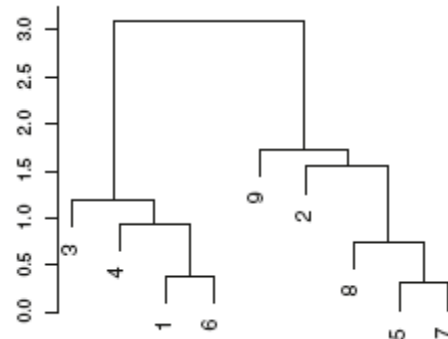


דוגמאות למבחן

1. נתונה הדנדרוגרמה הבאה:



linkagen בו השתמשנו כדי לכמת מרחק בין clusters הוא complete linkage. מה לא נכון?

1. המרחק בין 1 ל 6 קטן מהמרחק בין 4 ל 6.
2. המרחק בין 1 ל 4 גדול מהמרחק בין 1 ל 6.
3. המרחק בין 5 ל 7 קטן מהמרחק בין 5 ל 8.
4. המרחק בין 8 ל 7 קטן מהמרחק בין 5 ל 7.

נתונים נתוני gene expression של חולים המשתייכים לשתי קבוצות: Bi A. חוקר בדק שינויים ב gene expression בין החולים המשתייכים לשתי הקבוצות. על כל גן התבצע מבחן t לבדיקת הבדלים ברמות הביטוי. החוקר קבע את הרף לדחיית השערת האפס (אלפה) להיות 0.01.

1. מהי המשמעות של רף דחית השערת האפס?
2. בהנתן גן עם p value של 0.0005, מה הסיכוי לראות רמת ביטוי כזו של הגן תחת השערת האפס?
3. בהתאם לרף שנבחר, החוקר גילה 8,000 גנים שנמצאו סיגניפיקנטית שונים בין שתי הקבוצות. מה הבעיה במה שעשה החוקר?
4. כפי שנלמד בכיתה, כיצד ניתן לתקן שגיאה זו?
5. לאחר התיקון שביצע החוקר, כמות הגנים הסיגניפיקנטיים שנמצאו בין שתי הקבוצות ירד ל 0. החוקר החליט לנקוט באסטרטגיה אחרת ולהוריד את רף דחית השערת האפס תוך שיערוך של השגיאה מסוג ראשון. כיצד יעשה זאת?

תשובות.

שאלה ראשונה: **תשובה 4** – מאחר זוג הנקודות: 1,6 חוברו ראשונים בהיררכיה – המרחק ביניהם קטן מאשר מרחק כ"א מהנקודות עם הנקודות העליונות יותר בהיררכיה. כנ"ל לגבי זוג הנקודות: 5,7 ביחס לנקודה 8.

שאלה שניה:

1. סיכוי לטעות מסוג ראשון
2. 0.0005
3. השוואות מרובות ללא תיקון אלפה
4. תיקון בונפרוני לאלפה

5. ישערך שגיאה מסוג ראשון באמצעות פרמוטציות על הדוגמאות (FDR), ובהתאם יוריד את האלפה (יבחר אלפה עם FDR סביר).