

מטלה 2: סטטיסטיקה להנדסת תוכנה

שאלה זאת עוסקת בקובץ הנתונים "triglycerides.csv".
הנתונים מפרטים רמות שומנים בדם (טריגליצרידים) עבור קבוצה של 15 אנשים עם מחלת לב (congestive heart disease), ו-23 איש מביקורת עצמאית.
מטרת התרגיל היא לגלות האם לחולי הלב יש רמות שומנים בדם גבוהות מאשר לאנשים בקבוצת הביקורת.

שאלה 1

נסחו השערה סטטיסטית התואמת את שאלת המחקר.
בדקו האם מתקיימות הנחות המודל למבחן t לבדיקת שוויון תוחלות:
- הדגימו את התפלגות הנתונים המשותפת.
- בצעו מבחן לבדיקת שוויון שונויות.
- בצעו מבחן לבדיקת נורמליות של ההתפלגות המשותפת.
בצעו מבחן מתאים להשוואת התוחלות. נסחו את מסקנותיכם בהתבסס על תוצאות המבחן.

שאלה 2

רמת טריגליצרידים נורמלית בדם נחשבת בטווח ערכים של עד 170mg/ml כולל.
בהתחשב בהגדרה זאת, השוו את רמות הטריגליצרידים בדם (גבוהה/נורמלית) בשתי קבוצות המחקר.
שימו לב לקיום התנאים הדרושים לביצוע המבחן. פרטו את החישובים שהובילו לתוצאות המבחן, בצעו מבחן שווה ערך לבדיקת ההשערה ב-R והציגו את תוצאותיו. סכמו את הממצאים בהסקה סטטיסטית תואמת.

שאלה 3

דיון תאורטי:

א. במה דומות/שונות התוצאות שנתקבלו בשאלה 1 ובשאלה 2?
ב. מהם ההבדלים במאפייני המבחנים שהובילו להבדלים בתוצאות הסטטיסטיות?
התייחסו בתשובתכם להנחות כל אחד מהמודלים, להבדלים בכמות אינפורמציה של הנתונים עליה מבוסס כל מבחן ולהבדלים הנגזרים מכך ברמת המובהקות הסטטיסטית המתקבלת.