מטלה 4

סטטיסטיקה להנדסת תוכנה

א. בסקר שביעות רצון ממסגרות החינוך שנערך בערים חיפה ות"א נאספו הממצאים הבאים. כל תא מייצג את כמות אנשים שהשיבו על השאלה האם את/ה מרוצה ממסגרות החינוך? בתשובה מרוצה או לא מרוצה:

	חיפה	תל אביב	סה"כ
מרוצים	18	17	35
לא מרוצים	6	19	25
סה"כ	24	36	60

- 1. נסחו השערה סטטיסטית למציאת הבדל בין שביעות הרצון בין חיפה לתל אביב. הערה: ניתן לנסח את ההשערה עבור המרוצים, אין צורך עבור שתי הקטגוריות.
- 2. חשבו את האומדן הסטטיסטי לבדיקת ההשערות שנסחתם ב-1, והסיקו מסקנה ברמת מובהקות 2 $\chi^2_{(1,0.975)}=5.02, \quad \chi^2_{(1,0.95)}=3.84, \quad \chi^2_{(1,0.90)}=2.71$ יש לבחור את הערך המתאים)
 - 3. על סמך מסקנתכם בשאלה 2, מה ניתן לומר לגבי ה-Pvalue של המדגם?
 - 4. חשבו את ה-Odd-Ratio בהתבסס על נתוני המדגם. מה ניתן להסיק ממדד זה לגבי התפלגות שביעות הרצון בין הערים?

ב. בבדיקת הקשר בין גיל לפעילות גופנית, סוכמו הנתונים באופן הבא:

קבוצת הגיל	גודל המדגם (n _i)	פעילות גופנית	
		(\overline{Y}_l) ממוצע	(s¡) סטיית התקן
A: 18-25	13	9.623	1.69
B: 26-35	20	8.045	1.32
C: 36-50	15	7.640	1.63
D: 51-60	12	6.775	1.55
All	60	8.032	

דווח גם כי סכום ריבועי ההפרשי הממוצעים מהממוצע הכללי של מדד הפעילות הגופנית (SS_{between}) דווח גם כי סכום ריבועי ההפרשי הממוצעים מהממוצעת (MSE) היא 54.18,

- 1. נסחו השערת אפס והשערת אלטרנטיבית לבדיקת שוויון תוחלות מספר שעות הפעילות הגופנית הממוצעת בין קבוצות הגיל השונות.
 - 2. סכמו את הנתונים בטבלת ה-ANOVA הבאה: SSerror ,MSbetween- החסרים: ה-חסרים: את האומדן הסטטיסטי, F.

Source	df	SS	MS	F
Between groups	3	54.18		
Error	56		2.344	

3. מה משמעות מדד ה-MShetween?

נתון כי 2.769 =(0.95,df1=3,df2=56) האם ניתן לדחות את השערת האפס? נתון כי 2.769 =(0.95,df1=3,df2=56) נסחו את המסקנה הסטטיסטית.

.4 בהשוואות contrast בין תוחלות כל זוגות הקבוצות האפשריים, נמצאו רמות המובהקות הבאות. נתחו את ההבדלים בשיטת Holms (עבור 20.05) וסכמו את הממצאים.

Groups	Pv
D-A	0.000125
C-A	0.005909
B-A	0.021698
D-B	0.080907
D-C	0.300442
С-В	0.441925

ג. בבית חולים מסויים נמדדו המשקל והאורך של 30 תינוקות בלידה. (נתונים פיקטיביים) חלק מהתוצאות מפורטות בטבלה הבאה:

index	Height (cm)	Weight (Kg)	
1	66.13	2.56	
2	66.04	2.42	
30	71.00	3.74	
Sample Mean	68	3	
Sample standard deviation	3.837	0.596	

- 1. איזה תנאי צריך להתקיים על מנת שנוכל לבחון את ההתכנות הסטטיסטית למתאם בין Pearson המשתנים לפי מבחן
 - לצורך העניין, נניח שתנאי זה מתקיים.כתבו את ההשערה הסטטיסטית לאומדן המתאם בין האורך למשקל.מה מספר דרגות החופש של האומדן הסטטיסטי במבחן?
- הדגימו חישוב של מקדם המתאם עבור שורה 1.
 הסבירו מילולית ולוו במשוואה מתאימה כיצד הייתם מחשבים את מקדם המתאם בין המשתנים,
 על בסיס החישוב שהדגמתם עבור השורה הראשונה.
 - 4. נתון כי מקדם המתאם המחושב הוא 0.879, סטיית התקן של משתנה הגובה היא 3.837, סטיית התקן של משתנה משקל התינוק היא 0.596. כמו כן, נתון כי האומדן הקריטי לבדיקת תוחלת המתאם הוא 2.048 ברמת מובהקות של 5%. חשבו את האומדן הסטטיסטי למקדם המתאם.
 - 5. מה ניתן להסיק מאומדן זה? האם ניתן לדחות או לקבל את השערת האפס?
 - .0.82 מדגם ובו 28 תצפיות נאסף, עם ממוצע מדגם 3.8 וסטיית תקן במדגם של

משוער, כי התוחלת היא 4.5, והשונות באוכלוסיה אינה ידועה. (ני התוחלת היא 4.5, והשונות באוכלוסיה אינה ידועה) נתון: $t_{0.975.27}=2.0518,\;t_{0.95.27}=1.703$

- 1. נסחו השערה לבדיקת התכנות התוחלת המשוערת.
- 2. מהי הנחת המבחן הסטטיסטי לבדיקת השערה זאת?

- 3. חשבו את רווח הסמך סביב ממוצע המדגם. האם ניתן לדחות או לקבל את ההשערה על בסיס רווח הסמך שחישבתם?
- 4. חשבו את האומדן הסטטיסטי להתכנות ההשערה. האם ניתן לדחות או לקבל את השערת האפס α=0.05 לפי האומדן הסטטיסטי?