

國立中央大學100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)。

回答以下問題：

1. 舉例說明最簡單儲存兩組病患的基因表現的資料結構。(20 points)
2. 請使簡單的虛擬碼寫出兩組病患每個基因的表現的平均值(20points)與標準差(30points)。(標準差公式： $\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$)
3. 請使簡單的虛擬碼找出(正常人的基因平均表現量)低於(癌症病患基因平均表現量減掉兩個癌症基因表現標準差)的基因列表。(30points)

參考用