

類 科：統計

科 目：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。

一、今對一組樣本資料 (x_i, y_i) , $i = 1, \dots, 20$ 配適一簡單線性迴歸模型 $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ ，其中 ε_i 為 $i.i.d.N(0, \sigma^2)$ 。已知 $F_{1,18,0.05} = 3.01$, $\bar{x} = 6$, $\bar{y} = 15$, $\sum(x_i - \bar{x})^2 = 25$, $\sum(y_i - \bar{y})^2 = 208$, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 40$ 。

(一)試寫出此模型之變異數分析 (Analysis of Variance) 表。(10 分)

(二)試求此一迴歸線之斜率與截距。(10 分)

(三)試求此一迴歸線斜率 β_1 之 90% 信賴區間。(5 分)

二、若 (x_{1i}, x_{2i}, y_i) , $i = 1, \dots, n$ 彼此獨立且來自截距項為零的線性迴歸模型 $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$ ，其中 ε_i 為 $i.i.d.N(0, \sigma^2)$ 。

(一)試求 β_1 與 β_2 之最小平方估計量 $\hat{\beta}_1$ 與 $\hat{\beta}_2$ 。(15 分)(二)試求 $E(\hat{\beta}_1)$ 與 $\text{Var}(\hat{\beta}_1)$ 。(10 分)

三、某一研究想要知道房屋的價格 Y 與房屋的坪數 X_1 ，屋齡 X_2 ，房間數 X_3 與空屋率 X_4 的關係。今收集 30 間房屋的資料並對此資料配適一迴歸模型 $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \varepsilon_i$ ，其中 ε_i 為 $i.i.d.N(0, \sigma^2)$ 。已知 $F_{2,27,0.05} = 3.35$, $F_{2,25,0.05} = 3.39$ 。

Source of Variation (變異來源)	SS	df	MS
$SSR(X_1, X_2, X_3, X_4)$	300	4	75
$SSR(X_3, X_4 X_1, X_2)$	40	2	20
$SSR(X_3, X_4)$	150	2	75
$SSE(X_1, X_2, X_3, X_4)$	200	25	8

(一)假設迴歸模型中僅考慮房間數 X_3 與空屋率 X_4 。試就此模型在 $\alpha = 0.05$ 下檢定 $H_0: \beta_3 = \beta_4 = 0$ 。(請務必將完整之檢定寫出，包括 H_0, H_1 ，檢定量，拒絕區域，結論等) (10 分)

(二)假設迴歸模型中已考慮坪數 X_1 與屋齡 X_2 。試就此模型在 $\alpha = 0.05$ 下檢定 $H_0: \beta_3 = \beta_4 = 0$ 。(請務必將完整之檢定寫出，包括 H_0, H_1 ，檢定量，拒絕區域，結論等) (10 分)

(三)試解釋上面兩小題結果不盡相同之原因。(5 分)

四、在模型診斷時，我們常用 DFFITS, Cook's Distance, DFBETAS 方法辨認具有影響力的個案 (Influential cases)。

(一)試比較 DFFITS, Cook's Distance, DFBETAS 此三種方法之差異。(15 分)

(二)試說明此三種方法辨認具有影響力的個案之判定原則。(10 分)