

類 科：統計

科 目：抽樣方法與迴歸分析

考試時間：2小時

座號：\_\_\_\_\_

※注意：(一)可以使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

(三)本科目除專門名詞或數理公式外，應使用本國文字作答。

一、假設某城市共有人口為 $N = 24000$ 人，為了研究此城市的失業人數，於是  
以簡單隨機取樣自此城市中抽取 500 人，結果發現有 10 人失業。試求：  
(每小題 5 分，共 25 分)

(一)此城市人口失業率的估計值。

(二)此城市人口失業率估計量的標準誤的估計值。

(三)此城市人口失業率估計量的 95% 近似誤差界限 $B$ 。

(四)此城市人口失業率的 95% 近似信賴區間。

(五)若我們希望估計此城市人口失業率的誤差界限為 $B = 5\%$ ，則須要在此  
城市抽取多少人數才能達到我們的需求？(註： $Z_{0.025} = 1.96$ )。

二、給定一分層隨機樣本如下：

層 ( $h$ )	$N_h$	$n_h$	$\bar{y}_h$	$s_h$
1	200	20	150	30
2	200	40	100	20
3	100	20	200	50

其中 $N_h$ 表示第 $h$ 層的母體大小， $n_h$ 表示第 $h$ 層的樣本大小， $\bar{y}_h$ 表示第 $h$ 層的  
樣本平均數， $s_h$ 表示第 $h$ 層的樣本標準差。試求：(每小題 5 分，共 25 分)

(一)每一層母體平均數的分層估計值。

(二)抽樣母體平均數的分層估計值。

(三)分層估計量標準誤的估計值。

(四)抽樣母體平均數的 95% 近似誤差界限 $B$ 。

(五)抽樣母體平均數的 95% 近似信賴區間。(註： $Z_{0.025} = 1.96$ )。

三、對一組樣本大小  $n = 100$  的資料配適一複迴歸模型

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i,$$

其中  $\epsilon_i$  為 iid  $N(0, \sigma^2)$ 。進行分析後，得到如下結果：

$$\hat{\sigma} = 5, s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2} = 10。$$

(一) 依據前述資訊，完成下列變異數分析表（將此表繪製於試卷上，並寫出詳細計算過程，再將結果填入表中）：（9分）

變異來源	平方和	自由度	均方	F
迴歸	(a)	(d)	(g)	(i)
誤差	(b)	(e)	(h)	
總和	(c)	(f)		

(二) 若要檢定  $\beta_1, \beta_2, \beta_3$  是否同時為 0，請列出虛無假設、對立假設、F 值的分配（需標明自由度）、以及在顯著水準  $\alpha = 0.05$  下拒絕虛無假設的條件。（8分）

(三) 計算  $R^2$ ，並說明其意義。（8分）

四、考慮配適一簡單線性迴歸模型： $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, i = 1, \dots, n$ ，並假設  $\epsilon_i$  為 iid  $N(0, \sigma^2)$ 。

(一) 請寫出模型中的應變數與自變數。（5分）

(二) 請問 iid 是那三個英文文字的縮寫，是代表什麼假設？請詳細說明。（5分）

(三) 若以最小平方估計式  $\hat{\beta}_0$  及  $\hat{\beta}_1$  得到  $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ ，且令  $e_i = Y_i - \hat{Y}_i$ 。請問  $\sum_{i=1}^n X_i e_i = ?$  請詳細列出推導過程。（15分）