2},n-1}^{(n-1)}$  $-t_{\frac{1-\alpha}{2},n-1}$ $t_{\frac{1-\alpha}{2},n-1}$	כל ההכרעה: אזור הדחיה של H_0 :
H_0 - דוחים את ■	H_0 - דוחים את ■	H_0 - דוחים את ■	
$\bar{X} > \mu_0 + t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} < \mu_0 - t_{1-\alpha}^{n-1} \cdot \frac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + t_{\frac{1-\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$ או $\bar{X} < \mu_0 - t_{\frac{1-\alpha}{2}}^{n-1} \cdot \frac{S}{\sqrt{n}}$	חלופה לכלל הכרעה: נדחה H_0 אם מתקימים:

$$\text{סטטיטיסטי המבחן: } t_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$$

**התפלגות T:**

הינה התפלגות סימטרית בעומנית שהתוחלת שלה היא 0. ההתפלגות דומה לתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים. התפלגות T תלויות במושג שנקרא דרגות החופש.

דרגות החופש הן: $df = n - 1$.

כל שדרגות החופש עלות התפלגות הופכת להיות יותר גבוהה וצרה. כסדרות החופש שוואות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה (פתרון בהקלטה):

מפעל קיבל הזמנה לייצור משטחים בעובי של 0.1 ס"מ. כדי לבדוק האם המפעל עומד בדרישה נדגו 10 משטחים ונמצא שהעובי הממוצע הוא 0.104 עם אומדן לסטיתת תקן 0.002 ס"מ.

א. מהו השערות המתקרי?

ב. מה ההנחה הדורשה לצורך פתרון?

ג. בדוק ברמת מובהקות של 5%.

שאלות:

- 1)** משך זמן ההחלמה בלקיחת אנטיביוטיקה מסויימת הוא 120 שעות בממוצע עם סטיית תקן לא ידועה. מעוניינים לבדוק האם אנטיביוטיקה אחרת מקטינה את משך זמן ההחלמה. במדגם של 5 חולים שלקחו את האנטיביוטיקה האחראית התקבלו זמני ההחלמה הבאים: 125, 100, 95, 90, 80 שעות. מה מסקנתכם ברמת מובהקות של 5% מהי ההנחה הדרושה לצורך הפתרון?
- 2)** משרד הבריאות פרסם משקל ממוצע של תינוקות ביום היולדות בישראל 3300 גר'. משרד הבריאות רוצה לחקור את הטענה שנשים מעשנות בזמן ההריון يولדות תינוקות במשקל נמוך מהתוצאתם. במחקר השתתפו 20 נשים מעשנות בהריון. להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:
- $$n = 20$$
- $$\bar{x} = 3120$$
- $$S = 280$$
- מה מסקנתכם ברמת מובהקות של 5% מה יש להניח לצורך פתרון?
- 3)** ציוני מבחן אינטיליגנציה מתפלגים נורמלית. באלה"ב ממוצע הציונים הוא 100. במדגם שנעשה על 23 נבחנים ישראלים, התקבל ממוצע ציונים 104.5 וסטיית התקן המדגמית 16. האם בישראל ממוצע הציונים שונה מאשר באלה"ב? הסיקו ברמת מובהקות של 5%.
- 4)** באוכלוסייה מסוימת נדגמו 10 תכפיות והתקבלו התוצאות הבאות:
- $$\sum_{i=1}^{10} X_i = 750$$
- $$\sum_{i=1}^{10} (X_i - \bar{X})^2 = 900$$
- נתון שההתפלגות היא נורמלית.
בדוק ברמת מובהקות של 5% האם התוחלת של ההתפלגות שונה מ-80.

- 5) ליאור ורוני העלו את אותן השערות על ממוצע האוכלוסייה. כמו כן הם התבפסו על אותן תוצאות של מדגם. ליאור השתמש בטבלה של התפלגות Z. רוני השתמש בטבלה של התפלגות t. מה נוכל לומר בנוגע להחלטת המחקר שלהם? בחר בתשובה הנכונה.
- אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.
 - אם רוני ידחה את השערת האפס אז גם בהכרח ליאור.
 - שני החוקרים בהכרח הגיעו לאותה מסקנה.
 - לא ניתן לדעת על היחס בין דמיון השערת האפס של שני החוקרים.

- 6) נתון ש: $H_0: \mu = \mu_0$ ו- $H_1: \mu < \mu_0$. $X \sim N(\mu, \sigma^2)$ כמו כן נתונות ההשערות הבאות:
- חוקר בדק את ההשערות הללו על סמך מדגם שככל 10 תצפיות. σ^2 לא הייתה ידועה לחוקר. החוקר החליט לדוחות את השערת האפס ברמת מובייקות של 5% לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תצפיות וشكلל את תוצאות אלה גם למדגם כך שככל עכשו 15 תצפיות. בחר בתשובה הנכונה:
- cut בברור הוא ידחה את השערת האפס.
 - cut הוא דוקא קיבל את השערת האפס.
 - cut לא ניתן לדעת מה תהיה מסקנתו.

תשובות סופיות:

- 1) נדחה H_0 .
- 2) נדחה H_0 .
- 3) קיבל H_0 .
- 4) קיבל H_0 .
- 5) ב'.
- 6) ג'.

mobekot_tozacha - alfa_minimalit (shevona) האוכלוסייה לא ידועה):

רקע:

נזכיר שהמסקנה של המבחן תיקבע לפי העיקרון הבא: אם $\alpha \leq p_v$ דוחים את H_0 .
 mobekot_tozacha היא הסיכוי לקבל תוצאות המדגם וקיצוני מהתוצאות אלה בהנחה השערת האפס.
 • $p_v = P_{H_0}$ (לקבל את תוצאות המדגם וקיצוני)
 אם ההשערה היא דו צדדית:
 • $p_v = 2P_{H_0}$ (לקבל את תוצאות המדגם וקיצוני)

mobekot_tozacha היא גם האלפא המינימלית לדחינת השערת האפס.

$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	$H_0: \mu = \mu_0$	השערת האפס: השערה אלטרנטיבית:	
$H_1: \mu > \mu_0$	$H_1: \mu < \mu_0$	1. σ אינה ידועה או 2. מוגן מספיק גדול $X \sim N$			
$P_{H_0}(\bar{X} \geq \bar{x})$	$P_{H_0}(\bar{X} \leq \bar{x})$	$2 \cdot P_{H_0}(\bar{X} \geq \bar{x}) \leftarrow \bar{x} > \mu_0$ $2 \cdot P_{H_0}(\bar{X} \leq \bar{x}) \leftarrow \bar{x} < \mu_0$			
		p-value			

$$t_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$$

$$d.f = n-1$$

דוגמה:

ממוצע זמן הנסיעה של אדם לעובדה הינו 40 דקות. הוא מעוניין לבדוק דרך חלופית שאמורה להיות יותר מהירה. לצורך כך הוא דוגם 5 ימים שבהם הוא נוסע בדרך החלופית. זמני הנסיעה שקיבל בדיקות הם: 34, 40, 30, 32, 27. הנicho שזמן הנסיעה מתפלג נורמלית.

- רשמו את השערות המחקר.
- מצאו חסמים לモבಹקות התוצאה.
- מה המסקנה ברמת מובהקות של 5%?

פתרון:

אוכלוסייה: כלל הנסיעות לעובדה בדרך החלופית.

משתנה: $X =$ זמן נסעה בדיקות.

תנאים: $X \sim N$.

פרמטר: μ .

א. השערות:
 $H_0: \mu = 40$
 $H_1: \mu < 40$

ב. תוצאות המדגם:

$$n = 5, \bar{X} = \frac{\sum X_i}{n} = \frac{34 + 40 + \dots}{5} = 32.6$$

$$S^2 = \frac{\sum X_i^2 - n \cdot \bar{X}^2}{n-1} = \frac{34^2 + 40^2 + \dots - 5 \cdot 32.6^2}{5-1} = 23.4$$

$$S = \sqrt{23.4}$$

$$t_{\bar{X}} = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{32.6 - 40}{\frac{4.88}{\sqrt{5}}} = -3.39$$

$$P_V = P_{H_0} = (\bar{X} \leq 32.6) = P(t \leq -3.39)$$

$$d.f = 5 - 1 = 4$$

$$1\% < P_V < 2.5\%$$

$P_V < \alpha = 0.05$, לכן דוחים את H_0 .

מסקנה: בר"מ של 5% נכريع שהדרך החלופית מהירה יותר.

שאלות:

- 1)** קוו ייצור אריזות סוכר נארזות כך שהמשקל הממוצע של אריזות הסוכר צריך להיות אחד קילוגרם. בכל יום דוגמים מקו הייצור 5 אריזות במטרה לבדוק האם קו הייצור תקין. בבדיקה דגמו 5 אריזות סוכר ולהלן משקלן בגרמים: 1024, 996, 1005, 997, 1008.
 א. רשמו את השערות המחקר.
 ב. מהי מובಹקות התוצאות? הצג חסמים.
 ג. מה המסקנה ברמת מוב hawkות של 5%?
- 2)** חוקר בדק את הטענה כי פועלים העובדים במשמרתليل האיטיים יותר מפועלים העובדים ביום. ידוע כי משך הזמן הממוצע הדרוש לייצר מוצר מסוים ביום הוא 6 שעות. בדוגמא מיקרי של 25 פועלים שעבדו במשמרתليل נמצא כי הזמן הממוצע לייצר אותו מוצר הוא 7 שעות עם סטיית תקון של 3 שעות. מהי α -המינימלית שלפיה ניתן להחליט שכאן העובדים במשמרתليل האיטיים יותר?
- 3)** הגובה של מתגייםים לצה"ל מתפלג נורמלית. בדוגמא של 25 מתגייםים מדדו את הגבהים שלהם בס"מ והתקבלו התוצאות הבאות:

$$\sum (x_i - \bar{x})^2 = 2832, \bar{x} = 176.2.$$

 מטרת המחקר היא לבדוק האם תוחלת הגבהים של המתגייםים גבוהה מ-174 ס"מ באופן מובהק. מהי בקרוב מוב hawkות התוצאות ועל פייה מה תהיה המסקנה ברמת מוב hawkות של 6%?

תשובות סופיות:

- 1)** א. $H_0: \mu = 1000$. ב. $P_v \leq 20\%$.
 $H_1: \mu \neq 1000$.
- ג. ברמת מוב hawkות של 5% לא נוכל לקבוע שקו הייצור אינו תקין.
- 2)** 10%
- 3)** נקבל את $H_0, 1.01$

הקשר בין רוח סמך לבדיקה השערות על תוחלת (מומוצע):

רקע:

ניתן לבצע בדיקת השערות דו צדדיות ברמת מובהקות α על μ :

$$\mu_0 : \mu = \mu_1 , H_0 : \mu \neq \mu_0$$

על ידי בניית רוח סמך ברמת סמך של $\alpha - 1$ ל- μ :

אם μ_0 נופל ברווח \leftarrow קיבל את H_0 .

אם μ_0 לא נופל ברווח \leftarrow נדחה את H_0 .

דוגמה:

חוקר ביצע בדיקת השערות לתוחלת. להלן השערותיו :

$$H_0 : \mu = 80 , H_1 : \mu \neq 80 , \alpha = 5\%$$

החוקר בנה רוח סמך ברמה של 90% וקיבל: $84 < \mu < 79$.

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

פתרון (פתרון מלא בהקלטה):

רוח הסמך ברמת סמך של 90% מכיל "80".

ברמת סמך של 95% רוח הסמך יגדל וכייל "80".

לכן, ברמת מובהקות של 5% קיבל H_0 .

שאלות:

- 1)** חוקר רצה לבדוק את ההשערות הבאות: $H_0: \mu = 90$, $H_1: \mu \neq 90$. החוקר בנה רוח סמך לתוכלת ברמת סמך של 95% וקיבל את רוח הסמך הבא: (87, 97). אם החוקר מעוניין לבצע בדיקת השערות ברמת מובהקות של 1% האם ניתן להגיע למסקנה ע"י רוח הסמך? נמקו.
- 2)** חוקר מעוניין לבדוק השפעת דיאטה חדשה על רמת הסוכר בدم. ידוע כי מספר מיליגרים הסוכר בסמ"ק דם הוא משתנה מקרי שמתפלג נורמלית עם סטיית תקן 10.4 מ"ג. נלקח מדגם של 60 נבדקים שניזונו מדיאטה זו. נמצא כי ממוצע מספר המיליגרים סוכר היה 115.5 מ"ג לסמ"ק.
- א. בנה רוח סמך ברמת סמך 95% לתוכלת רמת הסוכר בדם אצל הניזונים מדיאטה זו.
- ב. ידוע שתוחלת רמת הסוכר בדם באוכלוסייה היא 90 מ"ג לסמ"ק. האם לדעתך ניתן להסיק על סמך תוצאת סעיף א' שהדיאטה משפיעה על רמת הסוכר בדם? הסבירו.
- 3)** יצרן אנטיביוטיקה רושם על גבי התרופות שכמות הפנצליין היא 200 מ"ג لكפסולה. משרד הבריאות ביצע מדגם של 8 קפסולות אקראיות מקו הייצור ומצא שבממוצע יש 196 מ"ג פנצליין لكפסולה עם סטיית תקן מדגמית של 5 מ"ג. בהנחה וכמות הפנצליין בקפסולה מתפלגת נורמלית.
- א. בנו רוח סמך ברמת סמך של 95% למומוצע כמות הפנצליין لكפסולה המיוצרת על ידי יצרן האנטיביוטיקה.
- ב. בדקו ברמת מובהקות של 5% האם יש אמת באינפורמציה המופיעukt על ידי הייצרן.

תשובות סופיות:

- 1)** קיבל השערת.
- 2)** א. $112.87 \leq \mu \leq 118.13$
ב. נזכיר שהדיאטה משפיעה על תוכלת רמת הסוכר בדם.
- 3)** א. $200.2 \leq \mu \leq 191.8$. ב. נזכיר שיש אמת בפרסום.

סטטיסטיקה ב

פרק 7 - רוח סmarketing להפרש תוחלות (ממוצעים) במדגים בלתי תלויים

תוכן העניינים

1. כישרונות האוכלוסייה ידועות	53
2. כישרונות האוכלוסייה לא ידועות ובהנחת שוויןון שונות	55

כשונות האוכלוסייה ידועות:

רקע:

המטרה היא לאמוד את פער התוחלות: $\mu_2 - \mu_1$, כולם ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$.

התנאים לבניית רוח הסמן:

1. σ^2_1, σ^2_2 ידועות.

2. $X_1, X_2 \sim N_{n_1, n_2}$ או $n_1, n_2 > 30$.

3. שני מוגדים בלתי תלויים.

$$\text{רוח סמן: } (\bar{x}_1 - \bar{x}_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}}$$

אם הערך אפס נופל בגבולות רוח הסמן נגיד שבביטחון של $\alpha-1$, לא קיים הבדל בין התוחלות.

דוגמה (פתרון בהקלטה):

נדגמו 100 תושבים מאזור A והמשכורת הממוצעת הייתה ש 9200 ₪. כמו כן נdagmo 120 תושבים מאזור B וממוצע המשכורות שהתקבל ש 8700 ₪. לצורך פתרון נניח שטטיית התקן של המשכורות באוכלוסיות שני האזוריים היא 1800 ₪. אמדנו ברמת סמן של 90% את הפרש השכר הממוצע בין אזור A לאזור B.

שאלות:

- 1)** מעוניינים לבדוק האם קיים הבדל בין ממוצע ציוני הפסיכומטרי של חילילים לממוצע ציוני הפסיכומטרי של תלמידי תיכון. ידוע שצינוי הפסיכומטרי מתפלג עם נורמלilit עס טיטית תקן 100. במדגם של 16 נבחנים חילילים התקבל ממוצע 543. במדגם של 20 תלמידי תיכון התקבל ממוצע 508. בנו רוח סמך לפער תוחלות הצינויים בין חילילים לתלמידי תיכון ברמת סמך של 90%. מה ניתן להסיק מרווח סמך זה?
- 2)** צינוי IQ מתוכנים כך שיתפלגו נורמללית עם סטיטית תקן של 15. במדגם של 20 נבחנים ישראלים התקבל ממוצע צינויים 104. במדגם של 23 נבחנים אמריקאים התקבל ממוצע צינויים 99.
- א. בנו רוח סמך ברמת סמך של 95% לפער בין ישראל לארה"ב בממוצע הצינויים בבחן IQ-ו.
- ב. האם קיים הבדל בין ישראלים אמריקאים מבחינת ממוצע הצינויים?
- 3)** חברה להנדסת בנין מעוניינת להשוות ברמת הקשיות של שני סוגי ברגים. ידוע שרמת הקשיות של ברגים מתפלגת נורמלית עם סטיטית תקן של 4 יחידות. במדגם של 15 ברגים מסווג א' התקבל רמת קשיות ממוצעת של 28 יחידות ובמדגם של 12 ברגים מסווג ב' התקבל רמת קשיות ממוצעת של 25. עברו אילו רמות בטחון יקבע שאיו הבדל בין שני סוגי הברגים מבחינת ממוצע רמת הקשיות שלהם?

תשובות סופיות:

- (1) (-20,90).
- (2) א. $\mu_2 - \mu_1 < 3.99$.
- ב. לא נוכל לטעון בביטחון של 95% שקיימים הבדל בין ישראל לארה"ב.
- (3) רמות בטחון הגבוחות מ-0.9476.

כשונות האוכלוסייה לא ידועות ובהנחה שוויון שונויות:

רקע:

המטרה היא לאמוד את פער התוחולות: $\mu_2 - \mu_1$, כולם ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$.

התנאים לבניית רוח סמך:

$$\cdot \sigma^2_1 = \sigma^2_2 \cdot 1$$

$$\cdot X_1, X_2 \sim N \cdot 2$$

3. מוגדים בلتוי תלויים.

השונות המשוקלلت: כיון שאנו מניחים שבין שתי האוכלוסיות השונות שווות אנו אומדים את השונות הזו על ידי שקולן שתי השונות של שני המוגדים על ידי

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

דרגות החופש: $d.f = n_1 + n_2 - 2$

$$\text{רוח סמך: } (\bar{x}_1 - \bar{x}_2) \pm t_{1-\frac{\alpha}{2}}^{n_1+n_2-2} \cdot \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

אם הערך אפס נופל בגבולות רוח סמך נגד שבביטחון של $\alpha - 1$, לא קיים הבדל בין התוחולות.

דוגמה (פתרון בהקלטה):

מחקר מעוניין לבדוק האם קיים הבדל בין תל אביב לברק שבע מבחן הכנסה הממוצעת של אקדמיים. להלן תוצאות המוגדים שנעשה:

טל אביב	ברק שבע	מספר האקדמאים
20	10	
11,000	9500	ממוצע הכנסות של אקדמאים
200	250	סטיית התקן של הכנסות אקדמיים

בנו רוח סמך ברמת ביטחון של 90% להפרש תוחולות הכנסה בשני האזוריים. הניחו שהשכר מתפלג נורמלית עם אותה שוננות בכל אחד מהאזורים.

שאלות:

- 1) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו ל מבחן IQ.
להלן תוצאות המדגמים:

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

מצאו רוח סמך ברמת סמך של 95% לסתיטה בין ממוצע הציונים בישראל לממוצע הציונים בארה"ב. רשמו את כל הנקודות הדרושים לצורך פתרון התרגילים.

- 2) להלן 4 תוצאות על משתנה X שמתפלג: $N(\mu_x, \sigma^2)$, ומשתנה Y שמתפלג: $N(\mu_y, \sigma^2)$.

X	22	20	21	25
Y	18	25	17	12

חשבו רוח סמך ל- $\mu_x - \mu_y$ ברמת הסמך 90%, בהנחה שני המדגמים בלתי תלויים.

תשובות סופיות:

- 1) הנקודות:
 1. השונות שווה.
 2. שהציונים מתפלגים נורמלית.
 3. המדגמים אינם תלויים זה בזה.
 $-5.52 < \mu_1 - \mu_2 < 17.52$
 $-9.6 < \mu_y - \mu_x < 1.6$ (2)

סטטיסטיקה ב

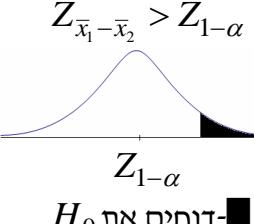
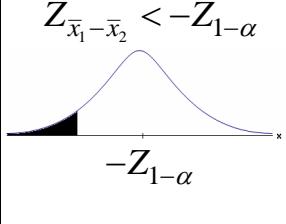
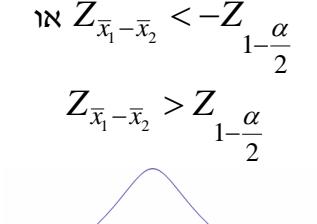
פרק 8 - בדיקת השערות על הפרש תוחלות במדגים בלתי תלויים

תוכן העניינים

57	1. כישרונות האוכלוסייה ידועות
61	2. כישרונות האוכלוסייה לא ידועות ומנחים שהן שוות

בדיקות השערות על הפרש תוחלות בדוגמים בלתי תלויים

כשהשונות של האוכלוסייה ידועות – רקע

$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	השערת האפס: השערת אלטרנטיבית:
$H_1 \quad \mu_1 - \mu_2 > c$	$H_1 \quad \mu_1 - \mu_2 < c$	$H_1 \quad \mu_1 - \mu_2 \neq c$	דוגמים בלתי תלויים σ_1, σ_2 $X_1, X_2 \sim N$ או דוגמים מספיק גודולים
$Z_{\bar{x}_1 - \bar{x}_2} > Z_{1-\alpha}$  $Z_{1-\alpha}$ -דוחים את H_0 ■	$Z_{\bar{x}_1 - \bar{x}_2} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ -דוחים את H_0 ■	$Z_{\bar{x}_1 - \bar{x}_2} < -Z_{\frac{1-\alpha}{2}}$ או $Z_{\bar{x}_1 - \bar{x}_2} > Z_{\frac{1-\alpha}{2}}$  $-Z_{\frac{1-\alpha}{2}} \quad Z_{\frac{1-\alpha}{2}}$ -דוחים את H_0 ■	כלל הכרעה: אזור הדחיה של H_0 : $Z_{\frac{1-\alpha}{2}}$

סטטיסטי המבחן:

$$Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

חלופה אחרת לכלל הכרעה:

נחתה אם מתקיים : H_0	
$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ או
$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	
$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	

התפלגות הפרש הממוצעים : $\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

$$\text{התקנו: } Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

דוגמה (פתרון בהקלטה) :

בשנת 2004 הופיע בין השכר הממוצע של הגברים לנשים היה 3000₪ לטובת הגברים. מעוניינים לבדוק האם כיום הצטמצם הופיע בין הגברים לנשים מבחינת השכר הממוצע. נגמו 100 עובדים גברים. שכרם הממוצע היה 4,072 ₪. נגמו 80 עובדים, שכרכו הממוצע היה 7809 ₪. לצורך פתרון נניח שסטיות התקן של השכר ידועות ושוות ל-2000₪ באוכלוסייה הנשים ו-3000₪ באוכלוסייה הגברים. מה המסקנה ברמת מבוהקות של 5%?

שאלות

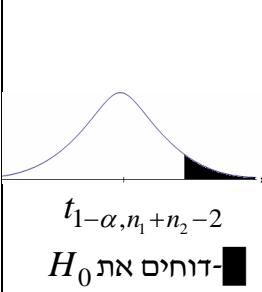
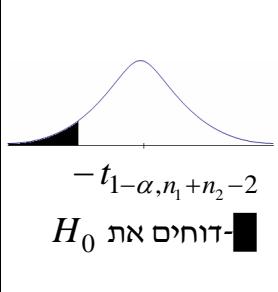
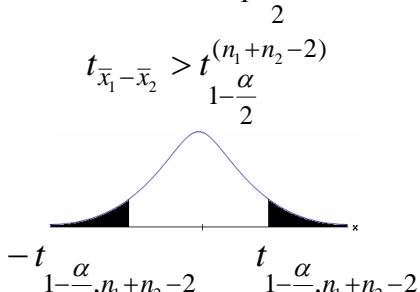
- 1) מחקר טוען שאנשים החיים במרכז הארץ צופים בממוצע בטלוויזיה יותר מ אנשים שלא חיים במרכז. נדגוו 100 אנשים מהמרכז ו-107 אנשים לא מהמרכז. אנשים אלה נשאלו כמה שעות ביום הם נהגים לצפות בטלוויזיה. במדגם של מרכז הארץ התקבל ממוצע 2.7 שעות. במדגם של מחוז למרכזי הארץ התקבל ממוצע 1.8 שעות. לצורך פתרון הניחו שככל אзор, סטיית התקן היא שעה 1 ביום. בדקו את טענת המחקר ברמת מובהקות של 1%.
- 2) ציוני פסיכומטרי מתפלגים נורמלית עם סטיית התקן 100. מכון ללימוד פסיכומטרי טוען שהוא יכול לשפר את ממוצע הציונים ביותר מ-30 נקודות. במדגם של 20 נבחנים שניגשו ל מבחן ללא הינה במכון התקבל ממוצע 508. במדגם של 25 נבחנים שעברו הינה במכון התקבל ממוצע ציונים 561. מה מסקנתכם ברמת מובהקות של 5%.

תשובות סופיות

- 1) נדחה H_0 .
- 2) לא נדחה את H_0 .

בדיקות השערות על הפרש תוחלות במדגמים בלתי תלויים

כששונוויות האוכלוסייה לא ידועות ומניחים שהן שווות – רקע

$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	$H_0 \quad \mu_1 - \mu_2 = c$	השערת האפס: השערת אלטרנטיבית:
$H_1 \quad \mu_1 - \mu_2 > c$	$H_1 \quad \mu_1 - \mu_2 < c$	1. מדגמים בלתי תלויים 2. σ_1, σ_2 לא ידועות אך שווות 3. המשתנים בכל אוכלוסייה מתפלגים נורמלית	
$t_{\bar{x}_1 - \bar{x}_2} > t_{1-\alpha}^{(n_1+n_2-2)}$ 	$t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\alpha}^{(n_1+n_2-2)}$ 	$t_{\bar{x}_1 - \bar{x}_2} < -t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$ או $t_{\bar{x}_1 - \bar{x}_2} > t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}$ 	אזור הדחיה של H_0

סטטיסטי המבחן:

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

השונות המשוקלلت:

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

חלופה אחרת לכל הכרעה:

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

דוגמה (פתרון בהקלטה) :

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

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

טמפרטורת ההתקה הממוצעת במתכת החדשה 1317 מעלות עם אומד חסר הטיה לשונות $S^2 = 260$.
 נניח לצורך פתרון שטמפרטורת ההתקה מתפלגת נורמללית עם אותה שונות במתכוות 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 שבסנת 2013 הייתה ירידה משמעותית בשטחי הדירות לעומת שנת 2012
 עבור רמת מובהקות של 5%.
 הניתנו שטחי הדירות בכל שנה מתפלגים נורמלית עם אותה שוננות.

2) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו ל מבחון IQ. להלן תוצאות

הדגם :	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 ב מבחון ה-IQ לטובת ישראל.
 רשמו את כל ההנחות הדרושים לצורך פתרון התרגיל.

3) להלן תוצאות מדגם הבדיקה אורך חיים של נורות מסוג W60 ומסוג W100.

אורך החיים מממד בשעות.

הקבוצה	60W	100W
\bar{x}	1007	956
S	80	72
n	13	15

- א. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקיות במעט.
 יותר מאשר נורות מסוג W100. רשמו את כל ההנחות הדרושים לפתרון.
- ב. עבור איזו רמת מובהקות ניתן לקבוע שנורות מסוג W60 דולקיות במעט
 יותר מאשר נורות מסוג 100?
- ג. בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקיות יותר
 מאשר נורות מסוג W100. רשמו את כל ההנחות הדרושים.

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 .
- ג. לא נדחה את H_0 .

סטטיסטיקה ב

פרק 9 - רוח סמרק לתוצאות (ממוצע) ההפרשים במדגמים מזוגיים

תוכן העניינים

1. רוח סמרק לתוצאות (ממוצע) ההפרשים במדגמים מזוגיים 65

רוח סמך לתוחלת (ממוצע) ההפרשים במדגים مزוגים:

רקע:

מבחן מזוג: מבחן אחד שבו יש n צמדים. כל תצפית במדגם תנפק זוג ערכים: X ו- Y . ניצור משתנה חדש: $D = y - x$.

הפרמטר שנרצה לאמוד: μ_D .

התנאים לבניית רוח הסמך:

1. $x, y \sim N$.

2. המבחן מזוג.

נוסחת רוח הסמך: $\bar{D} \pm t_{\frac{\alpha}{2}}^{n-1} \frac{S_D}{\sqrt{n}}$.

כאשר דרגות החופש: $df = n - 1$.

דוגמה (פתרון בהקלטה):

מעוניינים לבדוק האם יש הבדל בין מהירות הריצות של שתי תוכנות מחשב. נלקחו 5 קבצים אקראיים והריצו אותם בשתי התוכנות:

5	4	3	2	1	הקובץ
38	46	49	48	25	זמן בתוכנה הראשונה
48	40	42	46	27	זמן בתוכנה השנייה

הניחו כי זמי הריצות מתפלגים נורמלית.
מצאו רוח סמך של 95% להפרש תוחלת הזמן בין שתי התוכנות.

שאלות:

- 1)** נדגמו 5 סטודנטים ששסיימו את הקורס סטטיסטיקה ב'. להלן הציונים בסמסטר א' ו-ב':

82	75	90	68	74	סמסטר א'
100	76	87	84	80	סמסטר ב'

- נניח שהציונים מתפלגים נורמללית.
- א. בנו רוח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סמסטר א' לבין סמסטר ב'.
- ב. האם על סמך רוח הסמך קיים הבדל בין הסמסטרים מבחינת תוחלת הציונים?
- ג. מה צריך לשנות בתנאים כדי שהمدגמים יהיו בלתי תלויים?
- 2)** במטרה לבדוק האם קיים הבדל בין קווי זהב לבזק מבחינת ממוצע המחרירים לשיחות בין"יל. נדגמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקט שיחת. להלן התוצאות:

חברה/ מדינה	ארה"ב	קווי זהב - Y	בזק - X	יפן	סין	מצרים	פולין	הולנד	קנדה
4.2	3.2	3.5	3	2.2	2.1	1.5			
4.2	3.2	3.3	3.1	1.9	2	1.4			

בහנה והחרירים מתפלגים נורמלית עבור כל חברה, בנו רוח סמך ברמת סמך של 90% לתוחלת הפרש המחרירים של שתי החברות.

תשובות סופיות:

- 1)** א. $\mu_0 < 38$. ב. בביטחון של 95% לא קיים הבדל.
ג. ראה הסבר בסרטון.
- 2)** $-0.013 < \mu < 0.185$.

סטטיסטיקה ב

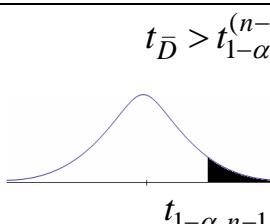
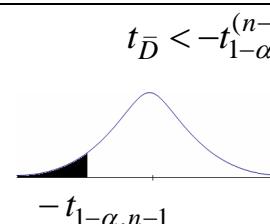
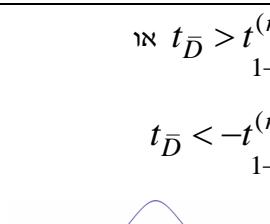
פרק 10 - בדיקת השערות לתוכלת ההפרש במדגמים מזוגים

תוכן העניינים

1. בדיקת השערות למדגמים מזוגים 67

בדיקות השערות על תוחלת הרפישים במדגמים مزוגים (תלויים)

בדיקות השערות למדגמים מזוגים – רקע

$H_0: \mu_D = C$	$H_0: \mu_D = C$	$H_0: \mu_D = C$	השערת האפס: השערת אלטרנטיבית:
$H_1: \mu_D > C$	$H_1: \mu_D < C$	$H_1: \mu_D \neq C$	תנאים:
		1. σ_D אינה ידועה 2. או מדגם מספיק גדול $D \sim N$	
$t_{\bar{D}} > t_{1-\alpha}^{(n-1)}$  $t_{1-\alpha, n-1}$	$t_{\bar{D}} < -t_{1-\alpha}^{(n-1)}$  $-t_{1-\alpha, n-1}$	$t_{\bar{D}} > t_{\frac{1-\alpha}{2}, n-1}^{(n-1)}$ או $t_{\bar{D}} < -t_{\frac{1-\alpha}{2}, n-1}^{(n-1)}$  $-t_{\frac{1-\alpha}{2}, n-1} \quad t_{\frac{1-\alpha}{2}, n-1}$	כלל הבדיקה: אזור הדחיה של H_0
$\bar{D} > C + t_{1-\alpha}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} < C - t_{1-\alpha}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} > C + t_{\frac{1-\alpha}{2}, n-1}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$ ו $\bar{D} < C - t_{\frac{1-\alpha}{2}, n-1}^{n-1} \cdot \frac{S_D}{\sqrt{n}}$	חלהה לכלל הבדיקה: נדחה H_0 אם מתתקיים:

$$S_D^2 = \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n-1} = \frac{\sum_{i=1}^n D_i^2 - n\bar{D}^2}{n-1}, \quad t_{\bar{D}} = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}$$

סטטיסטי המבחן:

דוגמה (פתרו בהקלטה):

חברה שיווקית מעוניינת לבדוק את טענת רשות השיווק "מגה בעיר" הטענה שמחירים נמוכים מהמחירים מרשות השיווק "שופרסל". לצורך הבדיקה נבחרו באקראי 4 מוצרים שונים. המחירים נבדקו בשתי הרשותות. להלן המחירים:

ה מוצר / רשות	מגה בעיר	שופרסל
18	17	שמפו
57	48	gil כביסה
35	35	עוגת גבינה
10	12	לחם
47	49	קפה נמס
142	113	בקבוק יין
26	20	גבינה בולגרית

בහנחה והמחירים מתפלגים נורמלית, בדקנו ברמת מובהקות של 5% את טענת רשות "מגה בעיר".

שאלות

- 1)** במטרה לבדוק האם קיימים הבדל בין חברת X לחברת Y מבחינת המחיר לשיחות בין-יל. נדגמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דלקת 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

בנהנה והמקרים מתפלגים נורמלית בכל חברת, בדקו ברמת מובהקות של 5% האם קיימים הבדל בין החברות מבחינת המחיר במומוץ?

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

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

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

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

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

פורסם שתלמידים שמשיכים את סטטיסטיקה ב' מושפרים ממוצע את הציונים ב-5 נקודות לעומת סטטיסטיקה 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 לא היפנוזה. לאחר מכן נערך לשניהם מבחון באנגלית. נניח שציוני המבחן מתפלגים נורמלית ללא ידיעת השונות האמיתית. 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 מזוגיים.

5) בتحقנת טיפת חלב מסויימת יש שני מכשירי שקילה. על מנת להשוות בין שני המשקלים נדגמו 4 תינוקות. כל תינוק בן חודשיים נשקל בכל אחד מהמשקלים.

להלן תוצאות השקליה (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 למדגים בלתי תלויים.
- ד. מבחן T למדגים מזוגים.

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 למדגים בלתי תלויים.
- ד. מבחן T למדגים מזוגים.

תשובות סופיות

- 1) לא נדחה H_0 .
- 2) לא נדחה H_0 .
- 3) א. לא נדחה H_0 . ב. $0.5 \leq p \leq 0.5$ ג. לא נדחה H_0 .
- 4) ד'.
- 5) ד'.
- 6) ג'.

סטטיסטיקה ב

פרק 11 - הקשר בין רוח סמרק לבדיקה השערות להפרש תוחלות

תוכן העניינים

1. הקשר בין רוח סמרק לבדיקה השערות להפרש תוחלות.....
71

הקשר בין רוח סמך לבדיקה השعروות על הפרש תוחלות

רקע

ניתן לבצע בדיקת השعروות דו צדדיות ברמת מובהקות α על $\mu_1 - \mu_2$:

$$H_0: \mu_1 - \mu_2 = C, \quad H_1: \mu_1 - \mu_2 \neq C$$

על ידי בניית רוח סמך ברמת סמך של $\alpha - 1$ על $\mu_1 - \mu_2$:

אם C נופל ברווח \leftarrow קיבל את H_0 .

אם C לא נופל ברווח \leftarrow נדחה את H_0 .

דוגמה (פתרו בהקלטה) :

חוקר ביצע בדיקת השعروות לתוחלת ההפרש במדגים מזוווג.

להלן השعروתיו : $H_0: \mu_D = 80, \quad H_1: \mu_D \neq 80, \quad \alpha = 5\%$

החוקר בנה רוח סמך ברמה של $90\% < \mu_D < 78$.

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

שאלות

1) נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיKA ב'. להלן ציוניהם בסMASTER A' ו-B':

סMASTER A	סMASTER B
80	74
84	68
87	90
76	75
100	82

א. בנו רוח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סMASTER A' לבין סMASTER B'.

ב. פורסם שתלמידים ממשיכים את סMASTER B' משפרים ממוצע את הציונים ב-5 נק' לעומת סMASTER A'. האם יש אמת בפרסום?

2) הוחלט להשוות הציונים אצל מרצה X ואצל מרצה Y. נבחרו באקראי 6 סטודנטים, 3 סטודנטים של מרצה X ו-3 סטודנטים של מרצה Y, עבורם התקבלו הציונים הבאים:

מרצה X	82	90	68
מרצה Y	68	68	81

א. חשבו רוח סמך ברמת סמך 90% להפרש בין התוחלות של הציונים אצל שני המרצים.

ב. האם ברמת מובהקות של 10% נכיר שיש הבדל בין תוחלות הציונים אצל שני המרצים?

שאלות רב-ברירה:

3) סטטיסטיKA נתקש לאמוד את הפרש הממוצעים של שני טיפולים לפי שני מוגדים מקרים בלתי תלויים.

הוא חישב רוח סמך להפרש ברמת סמך 0.98, וקיים את הרווח $\mu_2 - \mu_1 < 2$. אילו יתבקש החוקר לבדוק לפי אותן נתוניות את השערות:

$H_0: \mu_1 - \mu_2 = 0$; $H_1: \mu_1 - \mu_2 > 0$, מסקנתו תהיה:

א. לדחות את השערת האפס.

ב. לא לדחות את השערת האפס.

ג. שלא ניתן לדעת את המסקנה עבור רמת מובהקות 0.05.

ד. שלא נתנות בשאלת סטיות התקן של האוכלוסיות, ולכן לא ניתן להסיק דבר.

- (4) במטרה לבדוק האם קיימים הבדל בין קווי זהב לבזק מבחינות ממוצע המחרירים לשיחות ביןיל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקט שיחה. בהנחה והמחקרים מתפלים נורמלית בנו רוח סמך ממוצע ההפרשנים וקיים : $\mu_D < 0.0293 - 0.2145$, רוח הסמך הוא ברמת סמך של 95% .
לכן מסקנת המחקר היא :
- א. ברמת מובהקות של 5% לא נוכל לקבוע שקיימים הבדל בין החברות.
 - ב. ברמת מובהקות של 5% נקבע שקיימים הבדל מובהק בין החברות.
 - ג. לא ניתן לדעת מה המסקנה ברמת מובהקות של 5% כיון שלא נאמר מה ההגדרה של D .

תשובות סופיות

(1) א. $-3.8 \leq \mu_D \leq 19$
ב. נכרייע שיש אמת בפרסום.

(2) א. $-8.5 \leq \mu_X - \mu_Y \leq 26.5$
ב. נכרייע שאין הבדל.

(3) ג'.
א. (4)

סטטיסטיקה ב

פרק 12 - רוח סマー לפרופורציה

תוכן העניינים

74	1. רוח הסマー לפרופורציה
77	2. קביעת גודל מוגם

רוח הסמן לפרופורציה:

רקע:

המטרה היא לאמוד את P – פרופורציה באוכלוסייה.

האומד הנקודתי:

$$\hat{p} = \frac{y}{n} \quad Y - \text{מספר ההצלחות שבמדגם}.$$

$$\cdot \hat{p} \pm Z_{\frac{1-\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} : p$$

תנאי לבניית רוח סמן:

מדגם של לפחות 30 תוצאות (לעתים נתונים תנאי של מספר ההצלחות ומספר כשלונות לפחות 5 או לפחות 10).

האומד לטיעות התקן:

$$\cdot L = 2\varepsilon , \hat{P} = \frac{A+B}{2}$$

דוגמה (פתרון בהקלטה):

במטרה לאמוד את אחוז המובטלים במשק נגמו 200 אזרחים, מתוכם התקבלו ש-24 היו מובטלים.

א. בנו רוח סמן לאחוז המובטלים באוכלוסייה ברמת סמן של 95%.

ב. מהו האומד לטיעות התקן?

שאלות:

- 1) נדגמו 200 דירות בעיר חיפה. 48 מהתוכן נמצאו כבעלות ממ"ד.
 - א. בנו רוח סמך ברמת סמך של 95% לאחוז הדירות בחיפה עם ממ"ד.
 - ב. על סמך סעיף א' מה ניתן לומר על שגיאת האמידה המקסימאלית?
 - ג. בהנחה ובחיפה 80 אלף דירות, בנו רוח סמך ברמת סמך של 95% למספר הדירות בחיפה עם ממ"ד בפועל.

- 2) במדגם של 300 אנשי היי-טק התקבל ש-180 מהם אקדמיים.
 - א. בנו רוח סמך לפרוורציה אקדמאים ברמת סמך של 95% (בקרב אנשי היי-טק).
 - ב. כיצד רוח הסמך של סעיף א' תהיה משתנה אם היינו מקטינים את רמת הסמך?
 - ג. כיצד רוח הסמך תהיה משתנה אם היינו מגדילים את גודל המדגמים?

- 3) במדגם של 400 נוהגים התקבל רוח סמך לפרוורציה הנוהגים החדשניים:

$$0.08 < p < 0.18$$
 - א. כמה נוהגים במדגם היו נוהגים חדשים?
 - ב. מהי רמת הסמך של רוח הסמך שנבנה?

- 4) במסגרת מערכת הבחירה בארה"ב נשאלו 840 אנשים עבר איזה מועמד יקבעו. 510 אנשים ענו כי יקבעו בעד ברק אובמה. בסקר פורסם שתתכן סטייה של $\pm 3\%$ מתוצאות האמת. באיזו רמת ביטחון הסקר השתמש?

- 5) במדגם של 300 נשים בגילאי 40-45 נמצא ש-140 היו נשואות, 80 היו גרושות, 60 רווקות והיתר אלמנות.
 - א. מצאו רוח סמך ברמה של 90% לאחוז הגרושות באוכלוסייה הנחקרת.
 - ב. מצאו רוח סמך ברמה של 99% לסיcoli שבאוכלוסייה הנחקרת תמצא אישה לא נשואה?

- 6) ביצעו מדגם באוכלוסייה. שיעור ההצלחות במדגם היה 10% ורוח הסמך ניבנה ברמת סמך של 95%. אורכו הינו 8.3156%. מהו גודל המדגם שנלקח?

תשובות סופיות:

- (1) א. $18.1\% < p < 29.9\%$.
 ב. בביטחון של 95% שגיאת האמידה היא לכל היותר 0.059.
 ג. לא ניתן לדעת.
- (2) א. $0.545 \leq p \leq 0.655$.
 ב. 0.997.
 ג. לא ניתן לדעת.
- (3) א. 0.52.
 ב. 0.925.
 ג. לא ניתן לדעת.
- (4) א. 60.72% > $p > 45.91\%$.
 ב. 30.9% > $p > 22.5\%$.
 ג. לא ניתן לדעת.
- (5) א. 0.200.

קביעת גודל מוגן:

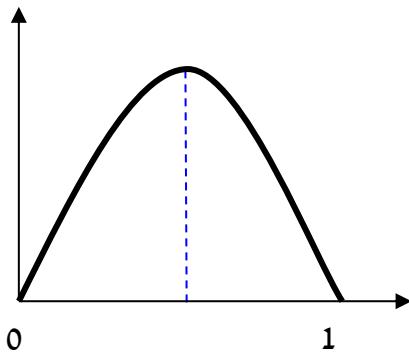
רקע:

בפרק זה נדונן איך קובעים גודל מוגן שבאים לאמוד פרופורציה באוכלוסייה מסוימת: החוקר קובע מראש את רמת הסמך הרצוי: $\alpha - 1$.

החוקר קובע מראש את הטעות הסטטיסטית המרבית שבה הוא מעוניין: ε (או את אורך רוח הסמך).

$L = 2\varepsilon$ - אורך רוח הסמך.

ε - טעות אמידה מרבית: המרחק המקסימלי (הסטטיה) בין הפרמטר (p) לאומד (\hat{p}).



$$\varepsilon = z_{\frac{1-\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

ויתעניין לדעת מהו גודל המוגן הרצוי לשם כך.

$$n \geq \left(\frac{2 \cdot Z_{\frac{1-\alpha}{2}} \cdot \sqrt{\hat{p}(1-\hat{p})}}{L} \right)^2 \quad \text{נקבל ש:}$$

הבעיה שאין אנו יודעים את \hat{p} .
נתבונן בביטויי: $(\hat{p}(1-\hat{p}))$.

כיוון שאין לנו ידע מוקדם על \hat{p} נציב את המקorra השמרני ביותר שמקסם את הביטוי עבור: $\hat{p} = 0.5$.

$$n \geq \left(\frac{2 \cdot z_{\frac{1-\alpha}{2}} \sqrt{0.5 \cdot 0.5}}{L} \right)^2 \Rightarrow n \geq \left(\frac{z_{\frac{1-\alpha}{2}}}{L} \right)^2$$

אך אם תהיה לנו אינפורמציה מוקדמת על הפרופורציה נציב את הערך הקרוב ביותר ל-0.5 האפשרי.

דוגמה (פתרון בהקלטה):

מעוניינים לאמוד את שיעור האבטלה במשק. האמידה צריכה להתבצע ברמת סמך של 90% ועם שגיאת אמידה שלא תעלה על 4%.

- מהו גודל המוגן המינימאלי שיש לקחת?
- חזור לסעיף א' אם ידוע שהابتלה לא אמורה לעלות על 20%.

שאלות:

- 1)** הממשלה אומדת מדי חדש את אחוז הتمיכה בה.
מהו גודל המדגם אשר יש לנקח אם דורשים שהאומדן לא יסטה מהאחוז האמתי באוכלוסייה ביותר מ-3%, וזאת בביטחון של 95%?
- 2)** משרד התקשורת מעוניין לדעת מה שיעור בתי האב עם אינטרנט.
א. כמה בתי אב יש לדגום אם מעוניינים שבביטחון של 90% אורך רוח הסמק לא עולה על 8%?
ב. חזו על סעיף א' אם ידעו שלפניהם חמיש שנים לפחות 80% מבתי האב היה אינטרנט וכיום יש להניח שיש יותר אינטרנט.
- 3)** עורך טלוויזיה מעוניין לאמוד את הריאטיבינג של העורך בפריים טיים. המטרה שבביטחון של 95% הסתירה המרבית בין האומדן לריאטיבינג האמתי לא תעלה על 4%.
א. כמה מכשירי PEOPLE METER יש להתקין לצורך האמידה?
ב. לפי הערכה מוקדמת הריאטיבינג של העורך לא יכול לעלות על 20%.
בנה ומכשיר כזה עולה 500 ש"ח ליחידה מה החיסכון הכספי
מאינפורמציה זאת?
- 4)** ענו על הסעיפים הבאים:
א. כמה אזרחים יש לדגום כדי לאמוד את אחוז הتمיכה הממשלה עם אורך רוח הסמק שלא עולה על 9% ברמת סמק של 90%?
ב. בנה ובצע מדגם שאט גודלו חישבתם בסעיף א' והתקבל שאותו הتمיכה הממשלה במדגם הננו 42%. בנו רוח סמק לאחוז הتمיכה במדינה ברמת סמק של 95%.
ג. על סמך סעיף ב', האם תקבלו את הטענה ששיעור האוכלוסייה תומך הממשלה?
- 5)** משרד הבריאות מתכוון לבצע מדגם שמטרתו לבדוק את הסיכוי לחילות בשפעת עם לקיחת חיסון נגד שפעת. הוא מעוניין שבסיכוי של 98% טעות האמידה לא תעלה על 3%.
א. כמה מחוסנים יש לדגום?
ב. משרד הבריאות ביצע את המדגם שאט גודלו חישבת בסעיף הקודם וקיבול-ש-15% מבין אלה שקיבלו חיסון נגד שפעת בכל זאת חלו במחלה החורף בשפעת. בנו ברמת סמק של 98% את הסיכוי לחילות בחורף בשפעת עם לקיחת חיסון נגד שפעת.
ג. בהמשך לסעיף הקודם. מהי טעות האמידה המרבית בביטחון של 98%?
מדוע הוא קטן מ-3%?

תשובות סופיות:

- .1068 (1)
.271 ב. .423 א. (2)
.108,000 ב. .601 א. (3)
.367 < p < 0.473 ב. .335 א. (4)
ג. בביטחון של 0.95 ניתן להגיד ששיעור האוכלוסייה תומך הממשלה.
(5) א. 1509. ב. 0.15 ± 0.02 ג. ראה סרטון.

סטטיסטיקה ב

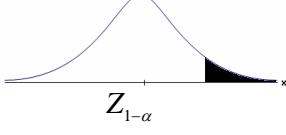
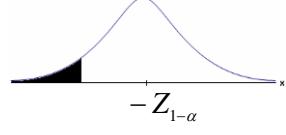
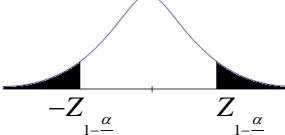
פרק 13 - בדיקת השערות על פרופורציה

תוכן העניינים

80	1. התהילה
83	2. מובהקות התוצאה - אלף מינימלית

התהילה:

רקע:

$H_0 : p = p_0$	$H_0 : p = p_0$	$H_0 : p = p_0$	השערת האפס: השערת אלטרנטיבית:		
$H_1 : p > p_0$	$H_1 : p < p_0$	$H_1 : p \neq p_0$	$np_0 \geq 5 \text{ \& } n(1-p_0) \geq 5$		
$Z_{\hat{p}} > Z_{1-\alpha}$  $Z_{1-\alpha}$ H_0 -דוחים את █	$Z_{\hat{p}} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ H_0 -דוחים את █	$Z_{\hat{p}} < -Z_{1-\frac{\alpha}{2}}$ או $Z_{\hat{p}} > Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}}$ $Z_{1-\frac{\alpha}{2}}$ H_0 -דוחים את █	כלל הבדיקה: אזור הדחיה: של H_0		

סטטיסטי המבחן:

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

חלופה אחרת לכלל הבדיקה:

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

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 אחוז האבטלה הוא כמו בתחילת השנה.

שאלות:

- 1)** במשך שנים אחזו המועמדים שהתקבל לפוקולטה מסוימת היה 25%. השנה מתוך מוגם של 120 מועמדים התקבלו 22. בرمת 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 מטבח 50 פעמים וקיבלו 28 פעמים עז. האם המטבח הוגן ברמת מובהקות של 5%?
- 4)** קפיטריה במכלה מסוימת מעירica כי אחזו הסטודנטים שוכנים קפה בקפיטריה הינו 20%. נערך סקר אשר כלל 200 סטודנטים. התברר כי 33 מהם רוכשים קפה בקפיטריה. מטרת הסקר הייתה לבדוק את אמינותה הערכה של הקפיטריה.
- רשמו את ההשערות.
 - בדקו את ההשערות ברמת מובהקות של 10%.
 - 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.
- מה מסקנתכם ברמת מובהקות של 5%?
 - אם 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 α_1 וחוקר ב' ברמת מובהקות α_2 החוקר הראשון דחה את H_0 ואילו החוקר השני קיבל את H_0 .
- שנייהם התבססו על אותן תוצאות של מוגם. בחר בתשובה הנכונה:
- $\alpha_1 = \alpha_2$.
 - $\alpha_1 > \alpha_2$.
 - $\alpha_1 < \alpha_2$.
 - המצב המתואר לא אפשרי.

תשובות סופיות:

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

МОובקהות התוצאה – אלפא מינימלית:

רקלע:

דרך נוספת להגעה להכרעות שלא דרך כלל הכרעה, היא דרך חישוב מוביקות התוצאה: באמצעות תוצאות המדגם מחשבים את מוביקות התוצאה שמסומן ב- p_v . את רמת המוביקות החוקר קובע מראש לעומת זאת, את מוביקות התוצאה החוקר יכול לחשב רק אחרי שייהו לו את התוצאות. המשקנה של המחקר תקבע לפי העיקרון הבא:

אם $p_v \leq \alpha$ דוחים את H_0 .

mobiyakot haTza'ah zeh haSicovi l'kbelat tzo'ot haMedgem v'kiitzoni matzo'ot ala b'hinchat haShurat haAfsh.

l'kbel at tzo'ot haMedgem v'kiitzoni) $\cdot p_v = P_{H_0}$.

אם ההשערה היא דו צדדית:

לקבול את תוצאות המדגם וקייטוני) $\cdot p_v = 2P_{H_0}$

mobiyakot haTza'ah hia g'm haAlfa haMinimilit laDchiat haShurat haAfsh.

השערת: הAfsh: השערה: אלטרנטיבית:	תנאים:	p-value	
$H_0: p = p_0$ $H_1: p > p_0$	$H_0: p = p_0$ $H_1: p < p_0$	$H_0: p = p_0$ $H_1: p \neq p_0$	$np_0 \geq 5 \& n(1-p_0) \geq 5$
$P_{H_0}(\hat{P} \geq \hat{p})$	$P_{H_0}(\hat{P} \leq \hat{p})$	$2 \cdot P_{H_0}(\hat{P} \geq \hat{p}) \leftarrow \hat{p} > p_0$ $2 \cdot P_{H_0}(\hat{P} \leq \hat{p}) \leftarrow \hat{p} < p_0$	

כאשר בהנחה השערת האפס: $\hat{P} \sim N\left(p_0, \frac{p_0(1-p_0)}{n}\right)$

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

דוגמה (פתרון בהקלטה):

ישנה טענה שיש הבדל בין אחוז הבנים ואחוז הבנות הפונאים ללימוד להנדסאי מחשבים. לשם כך נלקח מבחן מקורי של 200 תלמידים הלומדים מחשבים והתברר כי 112 מהם בניים.

- א. מהי מובاهקות התוצאות?
- ב. מה המסקנה ברמת מובاهקות של 5%?

שאלות:

- 1)** במשך שנים אוחז המועמדים שהתקבל לפקולטה מסוימת היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. רוצים לבדוק האם האם השנה הקשו על תנאי הקבלה.
- מהי מובהקות התוצאה?
 - מה תהיה המסקנה ברמת מובהקות של 1% וברמת מובהקות של 5%?
- 2)** נהוג לחשב ש-60% מהילדים בגיל שלוש קמים מהORITY במהלך הלילה לפחות פעם אחת. ישנה טענה שלאו שנות צהרים פחות מ-60% מהילדים בגיל זה יקומו לפחות פעם אחת במהלך הלילה. נדגו 80 ילדים בגיל 3 אשר אינם ישנים בצהרים מתוכם התקבל ש-41 קמו במהלך הלילה.
- מהי רמת המובהקות המינימלית עבורה תתקבל הטענה במחקר?
 - מהי רמת המובהקות המקסימלית עבורה לא תתקבל טענת המחקר?
 - עבור אילו רמות מובהקות קיבל את טענת המחקר?
 - מה תהיה מסקנת המחקר ברמת מובהקות של 6%?
- 3)** במטרה לבדוק האם מטבח הוגן מטילים אותו 80 פעמים. התקבל ש-60 מההטלות הראו עצ. רשמו את השערות המחקר, חשבו את מובהקות התוצאה והסיקו מסקנה ברמת מובהקות של 5%.
- 4)** בבדיקה השערות על פרופורציה התקבל שה- $p-value = 0.02$.
- מה תהיה מסקנת חוקר המשמש ברמת מובהקות 5%:
(בחרו בתשובה הנכונה)
- קיבל את השערת האפס.
 - דחה את השערת האפס.
 - לא ניתן לדעת כי אין מספיק נתונים.
- 5)** קבעו אם הטענה הבאה נכונה:
"בבדיקה השערות חד-צדדי התקבל ערך $p-value$ של 3%, לכן אם היינו מבצעים מבחן דו-צדדי (כאשר יתר הנסיבות ללא שינוי), היינו מקבלים ערך $p-value$ של 6%".
- 6)** במפעל 10% מהעובדים נפגעים לפחות פעם אחת בשנה מתאונות עבודה. לאור זאת, המפעל החליט לצאת בתוכנית לצמצום שיעור הנפגעים. תוכנית זו נוסתה על 100 עובדים. מתוכם 12 נפגעו בתאונות עבודה במשך השנה.
- מהי רמת המובהקות הקטנה ביותר עבורה יוחלט שהתוכנית יעילה?

תשובות סופיות:

(1) א. 0.0455

ב. ברמת מובהקות של 1% : לא דוחים את H_0 .ברמת מובהקות של 5% : נדחה את H_0 .

(2) א. 0.0548 ב. 0.0548 ג. מעל 0.0548

ד. נכרייע לטובת טענת המחקר.

(3) נדחה את H_0 , $p_v = 0$

(4) ב'.

(5) הטענה נכונה.

(6) 0.7486

סטטיסטיקה ב

פרק 14 - רוח סמך להפרש פרופורציות

תוכן העניינים

1. רוח סמך להפרש פרופורציות 87

רוח סמך להפרש פרופורציותות:

רקע:

המטרה: לאמוד את $p_1 - p_2$: הפרש פרופורציותות בין שתי אוכלוסיות שונות.

האומד הנקודתי: $\hat{p}_1 - \hat{p}_2$.

התנאי לבניית רוח סמך: כל מבחן מעל 30 או לבדוק שמספר ההצלחות ומספר הכשלונות בכל מבחן לפחות 5 בכל מבחן (יש案אלה שבזדים לפחות 10).

$$\text{רוח סמך: } (\hat{p}_1 - \hat{p}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

רק שאפס נופל בתחום רוח סמך להפרש הפרופורציה נאמר שלא ניתן לקבוע שקיים הבדל מובהק בין הפרופורציות באוכלוסיות.

דוגמה (פתרון בהקלטה):

במטרה להשוות בין שתי תרופות נדגמו 200 איש שלקחו תרופה X, מתוכם 180 טענו שהתרופה עוזרת להם. כמו כן, נלקחו 300 איש שלקחו את תרופה Y. מתוכם 150 טענו שהתרופה עוזרת להם. בנו רוח סמך להפרש אחוזי ההצלחה של התרופות ברמת סמך של 95%. מה ניתן לומר על סמך רוח סמך על ההבדלים בין התרופות?

שאלות:

- 1)** מתיוך 150 נשים שנדגמו באקראי 30% תמכו בהצעת חוק מסויימת. מתיוך 200 גברים שנדגמו באקראי 25% תמכו בהצעת החוק.
- בנו רוח סמך לפער בין אחוזי התמיכה של הנשים לעומת הגברים ברמת סמך של 96%.
 - בנו רוח סמך ברמת סמך של 95% לאחוז התמיכה בהצעת החוק.
- 2)** במחקר רפואי השתתפו 200 אנשים הסובלים מכאבים כרוניים. הם חולקו באקראי ל-2 קבוצות שוות בגודלן. קבוצה 1 קיבלה את תרופה A וקבוצה שנייה קיבלה את תרופה B. בקרב לוקחי תרופה A 90 טענו שמצום השתף. בקרב לוקחי תרופה B 70 טענו שמצום השתף.
- בנו רוח סמך ברמת סמך של 95% להפרש בין שיעורי ההצלחה של שתי התרופות.
 - האם על סמך סעיף א' ניתן לקבוע שקיים הבדל בין התרופות מבחינה שיעורי ההצלחה?
- 3)** נדגמו 200 משפחות מגוש דן. ל-70% מהתוכן מכשיר DVD בבית. נדגמו 300 משפחות מאזור הצפון ל-65% מהתוכן מכשיר DVD בבית.
- בנו רוח סמך ברמת סמך של 98% לפרופורציות המשפחות בגוש דן עם DVD בבית.
 - בנו רוח סמך ברמת סמך של 95% להפרש בין פרופורציות המשפחות בגוש דן עם DVD לבין פרופורציות המשפחות בצפון עם DVD.

תשובות סופיות:

$$\text{א. } .22.5\% < p < 31.8\% \quad \text{ב. } . -4.9\% < P_F - P_M < 14.9\% \quad (1)$$

$$\text{ב. כן. } .0.093 < P_A - P_B < 0.307 \quad \text{א. } (2)$$

$$.0.625 < p < 0.7754 \quad (3)$$

סטטיסטיקה ב

פרק 15 - בדיקת השערות על הפרש פרופורציות

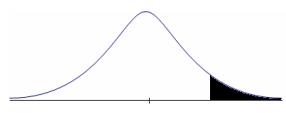
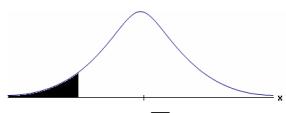
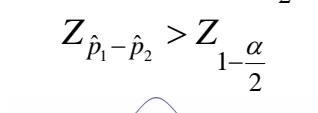
תוכן העניינים

1. כללי

89

בדיקות השערות על הפרש פרופורציות

פרק

השערת האפס השערה אלטרנטיבית:	תנאים:	כלל ההכרעה: אזור הדחיה של
$H_0: p_1 - p_2 = 0$ $H_1: p_1 - p_2 > 0$	$H_0: p_1 - p_2 = 0$ $H_1: p_1 - p_2 < 0$	$H_0: p_1 - p_2 = 0$ $H_1: p_1 - p_2 \neq 0$
2. מדגמים תלויים $Z_{\hat{p}_1 - \hat{p}_2} > Z_{1-\alpha}$  $Z_{1-\alpha}$ H_0 - דוחים את ■	$Z_{\hat{p}_1 - \hat{p}_2} < -Z_{1-\alpha}$  $-Z_{1-\alpha}$ H_0 - דוחים את ■	$Z_{\hat{p}_1 - \hat{p}_2} < -Z_{\frac{1-\alpha}{2}}$ או $Z_{\hat{p}_1 - \hat{p}_2} > Z_{\frac{\alpha}{2}}$  $-Z_{\frac{1-\alpha}{2}}$ $Z_{\frac{1-\alpha}{2}}$ H_0 - דוחים את ■

$$\text{סטטיטיסטי המבחן: } Z_{\hat{p}_1 - \hat{p}_2} \underset{H_0}{=} \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

$$\hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} \quad \text{כאשר הפרופורציה המשוקלلت:}$$

חלופה אחרת לכל הכרעה:

כל הכרעה: איזור הדמייה של H_0	
$\hat{p}_1 - \hat{p}_2 < 0 - Z_{1-\alpha} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$	$\hat{p}_1 - \hat{p}_2 > 0 + Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$ $\hat{p}_1 - \hat{p}_2 < 0 - Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$ או
$\hat{p}_1 - \hat{p}_2 > 0 + Z_{1-\alpha} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$	

$$\hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2}) \quad : \hat{p}_1 - \hat{p}_2$$

$$Z_{\hat{p}_1 - \hat{p}_2} = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}$$

תקנון:

$$Z_{\hat{p}_1 - \hat{p}_2 | H_0} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

דוגמה (פתרון בהקלטה) :

נדגמו 80 סטודנטים שנבחנו במיקרו-כלכלה. מתוכם 60 עברו את הבדיקה.
 נדגמו 100 סטודנטים שנבחנו בסטטיסטיקה א'. מתוכם 82 עברו את הבדיקה.
 האם שיעור העוברים את הבדיקה בסטטיסטיקה גבוהה מאשר מהבדיקה במיקרו
 כלכלה? בדקו ברמת מבוקחות של 10%.

שאלות

- 1)** במדגם של 200 גברים, 8% היו מובטלים. במדגם של 180 נשים, 10% מהן היו מובטלות. האם קיים הבדל מובהק בין פروفורציה המובטלים לפרופורציה המובטלות? בדקו ברמת מובהקות של 5%.
- 2)** אחוז בעלי רישיון נהיגה בקרב האוכלוסייה הבוגרת הינו 60%. במדגם של 300 בוגרים מתל אביב 204 היו בעלי רישיון נהיגה. במדגם של 220 בוגרים מירושלים 100 היו בעלי רישיון נהיגה.
א. ברמת מובהקות של 5% האם תקבלו את הטענה שאחוז בעלי הרישיון בתל אביב גבוה מהאחוז הארץ?
ב. ברמת מובהקות של 10% האם תקבלו את הטענה שאחוז בעלי הרישיון נהיגה בתל אביב גבוה מהאחוז בעלי רישיון הנהיגה בירושלים?
- 3)** נדגו 500 בוגרים מתוכם 200 גברים והיתר נשים. במדגם התקבל: מתוד הגברים ל-48% تعוזת בגרות. מתוך הנשים ל-58% تعוזת בגרות. מטרת המחקר היא לבדוק האם שיעור הזכאות לבגרות גבוהה משיעור הזכאים.
א. מהי מובהקות התוצאה?
ב. מה תהיה המסקנה ברמת מובהקות של 8%?
- 4)** במדגם שנערך על 100 פרות מחוות בדروم הארץ התקבל כי 20 פרות נשאות וירוס מסוים. במדגם שנערך על 200 פרות מחוות בצפון הארץ התקבל כי 10 מתוכן נשאות וירוס גם כן.
א. בנו מבחן ברמת מובהקות של 5% לבדיקת הטענה כי הוירוס תקף את פרות הדром באופן משמעוני יותר מאשר את הפרות בצפון הארץ.
ב. מהי המסקנה לבדיקת הטענה של סעיף א' ומהי הטועה האפשרית במסקנה?
ג. מהי עוצמתה המבחן אם שיעור הפרות בדروم עם הוירוס גבוהה ב-10% משיעור הפרות בצפון עם הוירוס?
ד. כיצד העוצמה תשתנה אם נגדיל את רמת המובהקות?

תשובות סופיות

- (1) לא נדחה את H_0 .
 ב. נדחה H_0 .
 (2) א. נדחה H_0 .
 ב. נדחה H_0 .
 (3) א. 0.0139
 ב. נדחה H_0 .
 (4) א. ראה סרטון.
 ד. תנצל.

סטטיסטיקה ב

פרק 16 - שאלות מסכמת על רוחי סマー

תוכן העניינים

1. שאלות מסכמת על רוחי סマー

93

שאלות מסכמת על רוחי סמך:

שאלות:

2) 200 אנשים נשאלו כמה פעמים ביום הם שותים כוס קפה.
להלן הפלגות התשובות:

מספר אנשים	מספר פעמים
86	0
34	1
28	2
22	3
20	4
10	5

א. תנו רוח סמך למספר כוסות הקפה שאנשים נהגים לשותות ביום. $\alpha = 0.05$.

ב. אדם השותה לפחות 4 כוסות קפה ביום נקרא "מכור לקפה". בנו רוח סמך לאחוז "המכורים לקפה". $\alpha = 0.1$.

3) חוקר בנה רוח סמך לאחוז האנשים שהתקררו לפחות פעם אחת בשנה. רוח הסמך שהתקבל הוא: $p < 91 < 81$. רוח הסמך הניל התבסס על מדגם של 500 איש.

א. כמה אנשים במדגם טענו שככל לא התקררו השנה?

ב. באיזו רמת סמך בונה רוח הסמך?

ג. בנו רוח סמך לאחוז האנשים שהתקררו לפחות פעם אחת השנה ברמת סמך של 96% על סמך תוצאות המדגם.

4) ציוני IQ באלה"ב מתפלגים נורמלית עם תוחלת 100. במדגם של 20 ישראלים שנבחנו ב מבחן ה-IQ התקבלו התוצאות הבאות:

$$\sum_{i=1}^{20} x_i = 2040, \quad \sum_{i=1}^{20} x_i^2 = 210740$$

א. אמדן ברמת ביטחון של 90% את ממוצע ציוני בחינת ה-IQ בישראל – מהי ההנחה הדרישה לפתרון?

ב. על סמך רוח הסמך של סעיף א' האם תקבלו את הטענה שבישראל ממוצע הציוניים שונה מארה"ב?

ג. מה היה קורה לרוח הסמך אם היינו מגדילים את רמת הסמך שלו?

5) להלן תוצאות מדגם שבדק עבור כל משפחה האם יש לה בבית מכשיר טאבלט:

אזור מגורים	גוש דן	שאר הארץ
גודל המדגם	200	240
מספר משפחות בעלי טאבלט	160	168

- א. בנו רוח סמך להבדל בין אחוז המשפחות עם טאבלט בגוש דן ואחוז המשפחות בעלי טאבלט בשאר חלקי הארץ. ברמת סמך של 98%.
- ב. בנו רוח סמך לפרווריות משפחות בעלות טאבלט בכלל הארץ ברמת סמך של 95%.

7) בנק מתלבט האם לפתח סניף באזור A או באזור B. לצורך פתרון נניח שסטיית התקן של המשכורת באזור A היא 1200 ובאזור B 1500. הבנק דגם 50 אנשים מאזור A, המשכורת המומוצעת שהתקבלה במדגם היא 6,800 ₪. כמו כן, נדגמו 40 אנשים מאזור B, המשכורת המומוצעת שהתקבלה במדגם היא 6,600 ₪.

- א. בנו רוח סמך ברמת סמך של 95% להפרש הממוצעים של המשכורות בשני האזוריים. האם על סמך רוח הסמך ניתן להמליך לבנק היכן לפתח את הסניף. אם כן, היכן?
- ב. בנו רוח סמך לתוחלת המשכורת באזור A ברמת סמך של 95%.

8) להלן מדגם של שכר הדירה ב-₪ של 5 דירות שלושה חדרים בשכונת בבלი בתל אביב:

שנת 2012	7500	6500	7000	7500	8000
שנת 2013	7700	6800	7800	8200	8000

בנו רוח סמך ברמת סמך של 95% לתוחלת עליית שכר הדירה משנה 2012 לשנת 2013 בשכונת בבלוי. ניתן להניח ששכר הדירה בשכונה מתפלג נורמלית.

תשובות סופיות:

- | | | | |
|---------------------------------|-----------------|--------------------------------------|---------------|
| . $10.85\% \leq p \leq 19.15\%$ | ב. | . $1.21 \leq \mu \leq 1.65$ | א. (2) |
| . $83\% < p < 89\%$ | . 0.9988 | . $.70$ | א. (3) |
| ג. גדל. | ב. לא. | . $97.4 \leq \mu \leq 106.6$ | א. (4) |
| . $0.704 \leq p \leq 0.786$ | ב. | . $0.5\% \leq p_1 - p_2 \leq 19.5\%$ | א. (5) |
| . $6467 \leq \mu_A \leq 7133$ | , לא. ב. | . $-372 \leq \mu_A - \mu_B \leq 772$ | א. (7) |
| | | . $-21 \leq \mu_D \leq 821$ | (8) |

סטייסטיקה ב

פרק 17 - שאלות מסכמת בבדיקה השערות

תוכן העניינים

1. שאלות רב ברירה (אמריקאיות)

96

שאלות סיוכם – שאלות רב ברירה על בדיקת השערות

1) בבדיקה השערה חד-צדדית ימנית ברמת מובהקות $\alpha = 0.01$, נדחתה השערת האפס. מה הייתה המסקנה לו נבדקה אותה ההשערה באמצעות אותם תנומים ברמת מובהקות $\alpha = 0.05$?

- א. השערת האפס הייתה נדחתה.
- ב. השערת האפס לא הייתה נדחתה.
- ג. ההשערה המחקנית הייתה נדחתה.
- ד. בהעדר תנומים נוספים, לא ניתן לדעת.

2) על מנת לבדוק האם ההסתברות לילדת בן הינה חצי, נבחר מוגם מקרי של 200 ילדים, ונמצא שישנם 120 בניים.
מהו ההשערה האלטרנטיבית להשערת האפס?

- א. $H_1: p = 0.5$
- ב. $H_1: p = 0.6$
- ג. $H_1: p > 0.5$
- ד. $H_1: p \neq 0.5$

3) לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני לא היפנוזה. לאחר מכן נערך לשניהם מבחן באנגלית. נניח שצינוי המבחן מתפלגים נורמלית ללא ידיעת השונות האמיתית. המבחן שיש לבצע כאן הוא :

- א. מבחן Z למוגם יחיד.
- ב. מבחן Z למדוגמים יחיד.
- ג. מבחן T למוגדים בלתי תלויים.
- ד. מבחן T למוגדים מזוגניים.

4) כדי לבדוק את הטענה שגברים רווקים שוקלים פחות מגברים נשואים ללח חוקר מוגם מקרי של 4 גברים ומדד את משקלם לפני נישואיהם ולאחר נישואיהם. הנה התוצאות :
מהו ההשערות הנבדקות? (ההפרש חושב $Y - X$)

68	82	93	69	X לפני הנישואין -
71	84	88	80	Y לאחר הנישואין -

- א. $H_1: \mu_d < 0, H_0: \mu_d = 0$
- ב. $H_1: \mu_X - \mu_Y < 0, H_0: \mu_X - \mu_Y = 0$
- ג. $H_1: \mu_X - \mu_Y < 0, H_0: \mu_X - \mu_Y = 0$
- ד. $H_1: \mu_d > 0, H_0: \mu_d = 0$

5) חוקר ביצע מחקר ובו עשה טעות מסווג שני לכך :

- השערת האפס נcona.
- השערת האפס נדחתה.
- השערת האפס לא נדחתה.
- אף אחת מהתובשות לא נcona בהכרח.

6) ידוע כי ילד בגיל שנתיים ישן בממוצע 9 שעות בלילה. במדגם של 20 תינוקות

בני שנתיים המתגוררים בצפון נמצא, כי ממוצע שעות השינה בלילה הינו 10 עם סטיית תקן של 1.1. במדגם של 10 תינוקות בדרום נמצא, כי ממוצע שעות השינה בלילה הינו 7.9 עם סטיית תקן של 1.1. על מנת להשוות בין ממוצע שעות השינה של ילדים אלה לבין ממוצע המתגוררים בצפון יש לערוך _____.

יש להניח שההנחה הדרושים מתקיימות.

- מבחן Z למדגם יחיד ; מבחן T למדגם יחיד.
- מבחן T למדגם יחיד ; מבחן T למדגמים תלויים.
- מבחן T למדגם יחיד ; מבחן T למדגמיםבלתי תלויים.
- מבחן T למדגמיםבלתי תלויים ; מבחן T ממוצע יחיד.

7) מובהקות התוצאה (PV) היא גם :

- רמת המובהקות המינימאלית לדחות השערת האפס.
- רמת המובהקות המקסימאלית לדוחית השערת האפס.
- רמת המובהקות שנקבעה מראש על ידי החוקר טרם קיבל את תוצאות המחקר.
- רמת המובהקות המינימאלית לאי דוחית השערת האפס.

כדי לבדוק את הטענה שגברים רווקים שוקלים פחות מגברים נשואים לעומת חוקר מדגם מקורי של 4 גברים ומדד את משקלם לפני נישואיהם ולאחר נישואיהם. הנה התוצאות:

	לפני הנישואין	לאחר הנישואין	
68	82	93	69
71	84	88	80

בדיקות

באיזה התפלגות משתמשים

ההשערות, ובכמה דרגות חופש :

- התפלגות Z ללא דרגות חופש.
- התפלגות T ו-3 דרגות חופש.
- התפלגות T ו-6 דרגות חופש.
- התפלגות χ^2 ו-3 דרגות חופש.

9) שני סטטיסטיקים בודקים השערות ברמת מובהקות $\alpha = 0.05$ על סמך אותו מבחן. סטטיסטיקי א' בודק את ההשערה: $H_0: \mu = 20$ נגד האלטרנטיבית $H_1: \mu \neq 20$ ומחליט לא לדוח את השערת האפס.

סטטיסטייקי ב' בודק את ההשערה $20 \leq \mu: H_0$ נגד

האלטרנטיבית $20 > \mu: H_1$ מה יחליט סטטיסטייקי ב'?

א. לדוח את השערת האפס.

ב. לא לדוח את השערת האפס.

ג. ללא נתונים נוספים אי אפשר לדעת מה יחליט.

10) חוקר בדק השערה מסוימת והחליט לדוח את השערת האפס ברמת מובהקות 5%. מה נכון לו?

א. הוא בודאות ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.

ב. הוא בודאות לא ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.

ג. הוא בודאות ידחה את השערת האפס ברמת מובהקות 9% וברמת מובהקות 2%.

ד. הוא בודאות לא ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.

11) רמת הcolesterol בדם של אנשים מתפלג נורמלית עם תוחלת של 180 מ"ג (ל 100 סמ"ק דם). וטיפות התקן של 10 מ"ג. מעוניינים לבדוק את הטענה שצמחיים הם בעלי רמתコレsterol נמוכה יותר.

נניח שטיפות התקן אצל צמחוניים זהה לטריפות התקן של כלל האנשים.

במבחן של 20 צמחוניים התקבל ממוצע רמתコレsterol 174.5 מ"ג.

אם הוחלט לקבל את הטענה שצמחיים הם בעלי רמתコレsterol נמוכה יותר
איזה סוג טעות אפשרית במסקנה?

א. טעות מסוג ראשון.

ב. טעות מסוג שני.

ג. טעות מסוג שלישי.

ד. לא ניתן לדעת כיון שאין לנו ליריעים מה התוחלת האמיתית אצל הצמחוניים.

12) בסקר שנערך התקבל ש 60% מתוך 220 נשאלים מבקרים אצל השיננית לפחות פעם אחת בשנה. עבור אילו רמות מובהקות ניתן יהיה לקבוע שרוב האוכלוסייה מבקרת אצל השיננית לפחות פעם אחת בשנה?

- א. רמת מובהקות הגדולה מ-5%.
- ב. רמת מובהקות הקטנה מ-5%.
- ג. רמת מובהקות הגדלה מ-0.0015.
- ד. רמת מובהקות הקטנה מ-0.0015.

13) שני חוקרים העוסקים בתחום מחקרי משותף החליטו להסתמך על נתונים של מדגם שפורסם על ידי הלשכה המרכזית לסטטיסטיקה.

חוקר א' ניסח השערה זו צדדית ואילו חוקר ב' ניסח השערה חד צדדית. מסקנתו של איזה מבין המשפטים הבאים הוא הנכון בנוגע למסקנות החוקרים?

- א. אם חוקר א' ידחה את השערת האפס לא ניתן לדעת מה יחליט חוקר ב' באוטה רמת מובהקות.
- ב. אם חוקר א' קיבל את השערת האפס גם חוקר ב' קיבל את השערת האפס באותה רמת מובהקות.
- ג. אם חוקר ב' ידחה את השערת האפס גם חוקר א' ידחה את השערת האפס באותה רמת מובהקות.
- ד. אם חוקר א' ידחה את השערת האפס גם חוקר ב' ידחה את השערת האפס בתנאי שרמת המובהקות כפולה בגודלה.

14) ידוע מנתוני העבר כי תוחלת הציונים בבחינה בפסיכולוגיה היא 79. הועלתה השערה כי תוחלת הציונים בקרב העולים החדשניים נמוכה יותר. לצורך בדיקת הטענה נלקח מדגם מקרי של 47 סטודנטים עולים ונמצא ממוצע של 75. מה משמעות הפרמטר בניסוח ההשערות?

- א. תוחלת ציוני העולים באוכלוסייה.
- ב. ממוצע ציוני העולים במדגם.
- ג. תוחלת ציוני האוכלוסייה מנתוני העבר.
- ד. ממוצע ציוני שאר האוכלוסייה במדגם.

15) חוקר ביצע מחקר וידוע כי עשה טעות מסווג 1. מה מהබאים נכון?

- א. החוקר דחה את השערת H_0 כאשר היא הייתה נכון.
- ב. החוקר דחה את השערת H_1 כאשר היא הייתה נכון.
- ג. החוקר לא דחה את השערת H_0 כאשר היא הייתה לא נכון.
- ד. המדגם של החוקר שייך בפועל להתפלגות הדגימה של H_1 .

16) חוקר ביקש לבחון האם תאומים זהים אשר הופרדו בילדותם שונים מתאומים זהים אשר גדלו יחדיו מבחינות מידת הפער בין התאומים בלחץ הדם. הוא דגם 20 זוגות תאומים מכל אוכלוסייה ומדד את הפרש בין לחץ הדם בכל זוג תאומים. מהו המבחן הסטטיסטי המתאים?

- מבחן D למדגים בלתי תלויים עם 38 דרגות חופש.
- מבחן T למדגים מזוגים, עם 39 דרגות חופש.
- מבחן D למדגים בלתי תלויים עם 39 דרגות חופש.
- מבחן T למדגים מזוגים עם 38 דרגות חופש.

17) בינוואר השנה פורסם שהשכר הממוצע במשק הוא 900,98₪. במדגם שנעשה בחודש יוני על 60 עובדים רשום עבר כל עובד במדגם האם השכר שלו נמוך או לא נמוך מהשכר הממוצע שפורסם בחודש ינואר. מהו המבחן המתאים כדי לבדוק שרוב העובדים בחודש יוני קיבלו שכר הנמוך מהשכר הממוצע שפורסם בחודש ינואר?

- מבחן Z על פרופורציה.
- מבחן T על תוחלת אחת.
- מבחן T על שתי תוחלות במדגים בלתי תלויים.
- מבחן T על שתי תוחלות במדגים תלויים.

18) שלושה חוקרים רצו לבדוק את השפעתו של שידור פרסומות נגד תאונות דרכיים על מהירות הנהיגה של נהגים בישראל (השינויים של מהירות הנהיגה בישראל אינה ידועה). עידו השווה את מהירות הנהיגה של קבוצת נהגים אחת, חודש לפני שידור הפרסומות וחודש לאחר שידור הפרסומות. רון השווה את מהירות הנהיגה של קבוצת נהגים, שראו את הפרסומות, ל מהירות הנהיגה של קבוצת נהגים, שלא ראו את הפרסומות. יואב השווה את מהירות הנהיגה הממוצעת בישראל על פי נתוני משרד התחבורה. הפרסומות, ומהירות הנהיגה הממוצעת בישראל להשתמש הם :

- שלושתם במבחן T למדגים בלתי תלויים.
- עידו במבחן T למדגים מזוגים, רון ויואב במבחן T למדגים בלתי תלויים.
- עידו במבחן T למדגים מזוגים, רון במבחן T למדגים בלתי תלויים ויואב במבחן T למדגם יחיד.
- עידו במבחן T למדגים מזוגים, רון ויואב במבחן T למדגם יחיד.

19) במחקר נמצא שתוצאה היא מובהקת ברמת מובהקות של 5%. מה תמיד נכון?

- א. הגדלת רמת המובהקות לא תשנה את מסקנת המחקר.
- ב. הגדלת רמת המובהקות תשנה את מסקנת המחקר.
- ג. הקטנת רמת המובהקות לא תשנה את מסקנת המחקר.
- ד. הקטנת רמת המובהקות תשנה את מסקנת המחקר.

20) חוקר ערך מבחן דו צדי ברמת מובהקות של α והחליט לדוח את השערת האפס. אם החוקר היה עורך מבחן חד צדי ברמת מובהקות של $\frac{\alpha}{2}$ איזי בהכרח:

- א. השערת האפס הייתה נדחתה.
- ב. השערת האפס הייתה לא נדחתה.
- ג. לא ניתן לדעת מה תהיה מסקנתו במקרה זה.

21) ליאור ורוני העלו את אותן השערות על ממוצע האוכלוסייה. כמו כן הם התבססו על אותן תוצאות של מדגמים.

ליאור השתמש בטבלה של התפלגות Z.

רוני השתמשה בטבלה של התפלגות T.

מה יוכל לומר בנוגע להחלטת המחקר שלהם?

- א. אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.
- ב. אם רוני תדחה את השערת האפס אז גם בהכרח ליאור.
- ג. שני החוקרים בהכרח הגיעו לאותה מסקנה.
- ד. לא ניתן לדעת על היחס בין דחינת השערת האפס של שני החוקרים.

22) נתון ש $(\sigma^2, \mu) \sim N$ כמו כן נתונים ההשערות הבאות: $H_0: \mu = \mu_0$, $H_1: \mu < \mu_0$.

חוקר בדק את ההשערות הללו על סמך מדגם שככל 10 תוצאות. σ^2 לא הייתה ידועה לחוקר. החוקר החליט לדוח את השערת האפס ברמת מובהקות של 5% לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תוצאות וشكל את תוצאות אלה גם למדגם כך שככל עכשו 15 תוצאות.

- א. כתע בברור הוא ידחה את השערת האפס.
- ב. כתע הוא דוקoa קיבל את השערת האפס.
- ג. כתע לא ניתן לדעת מה תהיה מסקנתו.

23) אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו איזי:

- א. הסיכוי לטעות מסוג ראשוני גדול.
- ב. העוצמה של המבחן גבוהה.
- ג. הסיכוי לטעות מסוג שני��大.
- ד. תשובות או-וב נכונות.

24) חוקר ביצע מחקר ובו עשה טעות מסווג שני לכך :

- השערת האפס נכונה.
- השערת האפס נדחתה.
- השערת האפס לא נדחתה.
- אף אחת מהתובשות לא נכון בהכרח.

25) מה המצב הרצוי לחוקר המבצע בבדיקה השערה :

- | | |
|----------|-------------|
| α | $1 - \beta$ |
| א. גדולה | גדולה |
| ב. גדולה | קטנה |
| ג. קטנה | גדולה |
| ד. קטנה | קטנה |

26) נערך שינוי בכלל החלטה של בדיקת השערה מסוימת ובקבוקתו אוזור דחיה H_0 קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה מכך :

- הן α , והן $(\beta - 1)$, יקטנו.
- α יישאר ללא שינוי ואילו $(\beta - 1)$ יגדל.
- α יגדל ואילו $(\beta - 1)$ יקטנו.
- הן α והן $(\beta - 1)$ יגדלו.

27) ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניה שלחץ הדם בקרוב עיתונאים גבוה יותר מה ממוצע באוכלוסייה. הואלקח מדגם של 60 עיתונאים וקיבל ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק שלחץ הדם בקרוב העיתונאים אינו גבוה יותר. מה הטעות האפשרית שהרופא עושה?

- טעות מסווג ראשון.
- טעות מסווג שני.
- טעות מסווג שלישי.
- אין טעות במסקنته.

28) בבדיקה השערות התקבל שה- $p-value = 0.02$. מה תהיה מסקנת חוקר המשמש ברמת מובהקות 1%? בחר בתשובה הנכונה :

- יקבל את השערת האפס בכל מקרה.
- ידחה את השערת האפס מקרה.
- ידחה את השערת האפס רק אם המבחן הינו דו צדדי.
- לא ניתן לדעת כי אין מספיק נתונים.

29) מובಹקות התוצאות (PV) היא גם :

- א. רמת המובಹקות המינימאלית לדוחות השערת האפס.
- ב. רמת המובಹקות המקסימאלית לדוחות השערת האפס.
- ג. רמת המובಹקות שנקבעת מראש על ידי החוקר טרם קיבל את תוצאות המחקר.
- ד. רמת המובಹקות המינימאלית לאי דוחות השערת האפס.

30) בבדיקה השערות מסוימת התקבל $p value = 0.0254$, לכן :

- א. ברמת מובಹקות של 0.01 אך לא של 0.05 נדחה את H_0 .
- ב. ברמת מובಹקות של 0.01 ושל 0.05 לא נדחה את H_0 .
- ג. ברמת מובಹקות של 0.05 אך לא של 0.01 נדחה את H_0 .
- ד. ברמת מובಹקות של 0.01 ושל 0.05 נדחה את H_0 .

31) רמת המובಹקות במחקר הייתה 2% לכן.

- א. בסיכוי של 2% נדחה את השערת האפס.
- ב. בסיכוי של 2% לא נדחה את השערת האפס.
- ג. בסיכוי של 2% השערת האפס לא נכונה.
- ד. אף תשובה לא נכונה.

32) נתון ש: $(\mu, \sigma^2) \sim N$. כמו כן נתונות ההשערות הבאות: $H_0: \mu = \mu_0$, $H_1: \mu < \mu_0$.

חוקר בדק את ההשערות הללו על סמך מדגם שכלל 10 תצפיות.

σ^2 לא הייתה ידועה לחוקר. החוקר החליט לדוחות את השערת האפס ברמת מובಹקות של 5%. אם הוא היה מגדיל את רמת המובಹקות ל-10% אזי:

- א. כעת בוורור הוא ידחה את השערת האפס.
- ב. כעת הוא דוחוק לקבל את השערת האפס.
- ג. כעת לא ניתן לדעת מה תהיה מסקנתו.

33) לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני ללא היפנוזה.

לאחר מכן נערכן לשניהם מבחן באנגלית. ננית שצינוי המבחן מתפלגים נורמללית ללא ידיעת השונות האטומית. מספר דרגות החופש במבחן הוא:

- א. 9
- ב. 19
- ג. 18
- ד. 8

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

משקל במיכס'ר 1	4.5	9.6	0.7	2.5
משקל במיכס'ר 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 מזוגים.

(35) כדי להשוות בין שני אצנים נדגו 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 מזוגים.

(36) סטטיסטיκאי ערך מבחן סטטיסטי. הוא חישב את עצמת המבחן וקיבל 0. המשמעות של תוצאה זו היא :

- א. לעולם לא לדוחות את השערת האפס כאשר היא לא נכונה.
- ב. תמיד לדוחות את השערת האפס כאשר היא נכונה.
- ג. לעולם לא לדוחות את השערת האפס כאשר היא נכונה.
- ד. תמיד לדוחות את השערת האפס כאשר היא לא נכונה.

(37) סטטיסטיκאי נתקבש לאמוד את הפרש הממוצעים של שני טיפולים לפי שני מדגמים מקרים בלתי תלויים. הוא חישב רוחסן סמך להפרש ברמת סמך 0.98 וקיבל את הרוחסן $\mu_2 - \mu_1 < 4.5$. אילו יתבקש החוקר לבדוק לפי אותן

נתונים את השערות : $H_1: \mu_1 - \mu_2 \neq 0$; $H_0: \mu_1 - \mu_2 = 0$,

ברמת מובהקות 0.05 מסקנתנו תהיה :

- א. לדוחות את השערת האפס.
- ב. לא לדוחות את השערת האפס.
- ג. שלא ניתן לדעת את המסקנה עבור רמת מובהקות 0.05.
- ד. שלא נתנות בשאלת סטיות התקן של האוכלוסיות, ולכן לא ניתן להסיק דבר.

38) במטרה לבדוק האם קיים הבדל בין קוווי זהב לבזק מבחינת ממוצע המחרירים לשיחות בינייל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקט שיחה. בהנחה והמחרירים מתפלים נורמלית בנו רוח סמך לממוצע ההפרשים וקיבלו: $\bar{x} = 0.2145$, $s_d = 0.0293$. רוח הסמך הוא ברמת סמך של 95%. לכן מסקנת המחקר היא:

- ברמת מובהקות של 5% לא נוכל לקבוע שקיים הבדל בין החברות.
- ברמת מובהקות של 5% נקבע שקיים הבדל מובהק בין החברות.
- לא ניתן לדעת מה המסקנה ברמת מובהקות של 5% כיון שלא נאמר מה ההגדרה של D .

39) אם רמת מובהקות של מבחן סטטיסטי הינה 0, הכוונה היא:

- תמיד נדחה H_0 כאשר היא נכונה, אך לא תמיד נדחה אותה כאשר היא לא נכונה.
- לא נדחה את H_0 אף פעם.
- לא נדחה את H_0 כאשר היא נכונה אך יתכן ונדחה אותה כאשר היא לא נכונה.
- כל התשובות לא נכונות.

40) חוקר ביצע ניסוי. הוא ניסח את ההשערות הבאות: $H_0: \mu = 10$, $H_1: \mu \neq 10$. לצורך בדיקה הואלקח מוגרבי בגודל 5 מתוך אוכלוסייה המתפלגת נורמלית עם שונות לא ידועה. על סמך תוצאות המוגרם הוא חישב וקיבל: $t_{\bar{x}} = -2.63$. לכן המסקנה היא:

- הוא ידחה H_0 ברמת מובהקות 0.1 אך לא כן ברמת מובהקות 0.05.
- הוא ידחה H_0 ברמת מובהקות 0.05 אך לא כן ברמת מובהקות 0.025.
- הוא ידחה H_0 ברמת מובהקות 0.025 אך לא כן ברמת מובהקות 0.01.
- הוא לא ידחה H_0 ברמת מובהקות 0.1.

41) האיגוד האמריקני לרפואת ילדים מפרסם הנחיות חדשות הקובעות כי יש ליטול תוספת יוד במהלך תקופת ההריון וההנקה. מחסור במינרל זה עלול לגרום לפגיעה מוחית אצל העובר והתינוק. החלטה זו נקבעה על סמך מחקר בו השתתפו 1050 נשים שנטלו יוד במהלך תקופת ההריון וההנקה.

מתוך הנשים שהשתתפו במחקר, רק ל-21 נמצאו ילדים בעלי פגיעה מוחית לעומת 3% באוכלוסייה הכללית. בנוספ', פורסם שהאיגוד האמריקאי מגיע למסקנותיו על סמך רמת מובהקות של 0.5%. מה הסיכוי לבצע טעות מסווג ראשון במחקר?

- 0.005
- 0.03
- 0.0287
- 0.05

(42) חוקרת שיערה, כי משקלן של נשים כשנה לאחר החתונה גבוהה ממשקלן בעת החתונה. החוקרת דגמה 15 נשים, ובדקה את משקלן בשתי נקודות הזמן (בעת החתונה, ו שנה לאחריה), אך לא מצאה הבדל מובהק ברמת מובהקות 0.01. בהנחה, כי במציאות השערתה של החוקרת נכונה, סביר כי אם היא תגדיל את גודל המדגם, אז:

- יקטן הסיכוי לטעות מסוג שני (β).
- תגדל רמת הביטחון ($\alpha - 1$).
- אף תשובה לא נכונה.
- כל התשובות נכונות.

(43) איזה מה המשפטים הבאים נכון תמיד?

- $\text{POWER} + \alpha + \beta = 1$
- $\text{POWER} = 0.5 - \beta$
- $\text{POWER} + \alpha = 1$
- $\beta + \alpha = 1$
- הقول לא נכון.

(44) מה נכון לגבי הנחת שיוויון השווניות ב מבחן T למדגים בלתי תלויים?

- היא אומרת שהשווניות המדגימות שוות.
- בלעדיה אין שום דרך לבדוק השערת על הפרש בין תוחלות.
- היא חשובה הן עבור מדגים מזוגים והן עבור מדגים בלתי תלויים.
- אף תשובה אינה נכונה.

(45) חוקר החליט לא לדוחות השערת ברמת מובהקות של α . במידה וחוקר זה היה בודק השערת זו ברמת מובהקות של $\alpha/2$ על סמך אותם נתונים, האם ההשערה תדחה?

- ההשערה תדחה.
- ההשערה לא תדחה.
- התשובה תליה בעוצמת המבחן.
- לא ניתן לדעת בוודאות אם ההשערה תדחה או לא.

(46) חוקרת שיערה, כי בגילאי הגן בנות יותר תקשורתית מבנים. אם החוקרת תדגום אקרואית 30 בנים ו-30 בנות, ובמדגם יתקבלו ממוצע של ציון תקשורת. סטטיסטי המבחן יהיה:

- אפס
- חיובי
- שלילי
- לא ניתן לדעת

(47) עצמה שווה ל-1 פרושה :

- לעולם לא לדחות את השערת האפס כאשר היא נכונה.
- תמיד לדחות את השערת האפס כאשר היא נכונה.
- לעולם לא לדחות את השערת האפס כאשר היא לא נכונה.

(48) מה מהבאים נכון לגבי מבחן T מדגמים מזוגים?

- כל התצפיות במחקר אינן תלויות זו בזו.
- כל התצפיות במחקר תלויות זו בזו.
- כל הצמדים של תצפיות במחקר אינם תלויים זה זה.
- התצפיות בתוך כל צמד אינן תלויות זו בזו.

(49) לבדיקת ההשערה חד צדדית על התוחלת של התפלגות נורמלית $\mu \geq 10$, $H_0: \mu < 10$. נלקח מבחן והתקבלה רמת מובהקות מינימאלית לדחיה השערת האפס 0.058 . לו רצינו לבדוק את ההשערה הדו צדדית $\mu = 10$, $H_1: \mu \neq 10$, אז על סמך תוצאות אותו המבחן ברמת מובהקות 0.05 :

- ניתן להכיריע בין ההשערות רק אם שונות האוכלוסייה נתונה.
- מקבלים את השערת האפס.
- דוחים את השערת האפס.
- לא ניתן להכיריע בין ההשערות שכן חסרים נתונים.

(50) לבדיקת ההשערה חד צדדית ימנית $\mu = 55$, $H_0: \mu = 55$, $H_1: \mu > 55$. נלקח מבחן מקרי בגודל n מאוכלוסייה בעלת התפלגות נורמלית ושונות σ^2 . רמת המובהקות היא 5% . נמצא שהעוצמה היא 0.9 . להלן 3 טענות :

- עבור מבחן בגודל n ורמות מובהקות 5% לבדיקת ההשערות :
- עבור מבחן בגודל $n=1$ $H_0: \mu = 55$, $H_1: \mu = 60$ העוצמה תהיה גדולה מ-0.9.
- עבור מבחן בגודל $n=2$ ורמות מובהקות 5% לבדיקת ההשערות :
- עבור מבחן בגודל $n=2$ $H_0: \mu = 55$, $H_1: \mu = 65$ העוצמה תהיה גדולה מ-0.9.
- עבור מבחן בגודל $n=3$ ורמות מובהקות 10% לבדיקת ההשערות :
- עבור מבחן בגודל $n=3$ $H_0: \mu = 55$, $H_1: \mu = 65$ העוצמה תהיה קטנה מ-0.9.

- שלושת הטענות אינן נכונות.
- טענות 2 ו-3 אינן נכונות וטענה 1 נכונה.
- טענת 1 ו-2 נכונות וטענה 3 אינה נכונה.
- טענות 1 ו-3 אינן נכונות וטענה 2 נכונה.

תשובות סופיות:

שאלה	תשובה	שאלה	תשובה	שאלה
א	26	א	א	1
ב	27	ד	ד	2
א	28	ד	ג	3
א	29	א	א	4
ג	30	ג	ג	5
ד	31	ג	ג	6
א	32	א	א	7
א	33	ב	ב	8
ד	34	ג	ג	9
ג	35	א	א	10
א	36	א	א	11
ג	37	ג	ג	12
א	38	א	א	13
ג	39	א	א	14
א	40	א	א	15
א	41	א	א	16
א	42	א	א	17
ה	43	ג	ג	18
ד	44	א	א	19
ד	45	ג	ג	20
א	46	ב	ב	21
ד	47	ג	ג	22
ג	48	ד	ד	23
ב	49	ג	ג	24
ד	50	ג	ג	25

סטטיסטיקה ב

פרק 18 - מבחני חי בربוע

תוכן העניינים

1. מבחון טיב התאמה	109
2. מבחון טיב התאמה והקשר שלו לבדיקת השערות על פרופורציה אחת	114
3. מבחון לאי תלות.....	116
4. קשר בין מבחון אי תלות לבדיקת השערות להפרש פרופורציות.	121

מבחני חי בריבוע

מבחן טיב התאמה – רקע

מבחון זה בא לבדוק האם אוכלוסייה מסוימת מתפלגת לפי התפלגות נתונה. המשתנה הנחקר מחולק למספר קטגוריות ויש לבדוק האם תוצאות המדגם תואמות לתפלגות הנתונה.

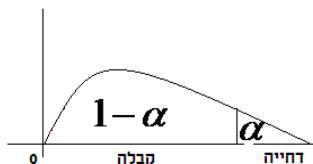
מבנה המבחן:

השערות:

- . המשתנה מתפלג לפי התפלגות מסוימת - H_0
- . אחרת - H_1

כלל הכרעה:

הערך הקרייטי נקבע על סמך התפלגות חי בריבוע. התפלגות זו היא אסימטרית חיובית ותלויה בדרגות החופש $d.f = K - 1$, כאשר K - מספר הקטגוריות.



הערך הקרייטי הוא: $\chi^2_{1-\alpha, K-1}$, כלומר האחוזון $1 - \alpha$ בתפלגות חי בריבוע שדרגות החופש הן $K - 1$.
אם $\chi^2 > \chi^2_{1-\alpha, K-1}$, דוחים את השערת האפס.

$$\text{סטטיסטי המבחן: } \chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

O_i - שכיחות שנצפתה במדגם בקטgorיה i .

p_i - הסתברות לקטgorיה i לפי השערת האפס.

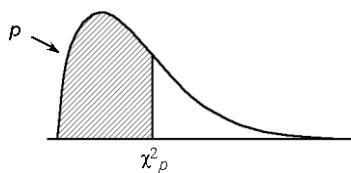
$E_i = np_i$ - שכיחות צפואה במדגם לקטgorיה i בהנחה השערת האפס.

הערה:

תנאי כדי לבצע את המבחן הוא $E_i \geq 5$ לכל i . במידה ותנאי זה לא מתקיים יש אפשרות לאחד קטגוריות סמכות עד שהתנאי יתקיים.

דוגמה (פתרון בהקלטה) :

במדינה מסוימת שלוש מפלגות. בפרלמנט הנוichi התפלגות מספר המושבים היא 30% למפלגה A, 60% למפלגה B ו-10% למפלגה C.
לקראת הבחירות המתוכנות בשבוע הבא נעשה סקר شامل 300 אזרחים.
בסקר התקבל ש-40% יצביעו למפלגה A, 50% למפלגה B ו-10% למפלגה C.
אם תוצאות הסקר תואמות להפלגות המושבים בפרלמנט הנוichi?
בדקו ברמת מובהקות של 5%.

טבלת התפלגות חי-בריבוע – ערכי החלוקה


df	<i>p</i>												
	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	
1	0.0 ⁴ 393	0.0 ³ 157	0.0 ³ 982	0.0 ² 393	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

שאלות

- 1)** במטרה לבדוק האם קובייה הוגנת, מטילים אותה 120 פעמים. התקבל 17 פעמים 1, 23 פעמים 2, 20 פעמים 3, 25 פעמים 4, 18 פעמים 5 ו-17 פעמים 6. מה מסקנתכם ברמת מובהקות של 5%?
- 2)** מפעל מייצר סוכריות בצבעים כחול, אדום, יroke וכתום. מעוניינים לבדוק שפרופורציות הסוכריות הכהולות גדולות פי 2 מכל צבע אחר. לצורך כך נדגמו באקראי 200 סוכריות והתקבל: 70 כחולות, 50 אדומות, 40 יrokeות והיתר כתומות. מה מסקנתכם ברמת מובהקות של 5%?
- 3)** משרד החינוך צוען שבקרב השכירים במשק היחס בין השכירים בעלי השכלה נמוכה, תיכונית ואקדמית הוא 1:2:1 בהתאם. במדגם של 200 שכירים התקבלו 56 אנשים בעלי השכלה נמוכה, 105 בעלי השכלה תיכונית והיתר בעלי השכלה גבוהה.
 א. על סמך תוצאות המדגם, האם התפלגות ההשכלה היא כמו משרד החינוך מפרש? בדוק ברמת מובהקות של 5%.
 ב. בנו רוח סמך ברמת סמך של 95% לפרופורצית השכירים במשק בעלי השכלה אקדמית.
- 4)** 200 איש נקבעו לבחור ספרה באקראי והנה התוצאות שהתקבלו:
 18 איש בחרו בספרה 0, 24 איש בחרו בספרה 1, 17 איש בחרו בספרה 2, 19 איש בחרו בספרה 3, 20 איש בחרו בספרה 4, 18 איש בחרו בספרה 5, 22 איש בחרו בספרה 6 והיתר בחרו בספרות 7-9.
 א. על סמך התוצאות הללו האם בחרית הספרות אקראית?
 בדקו ברמת מובהקות של 2.5%.
 ב. תננו הערקה למובהקות התוצאה.
 ג. אם נגדיל את גודל המדגם פי 2 ונשמר על אותם יחסים של כמות האנשים במדגם שבחרו בספרות, כיצד הדבר ישפיע על ערכו של הסטטיסטי χ^2 ? מה תהיה המסקנה במקרה זה?
- 5)** מעוניינים לבדוק האם קובייה היא הוגנת. הטילו את הקובייה פעמיים והתבוננו בסכום הוצאות. חזרו על התהליך 72 פעמים.
 להלן התוצאות שהתקבלו במדגם:
 מה המסקנה ברמת מובהקות של 5%?

סכום הוצאות	מספר הטלות
20	2-5
17	6-8
20	9-10
15	11-12

6) בפנס יש 4 סוללות. בבדיקה שנערכה ב-400 פנסים נמצאו סוללות פגומות לפי

	0	1	2	3 ומעלה	מספר הסוללות הפגומות
שכיחות	276	104	12	8	

השכיחיות הבאות:

מעוניינים לבדוק על סמך תוצאות מבחן אלה האם הסיכוי לסוללה פגומה הוא 20%. מהי רמת המובהקות המינימלית עבורה נזכיר שהסיכוי לסללה פגומה אינו 20%?

7) להלן השערות מחקר: $H_0 : X \sim N(40, 2^2)$, $H_1 : \text{else}$: תוצאות המבחן הן:

	40-44	36-40	36	מתחת 36	X	מספר הדגימות
2A	45A	50A	3A			

מהו ערכו המקסימלי של A עבורו קיבל את H_0 ברמת מובהקות של 5%?

תשובות סופיות

1) לא נדחה H_0 .

2) לא נדחה H_0 .

3) א. לא נדחה H_0 . ב. (.0.14,0.25)

4) א. לא נדחה H_0 . ב. בין 0.95 ל-0.975

ג. יגדל פי 2; המסקנה לא תשתנה.

5) נזכיר שהקוביות אינה הוגנת.

6) 0.005

7) .14

הקשר בין מבחן טיב התאמה לבדיקת השערות על הפרופורציה – רקע

אם אנו רוצחים לבצע מבחן טיב התאמה על משתנה שיש לו שתי קטגוריות בלבד (משתנה דיקוטומי), הדבר זהה לתהליך של בדיקת השערות דו צדדיות על פרופורציה בודדת.

דוגמה (פתרון בהקלטה) :

הטילו מטבע 80 פעמים וקיבלו 48 פעמים את התוצאה "ראש".
בדקו האם המטבע הוא הוגן ברמת מובהקות של 5%.

- א. באמצעות מבחן טיב התאמה.
- ב. באמצעות מבחן Z לפרופורציה בודדת.

שאלות

- 1)** בסקר שנעשה על 320 נשאלים, 43.75% טענו שהחיה המועדף עליהם היא כלב. עד היום היה נהוג לחשב ש40% מהאנשים מעדיפים כלבים. בדקו ברמת מובהקות של 5% האם הסקר יישנה את הסברת שהייתה נהוגה עד היום לגבי העדפת כלב.
- באמצעות מבחן טיב התאמה.
 - באמצעות מבחן על פרופורציה.
- 2)** לSOCNOOT מכוניות שלושה סניפים ברחבי הארץ. המכוניות נמכרות בסניפים השונים. מתוך 100 מכוניות נמצא ש-65 נמכרו בסניף תל-אביב, 23 בסניף ירושלים והיתר בסניף חיפה.
- בדקו ברמת מובהקות של 5% האם שיעור המכוניות שנמכרות בסניף ת"א גדול פי 2 מכל סניף אחר.
 - בדקו באמצעות מבחן טיב התאמה האם 60% מהמכוניות נהוגות להימכר בסניף תל אביב. האם יש דרך לבדוק את ההשערה?
- 3)** בתחרות ריצה בית ספרית שלושה מסלולי ריצה. ב-50 תחרויות בדקו באיזה מסלול היה הניצחון. התוצאות שהתקבלו מסוימות בטבלה הבאה:
- | מסלול | 1 | 2 | 3 |
|----------------|----|----|----|
| מספר הניצחונות | 20 | 15 | 15 |
- בדקו ברמת מובהקות של 5% האם יש מסלול מועדף לניצחון.
 - בדקו ברמת מובהקות של 5% האם הסיכוי לנצח במסלול מס' 1 גבוה מ- $\frac{1}{3}$.

תשובות סופיות

- לא נזחה H_0 .
- א. נזחה H_0 .
- א. לא נזחה H_0 .

מבחן חי בריבוע לאי תלות בין משתנים – רקע

מבחן לאי תלות מטרתו לבדוק האם קיים קשר בין שני משתנים. שני המשתנים שנבדקים צריכים להיות מחולקים למספר קטגוריות.

מבנה המבחן:

השערות:

אין תלות בין המשתנים H_0 .

יש תלות בין המשתנים H_1 .

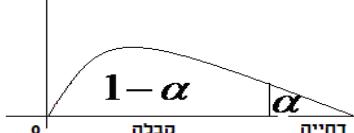
כלל הכרעה:

הערך הקритי נקבע על סמך התפלגות חי בריבוע. התפלגות זו היא אסימטרית חיובית ותלויה בדרגות החופש $(r-1)(c-1)$.
 כאשר: r - מספר הקטגוריות של המשתנה שبشורות.
 c - מספר הקטגוריות של המשתנה שבעמודות.

הערך הקритי הוא: $\chi^2_{1-\alpha,(r-1)(c-1)}$, קלומר האחזוון ה- α – בהתפלגות חי בריבוע שדרגות החופש הן $(r-1)(c-1)$ אם $\chi^2 > \chi^2_{1-\alpha,(r-1)(c-1)}$ או דוחים את השערת האפס.

$$\text{סטטיסטי המבחן: } \chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

כאשר:



O_i - השכיחות נצפית במדגם בתא i .

E_i - שכיחות צפואה במדגם בתא i בהנחה השערת האפס.

$$E_i = \frac{f(x) \cdot f(y)}{n}$$

הערה:

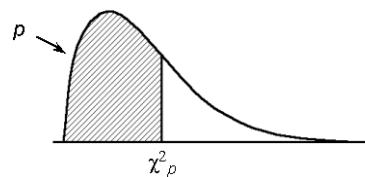
תנאי כדי לבצע את המבחן הוא $5 \geq E_i$ לכל i . במידה ותנאי זה לא מתקיים יש אפשרות לאחד קטגוריות סמכות עד שהתנאי יתקיים.

תנאי חלופי: אין E קטן מ-1 וגם אין יותר מ 20% מהתאים E קטנים מ-5.

דוגמה (הפתרון בהקלטה) :

האם יש תלות בין המגדר לבין דעה מסוימת?
יש לבדוק ברמת מובהקות של 5% על סמך תוצאות הסקר :

המגדר / דעתה	סה"כ	נמנע	נגד	بعد	גברים
נשים	20	60	20	50	10
סה"כ					

טבלת התפלגות חי-בריבוע – ערכי החלוקה χ^2_p


df	p												
	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995
1	0.04393	0.03157	0.03982	0.0393	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

שאלות

1) נבדקה ה תלות בין גודל הארגון לבין שביעות הרצון של העובדים.

להלן התוצאות:

סה"כ	גבואה	ביןונית	נמוכה	שביעות רצון	גודל המפעול
600	215	203	182		גודל
400	136	110	154		קטן
1000	351	313	336		סה"כ

מה המסקנה ברמת מובהקות של 2.5%?

2) מפעל עובד בשלוש משמרות. להלן מספר המוצרים הפוגמים והתקינים בכל

אחד מן המשמרות לפי דגם שנעשה:

פוגמים	לילה	יום	ערב
50	60	70	
600	700	800	

האם יש הבדל בין שיעורי הפוגמים במשמרות השונות?

הסיקו עבור רמת מובהקות של $\alpha = 0.05$.

3) נדגמו 50 מוצרים ממפעל מסוים מתוך 30 מוצרים שיוצרו ביום 17 נבחרו

לייצוא מתוך המוצרים שיוצרו בלילה 10 נבחרו לייצוא.

האם יש קשר בין להיות מוצר לייצוא למועד שבו הוא יוצר?

בדקו ברמת בטחון של 95%.

4) במטרה לבדוק האם השתנו דפוסי החכבה למפלגות השונות בשבוע שעבר

לשבוע נלקחו שני סקרים אחד מהשבוע שעבר והאחר מהשבוע.

להלן דפוסי החכבה שהתקבלו בסקרים אלה.

א. מהי רמת המובהקות המינימלית עבורה ניתן להחליט שהשתנו דפוסי החכבה בשבוע שעבר לשבוע באופן מובהק?

ב. כיצד הייתה התשובה לסעיף א' משתנה אם כל השכיחויות בטבלה של תוצאות המדגם היו מוכפלות פי 2?

ג. בנו רוח סמך לשיעור המצביעים למפלגה א' בשבוע ברמת סמך של 95%.

סה"כ	מפלגה א'	מפלגה ב'	מפלגות אחרות	שבוע שעבר
550	253		143	השבוע
1050		314	243	סה"כ

- 5) בוחנות בגדים A בדקו את התפלגות הצבעים של הבגדים הנמכרים ביום מסוים. כמו כן בדקו את התפלגות הצבעים בוחנות שכנה B :

מספר פריטים / צבע	כחול	אדום	לבן	שחור	
חנות A	50	15	20	15	
חנות B	20	10	20	60	

- A. בדקו ברמת מובהקות של 5% האם התפלגות הצבעים בוחנות A היא ביחס של 1:1:1:3 לטובת הכחול.
- B. בדוק ברמת מובהקות של 2.5% האם קיים הבדל בין החניות מבחינה התפלגות הצבעים של הפריטים הנמכרים.

- 6) סטודנט קיבל בבדיקה השערות ערך χ^2 (chi-square) השוו לאפס. הסטודנט הסיק כי לא קיימת תלות בין שני המשתנים שבדק, בכל רמת מובהקות. נכון / לא נכון? נמקו.

- 7) להלן טבלה O של שני משתנים שהתקבל במדגם כלשהו :

$f(x)$	Y_4	Y_3	Y_2	Y_1	
200					X_1
200					X_2
	160	120	60	60	$f(y)$

מה צרכות להיות השכיחויות בתוך הטבלה כדי שМОבהקות התוצאה (PV) תהיה 100%?

תשובות סופיות

- 1) נסיק שיש קשר בין גודל הארגון לשביעות הרצון של העובדים.
- 2) נסיק שאין הבדל מובהק בין שיעור הפגומים במשמרות השונות.
- 3) נסיק שאין קשר בין היות מוצר ליצוא למועד שבו הוא יוצר.
- 4) א. 10% ב. קטן ג. (0.223,0.297)
- 5) א. נסיק שההתפלגות הצבעים בוחנות היא כמו שמצוין.
ב. נסיק שיש הבדל בין החניות מבחינה התפלגות הצבעים.
- 6) נכון
7) להלן טבלה :

$f(x)$	Y_4	Y_3	Y_2	Y_1	
200	80	60	30	30	X_1
200	-8	60	30	30	X_2
400	160	120	60	60	$f(y)$

הקשר בין מבחנים לאי תלות ובדיקה השערות להפרש פרופורציות – רקע

מבחן לאי תלות שבו לכל משתנה יש שתי קטגוריות שקול לבדיקת השערות זו צדדיות על הפרש פרופורציות כאשר השערת האפס היא שהפרופורציות שוות. כל זאת, כמובן, אם התנאים למבחנים מתקיימים.

דוגמה (פתרון בהקלטה) :

בקרוב מוגדר מוגדר של 200 נשים 120 טענו שהן תצבעהו למועמד R בראשות העיר. בקרוב מוגדר של 200 גברים 80 טענו שהם יציבו למועמד R האם קיימן הבדל בין דפוס ההצבעה של הנשים ושל הגברים? האם אפשר לבדוק זאת גם על ידי מבחן לאי תלות וגם על ידי בדיקת השערות לשתי פרופורציות?

שאלות

- 1) בקרוב מוגדר של 200 נשים 120 טענו שהן תצבעהו למועמד R בראשות העיר. בקרוב מוגדר של 200 גברים 80 טענו שהם יציבו למועמד R האם קיימן הבדל בין דפוס ההצבעה של הנשים ושל הגברים? בדוק ברמת מובהקות של 5%.
 - א. על ידי מבחן לאי תלות.
 - ב. על ידי בדיקת השערות לשתי פרופורציות.

- 2) נלקחו 200 אנשים שמתוכם 60 הצביעו שהם עוסקים בפעילויות גופנית סידירה. מתווך אלו שעוסקים בפעילויות גופנית סידירה 50 נמצאו במצב בריאותי תקין. מתווך אלו שלא עוסקים בפעילויות גופנית סידירה 90 נמצאו במצב בריאותי תקין.
 - א. בנו טבלת שכיחות משותפת לנוטונים שהוצעו בשאלת.
 - ב. האם ניתן להגיד שהסיכוי להימצא במצב בריאותי תקין גבוה יותר כאשר עוסקים בפעילויות גופנית סידירה לעומת המצב שלא עוסקים בפעילויות גופנית סידירה? בדוק ברמת בטחון של 90%.

תשובות סופיות

- 1) נדחה את השערת האפס לפירוט הדרך ראה וידאו.
 2) א. להלן טבלה : ב. נדחה את השערת האפס.

	לא תקין	תקין	פעילות/מצב בריאותי
60	10	50	סדייה
140	50	90	לא סדייה
200	60	140	

סטטיסטיקה ב

פרק 19 - ניתוח שונות חד כיוונית

תוכן העניינים

1. כללי

123

ניתוח שונות חד כיוונית

רקע תיאורטי

ניתוח שונות (חד כיווני) הוא מבחן להשוואת תוחלות (μ_1, \dots, μ_k) של k אוכלוסיות שונות. לכן, בניתוח שונות, השערות המחקר הן:

$H_0: \mu_1 = \mu_2 = \dots = \mu_k$ (התוחלות של כל האוכלוסיות שוות)

$H_1:$ אחרת (פחות שתיים מהתוחלות שונות)

הנחהות הדרושות לביצוע התהילה:

- 1) בכל אוכלוסייה מטוך k האוכלוסיות ההתפלגות נורמלית.
- 2) כל האוכלוסיות הן עם אותה שונות σ^2 .
- 3) המדגמים בלתי תלויים זה בזו.

ישנו משתנה המבדיל בין הקבוצות השונות, הוא המשתנה הבלתי תלוי הנקרא גורם (factor). משתנה זה הוא קטגוריאלי עם k רמות (levels). כדי לבצע את התהילה יש לבצע מוגם מכל אוכלוסייה: נסמן ב- n_i את גודל המוגם בקבוצה i .

$$\sum_{i=1}^k n_i = n - \text{מספר הרציפות סך הכל (בכל המדגמים)}.$$

\bar{X}_i - ממוצע המוגם הראשון, ..., \bar{X}_k - ממוצע המוגם ה- k -י.
 \bar{X} - ממוצע כללי (של כל המדגמים).

$$\text{סכום ריבועים בין הקבוצות: } SS_B = \sum_{i=1}^k n_i [\bar{X}_i - \bar{X}]^2$$

$$\text{סכום ריבועים בתוך הקבוצות: } SS_W = \sum_{i=1}^k n_i [n_i - 1] \cdot \hat{s}_i^2$$

$$\text{סכום ריבועים כללי: } SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} [X_{ij} - \bar{X}]^2$$

$$SST = SSB + SSW$$

יש למלא את טבלת ניתוח השונות הבאה :

מקור השונות	סכום הריבועים SS	דרגות חופש df	ממוצע הריבועים MS	F
B - בין הקבוצות	SSB	$k-1$	$\frac{SSB}{k-1}$	$\frac{MSB}{MSW}$
W - בתוך הקבוצות	SSW	$n-k$	$\frac{SSW}{n-k}$	
T - סה"כ	SST	$n-1$		

$$F = \frac{\frac{SSB}{k-1}}{\frac{SSW}{n-k}} \sim F(k-1, n-k)$$

$$\text{אזור דחיה: } H_0 : F > F_{(k-1, n-k)} : 1-\alpha$$

שאלות

- 1)** מבחן מעוניין להשוות בין שלוש תרופות לשיכוך כאבים במטרה לבדוק האם קיים הבדל בין התרופות מבחינת הזמן בדיקות שלוקח עד שהתרופה משפיעה. לצורך הבדיקה נלקחו 15 אנשים שסובלים מכאבי ראש. אנשים אלה חילקו באקראי לשולש: קבוצה 1 קיבלה "אקסמול" קבוצה 2 קיבלה "אופטיגין" קבוצה 3 קיבלה "נוורופן". כל אדם במחקר מסר את מספר הדיקות עד שהתרופה השפיעה עליו.
- מהו המשנה תלוי ומהו המשתנה הבלתי תלוי במחקר?
 - מהו המבחן הסטטיסטי המתאים כאן? רשמו את ההשערות.
 - מה הן הנקודות הדרושות כדי לבצע את המבחן הסטטיסטי שהצעת בסעיף הקודם?
- 2)** בעיר מסוימת שלושה בתים ספר תיכון. ראש העיר התעניין לבדוק האם קיים הבדל בהצלחה של בתים הספר במקצוע מתמטיקה. לצורך כך הוא דגם מספר תלמידים שנבחנו בבחן הבגרות במתמטיקה ברמה של 3 יחידות בעיר ובדק עבור כל תלמיד מה ציון הבגרות שלו במתמטיקה. להלן הציונים שהתקבלו:
- | "הס" | "הטמיד" | "רבין" |
|------|---------|--------|
| 85 | 98 | 78 |
| 83 | 62 | 65 |
| 74 | 55 | 70 |
| 85 | 80 | 90 |
| 75 | | 56 |
- מהו המבחן הסטטיסטי המתאים? רשמו את ההשערות ואת הנקודות של המבחן.
 - מהו גודל המדגמים? מהו המשנה הבלתי תלוי (factor) כמה רמות יש לו?
 - חשבו את הממוצע ואת סטיית התקן של הציונים בכל אחד מהمدגמים.
 - מלאו את טבלת ANOVA.
 - רשמו את כלל ההכרעה לבחן שהוצע בסעיף א ברמת מובהקות של 5%.
 - האם קיים הבדל בין בתים הספר בעיר מבחינה רמת הצלחת התלמידים במקצוע המתמטיקה? ענה על סמן הסעיפים הקודמים.
- 3)** מעוניינים לבדוק האם יש הבדל בהשפעה של שיטות טיפול שונות על לחץ הדם הסיסטולי (SBP) באוכלוסייה של קשישים. נבדקו 4 שיטות שונות.
- | D | C | B | A | השיטה |
|-----|-----|-----|-----|------------|
| 12 | 8 | 14 | 12 | גודל המדगם |
| 182 | 180 | 172 | 178 | הממוצע |
| 3 | 5 | 8 | 4 | סטיית התקן |
- בטבלה המצורפת מ羅וצזים מצאי המחקר.
- רשמו את השערות המחקר והנקודות הדרושות כדי לבצע את המבחן המתאים.
 - מה מסקנת המחקר ברמת מובהקות של 5%?
 - האם יש צורך לבצע השוואות מרובות?

- 4) שלושה אופים נתקשו להכין עוגת שוקולד. לכל אופה בדקו את משך הזמן ההכנה בדקות. כל אופה נדרש לאפות בכל יום 4 עוגות. האם קיים הבדל בין האופים מבחינת תוחלת זמן ההכנה של העוגות?

בדקו ברמת מובהקות של 5%.

שלוט	מוסס	ניר	האופה
			סכום הזמנים
			סכום ריבועי הזמנים
182	212	206	
8982	11250	10644	

- 5) להלן טבלת ניתוח שונות חד כיוונית. במחקר בחנו 4 סוגי סוללות. רצוי לבדוק האם לסוג הסוללה השפעה על תוחלת אורץ החיים שלה. הפעלו את כל הסוללות על אותו מכשיר ובדקו את אורץ החיים של כל סוללה בשעות. מה המסקנה ברמת מובהקות של 10%? רשמו את ההשערות וההנחות הדרשיות.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10.317	3	3.439	1.361	.279
Within Groups	60.648	24	2.527		
Total	70.964	27			

- 6) להלן טבלת ANOVA בטבלה הושמו חלקים. השלימו את החלקים בטבלה שהושמו ומסומנים באותיות.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	357.450	ב	ג	ה	.000
Within Groups	א	17	ד		
Total	522.950	19			

7) חברת תרופות לκחה 15 אנשים ברמת בריאות דומה. החברה חילקה את האנשים לשולש קבוצות שוות בגודלו. לכל קבוצה ניתנה אותה תרופה במינון שונה (dosage). המינונים שניתנו הם : 10 מ"ג, 20 מ"ג ו-30 מ"ג. לאחר שעה מזמן לקיחת התרופה ניבדק קצב פעימות הלב של כל אדם (pulse). הנתונים הוזנו לתוכנה סטטיסטית והתקבלו התוצאות הבאות:

ANOVA						pulse			
pulse						Tukey HSD ^a			
	Sum of Squares	df	Mean Square	F	Sig.	dosage	N	Subset for alpha = 0.05	
Between Groups	414.400	2	207.200	19.733	.000	30.00	5	71.0000	
Within Groups	126.000	12	10.500			20.00	5	80.2000	
Total	540.400	14				10.00	5	83.4000	

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

Post Hoc Tests

Multiple Comparisons

		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
(I) dosage	(J) dosage				Lower Bound	Upper Bound
10.00	20.00	3.20000	2.04939	.299	-2.2675	8.6675
	30.00	12.40000*	2.04939	.000	6.9325	17.8675
20.00	10.00	-3.20000	2.04939	.299	-8.6675	2.2675
	30.00	9.20000*	2.04939	.002	3.7325	14.6675
30.00	10.00	-12.40000*	2.04939	.000	-17.8675	-6.9325
	20.00	-9.20000*	2.04939	.002	-14.6675	-3.7325

*. The mean difference is significant at the 0.05 level.

- בדקו ברמת מובהקות של 5% האם קיים הבדל בין המינונים השונים מבחינית תוחלת הדופק של האנשים? רשמו את ההשערות וההנחות הדרושים לצורך פתרון.
- הסבירו ללא חישוב כיצד הייתה משתנה התשובה לסעיף הקודם אם היינו מעלים את הדופק של כל התצפיות במחקר ב-2.
- אם יש צורך במחקר בהשואת מרובות. נמקו!
- לטבלת ANOVA צורפו טבלאות של השוואות מרובות בשיטה הנkrat "טוקי". ברמת בטחון של 95% מה הם הממצאים לפי שיטה זו?

- 8) בעיר מסוימת רצוי לבדוק האם קיימים הבדל ברמה של התלמידים בין בתיה הספר השונים בעיר. ביצעו מוגם מכל בית ספר ונתנו מבחן זהה לכל הנדגמים. לאחר מכן ריכזו את הנתונים בתוכנה סטטיסטית והפיעלו ניתוח שונות. מצורפים הפלטים שהתקבלו. ענו על השעיפים הבאים:
- כמה בתיה ספר יש בעיר?
 - כמה תלמידים השתתפו בסך הכל במחקר?
 - האם קיימים הבדל בין בתיה הספר בעיר מבחינה רמת הציונים? בדקו ברמת מובהקות של 1%
 - בביחוון של 95% אילו בתיה ספר שונים זה מזה ברמת התלמידים? נמקו והסבירו.

Oneway

ANOVA

grade	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7799.600	4	1949.900	13.586	.000
Within Groups	2870.400	20	143.520		
Total	10670.000	24			

Post Hoc Tests

Multiple Comparisons

grade

Scheffe

(I) school	(J) school	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	5.40000	7.57681	.971	-20.2543	31.0543
	3.00	36.80000*	7.57681	.003	11.1457	62.4543
	4.00	36.40000*	7.57681	.003	10.7457	62.0543
	5.00	-2.60000	7.57681	.998	-28.2543	23.0543
2.00	1.00	-5.40000	7.57681	.971	-31.0543	20.2543
	3.00	31.40000*	7.57681	.011	5.7457	57.0543
	4.00	31.00000*	7.57681	.013	5.3457	56.6543
	5.00	-8.00000	7.57681	.888	-33.6543	17.6543
3.00	1.00	-36.80000*	7.57681	.003	-62.4543	-11.1457
	2.00	-31.40000*	7.57681	.011	-57.0543	-5.7457
	4.00	-.40000	7.57681	1.000	-26.0543	25.2543
	5.00	-39.40000*	7.57681	.001	-65.0543	-13.7457
4.00	1.00	-36.40000*	7.57681	.003	-62.0543	-10.7457
	2.00	-31.00000*	7.57681	.013	-56.6543	-5.3457
	3.00	.40000	7.57681	1.000	-25.2543	26.0543
	5.00	-39.00000*	7.57681	.001	-64.6543	-13.3457
5.00	1.00	2.60000	7.57681	.998	-23.0543	28.2543
	2.00	8.00000	7.57681	.888	-17.6543	33.6543
	3.00	39.40000*	7.57681	.001	13.7457	65.0543
	4.00	39.00000*	7.57681	.001	13.3457	64.6543

*. The mean difference is significant at the 0.05 level.

Homogeneous Subsets

Scheffe ^a	school	N	grade	
			Subset for alpha = 0.05	
			1	2
	3.00	5	45.0000	
	4.00	5	45.4000	
	2.00	5		76.4000
	1.00	5		81.8000
	5.00	5		84.4000
Sig.			1.000	.888

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

תשובות סופיות

ב. ניתוח שונות חד כיווני

1) א. משתנה בלתי תלוי : סוג התרופה.

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \text{otherwise}$$

משתנה תלוי : הזמן עד להשפעת התרופה בדקות.

ג. 1. מדגמים בלתי תלויים.

2. שוויון שוניות.

3. משתנים מתפלגים נורמלית.

2) א. המבחן לניתוח שונות חד כיוונית.

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \text{otherwise}$$

הנחות :

1. מדגמים בלתי תלויים.

2. משתנים מתפלגים נורמלית.

3. שוויון שוניות.

ב. גודל המדגם : 14. משתנה בית' : בית הספר, בעל 3 רמות.

$$\bar{X} = 80.4, \hat{S} = 5.46, \bar{X} = 73.75, \hat{S} = 19.29, \bar{X} = 71.8, \hat{S} = 12.93$$

ג. להלן טבלה :

F	MS	df	SS	מקור השונות
	100.3	2	200.6	B
	173.2	11	1904.75	W
0.58		13	2105.35	סה"כ

ה. $F > 3.98$.

ו. קיבל את H_0 .

ג. כן.

ב. נדחה את H_0 .

3) א. $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$
 $H_1: \text{otherwise}$

הנחות :

1. מדגמים בלתי תלויים.

2. שוויון שוניות.

3. משתנים מתפלגים נורמלית.

(4) קיבל את H_0 : נכריע שאין הבדל מובהק בין האופים מבחינת תוחלת זמן הכנה.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \quad (5)$$

$$H_1: \text{otherwise}$$

הנחות :

1. מדגמים בלתי תלויים.

2. שוויון שוניות.

3. משתנים מתפלגים נורמלית.

נקבל את H_0 : לסוג סוללה אין השפעה של תוחלת החיים ברמת ביטחון של 10%.

(6) להלן טבלה :

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	357.450	2	178.725	ג. 18.36	.000
Within Groups	165.5	17	9.735	ד.	
Total	522.950	19			

$$H_0: \mu_1 = \mu_2 = \mu_3 \quad (7) \text{ א.}$$

$$H_1: \text{otherwise}$$

הנחות :

1. מדגמים בלתי תלויים.

2. משתנים מתפלגים נורמלית.

3. שוויון שונייה.

נדחה את H_0 : ברמת ביטחון של 5% קיים הבדל במינונים השונים מבחינת תוחלת הדופק.

$$\text{ב. ראה וידעו.} \quad \text{ג. כן.} \quad \text{ד. } \mu_{30} > \mu_{10} \quad \text{ה. 25.} \quad \text{א. 5.}$$

ג. נדחה את H_0 : יש לפחות שני בתים ספר בעיר עם תוחלת רמת ציונים שונה.

$$\text{ד. } (\mu_3 = \mu_4) < (\mu_1 = \mu_2 = \mu_3).$$