國立臺灣大學 103 學年度碩士班招生考試試題

I. 計算與證明題，每題 5 分 (共 35 分，請附上計算過程)

1. 機率密度函數為 $f(x)=1$ 且 $-0.5<x<0.5$。已知 $y=x^2$，請求 y 的機率密度函數 $g(y)$。
2. 小霞蒐集下列時間序列資料 y_t，建立自我迴歸(autoregressive regression, AR) 模型(t 為時間): $y_t=c+dy_{t-1}+e_t$，請問最小平方方法估計下的估計式 $\hat{c}$ 和 $\hat{d}$ 為何?

t	1	2	3	4	5	6	7	8	9
y_t	1	2	2	3	2	8	3	15	7

3. 承上題，請問 $\hat{d}$ 的標準差為何?
4. 承上題，若小霞認為 y_t 有變異數不齊一問題，欲用 White (1980)建議的估計式，請問 White (1980) 估計式之下 $\hat{d}$ 的標準差是多少?
5. 小智研究神奇寶貝訓練家的性別和訓練家對傑尼龜的喜好與否是有關，小智蒐集樣本如下。

	喜歡	不喜歡
男性	80	20
女性	60	40

小智欲以邏輯迴歸(logistic regression)分析之，令 $x=1$ 表示男性(0 為女性)，$y=1$ 表示喜歡(0 表示不喜歡)。請寫下 $Prob(y=1)=L(a+bx)$的邏輯迴歸方程式。

6. 承上題，請求解 a 和 b 迴歸估計式(提示: ln(4)=1.386，ln(1.5)=0.405)。
7. 請證明才比雪夫不等式(Chebyshev inequality)成立

II. 是非題，每題 5 分 (共 15 分，錯誤須說明並更正)

1. 建立自我迴歸模型 AR(1): $y_t=a_0+a_1y_{t-1}+e_t$, $e_t\sim(0,1)$，我們得到 $E(y)=a_0$。
2. 給一隨機變數 x，樣本標準差為 $\hat{s}=\sqrt{\dfrac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N-1}}$ 為母體標準差 σ_x 的不偏估計式。
3. 若 X 和 Y 為兩組 $i.i.d.$ 常態隨機變數，其變異數未知但 X 和 Y 的變異數不相等，X 和 Y 樣本數分別為 n 和 m；則樣本平均數差 $(\bar{X}-\bar{Y})$ 遵循 t 分配自由度$(n+m-2)$。

III. 問答與演算題 (共 50 分)

1. 設 $\{X_i\}^3 \overset{i.i.d.}{\sim} Pareto(\alpha,\sigma)$ 為一組來自 Pareto 母體之隨機樣本，其機率密度函數(pdf)為

$$f_X(x)=\begin{cases}(\alpha/\sigma)(x/\sigma)^{-\alpha-1}, & x\geq\sigma\\ 0 & \text{，其他範圍。}\end{cases}$$

令 Y_i 為 $\{X_i\}^3$ 之第 i 階順序統計量，$i=1,2,3$,

(1) 求(Y_1,Y_2,Y_3)之聯合 pdf。 (3 分)
(2) 求極小統計量 Y_1 之 pdf，並註明其分配名稱及參數。(4 分)
(3) 求樣本中位數(sample median)之 pdf。(3 分)

見背面

2. 令$\{X_i\}^n$為一組由二項分配$B(5,p)$母體所抽出之隨機樣本，令$\overline{X}_n=\sum_{i=1}^{n}X_i/n$表示樣本平均數。

(1) 當樣本數大小 n<30 且固定，試推導$\overline{X}_n$之精確分配 pmf (probability mass function) (5 分)

(2) 當樣本數大小 $n\geq 30$ 且固定，試寫出$\overline{X}_n$之漸近分配，需說明理由及註明分配名稱與參數。(5 分)

(3) 設參數p未知，試求p之最大概似估計元，記為$\hat{p}$。
請說明$\hat{p}$是否具有 UMVUE (uniformly minimum variance unbiased estimator)? (5+10 分)

3. 設$\{X_i\}^n \overset{i.i.d.}{\sim} N(\mu,1)$為一組來自常態分配之隨機樣本，設$\mu$未知

(1) 寫出μ之參數空間Ω，及$\sum_{i=1}^{n}(X_i-\mu)^2$與$\sum_{i=1}^{n}(X_i-\overline{X})^2$之分別分配名稱與參數。(1+2+2 分)

(2) 以顯著水準 α=0.05，檢定 H_0: μ=0 v.s H_1: μ=1，試推導一最佳拒絕域(The best critical region)之檢定。(10 分)