

國立中央大學101學年度碩士班考試入學試題卷

所別：系統生物與生物資訊研究所碩士班 不分組(一般生) 科目：計概與資料結構 共 1 頁 第 1 頁  
本科考試禁用計算器 \*請在試卷答案卷(卡)內作答

人類有 3 萬個基因( $G_1, G_2, G_3, \dots, G_{30000}$ )，使用微陣列技術取得  $k$  個正常人 ( $N_1, N_2, N_3, \dots, N_k$ ) 與  $r$  個癌症病患 ( $C_1, C_2, C_3, \dots, C_r$ ) 的組織檢體的基因表現量(每個基因表現量為一實數)。每個人的基因表現使用一個矩陣儲存。

回答以下問題:

1. 請使簡單的虛擬碼寫出如何排序某個人的所有基因表現量演算法(30points)，並評估你的演算法的時間複雜度。(5point)
2. 承一，每個人每個基因都有一個排序值，請使簡單的虛擬碼寫出正常組病患每個基因的表現排序值(rank function)在這  $k$  個病患中的平均值(10points)，並評估你的演算法的時間複雜度。(5point)
3. 承二，請使簡單的虛擬碼寫出基因表現值在正常/癌症表現值前後 10% 中呈現相反現象基因。亦即:
  - (1) 正常組病患基因的表現的排序平均值的前 10% 高與癌症組病患基因表現的排序平均值的後 10% 中相同的基因。(20 point)。
  - (2) 正常組病患基因的表現的排序平均值的後 10% 高與癌症組病患基因表現的排序平均值的前 10% 中相同的基因。(20 point)。
4. 請評估以上第三題的演算法的時間複雜度。(10 point)。