

第一題 (4 小題，共 50 分) ※ 注意：請於試卷內之「非選擇題作答區」依序作答，並應註明作答之大題及小題題號。

指數分佈與布瓦松分佈都是處理與時間相關統計分佈。請應用此兩種分佈於 COVID-19 流行確診個案接觸調查。假設 100 位接觸者在與確診個案不同時間接觸 14 天內分別有 30 位感染確診，請自訂符號並用於回答以下問題。

- (1) 寫出指數分佈及布瓦松分佈機率密度(質量)函數及累積分佈；說明兩者共同參數，並解釋此參數代表意義及所對應參數於兩種分佈之下的充分統計量(15%)。
- (2) 請利用上述資料寫出兩種分佈概似函數(Likelihood function)；如何求最大概似估計值(Maximum Likelihood Estimate, MLE) (10%)。
- (3) 證明若接觸者無任何感染下兩者機率密度(質量)函數分佈相同；導出半衰期(半數接觸者發生感染的時間)與參數之關係；預測 10 天內接觸者感染機率(15%)。
- (4) 如何應用此兩種分佈來估計在控制年齡後接觸者打疫苗及未打疫苗兩者感染狀況不同 (10%)。

第二題 (5 小題，每小題 10 分，共 50 分)

欲研究某「單核苷酸多態性」(Single-Nucleotide Polymorphism, SNP) 是否與「身體質量指數」(body mass index, BMI, 單位：公斤/公尺平方) 有關。該 SNP 有三種基因型 AA, Aa, aa，下表列出三種基因型下的 BMI 資料：

基因型 AA	基因型 Aa	基因型 aa
27	25	24
25	25	24
24	24	23
25	23	23
25	24	23
24	22	24
27	24	25
24	23	
26		
26		

- (1) 請以「變異數分析」(Analysis of variance, ANOVA)檢定該 SNP 是否與 BMI 有關。顯著水準設為 0.05。請寫出虛無假說、對立假說、檢定統計量、虛無假說下該檢定統計量之分布、拒絕域。如有自設符號請定義。
- (2) 請以「迴歸分析」(regression analysis)檢定該 SNP 是否與 BMI 有關。顯著水準設為 0.05。請寫出虛無假說、對立假說、迴歸係數的意義、檢定統計量、虛無假說下該檢定統計量之分布、拒絕域。如有自設符號請定義。
- (3) 如定義 BMI  $\geq 25$  (公斤/公尺平方) 為肥胖，表中連續型資料可轉為二元資料 (肥胖 vs. 正常)。請以「羅吉斯迴歸」(logistic regression)檢定該 SNP 是否與肥胖有關。顯著水準設為 0.05。請寫出虛無假說、對立假說、迴歸係數的意義、檢定統計量、虛無假說下該檢定統計量之分布、拒絕域。如有自設符號請定義。
- (4) 承上一小題，請寫出「羅吉斯迴歸」中迴歸係數的「概似函數」(likelihood function)。並請推導如何求得「羅吉斯迴歸」中迴歸係數的「最大概似估計值」(maximum likelihood estimate, 簡寫為 MLE)。
- (5) 某研究者論述：「若該 SNP 與某環境因子間存在交互作用，則 BMI 在 AA, Aa, aa 三組的變異數將不等。」請推導證明此論述是否正確。

試題隨卷繳回