

98 年公務人員高等考試三級考試試題

代號：31750 全一張
(正面)

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、若 $(y_i, x_i), i = 1, 2, \dots, n$ ，彼此獨立並滿足線性迴歸模型 $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ ，其中截距項 β_0 是已知的， ε_i 為 i.i.d. $N(0, \sigma^2)$ 。

(一)試求 β_1 之最小平方估計量 (least squares estimator) $\hat{\beta}_1$ 。(10 分)

(二)試求 $E(\hat{\beta}_1)$ 和 $Var(\hat{\beta}_1)$ 。(10 分)

(三)試求當另一獨立解釋變數之觀測值 $x = x_0$ 時，其反應變數 y_0 之 $(1-\alpha)100\%$ 的預測區間。(10 分)

二、假設樣本資料 $(y_i, x_i), i = 1, 2, \dots, n$ ，配適簡單線性迴歸模型 $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ ，其中 ε_i 為 i.i.d. $N(0, \sigma^2)$ 。若樣本數 $n = 20$ ， $\bar{x} = \sum_{i=1}^n x_i / n = 5$ ， $\bar{y} = 3$ ， $\sum_{i=1}^n (x_i - \bar{x})^2 = 160$ ， $\sum_{i=1}^n (y_i - \bar{y})^2 = 83.2$ 且 $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 80$ 。

(一)試寫出模型之 ANOVA (analysis of variance) 表。(10 分)

(二)若解釋變數 $x_i = 2, i = 1, 2, 3, 4; x_i = 4, i = 5, \dots, 8; x_i = 6, i = 9, \dots, 12; x_i = 8, i = 13, \dots, 16$ 及 $x_i = 10, i = 17, \dots, 20$ ，且其純誤差平方和 (pure error sum of squares) 為 23.2。試問此時(一)中的模型是否仍恰當？請寫出檢定統計量之分布和自由度。(臨界值 (critical value) = 3.29。)(15 分)

三、若 $(y_i, x_{i1}, x_{i2}, x_{i3}), i = 1, 2, \dots, 20$ ，滿足迴歸模型

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i,$$

其中 ε_i 為 i.i.d. $N(0, \sigma^2)$ 。且其 SS_{Reg} (sum of squares for regression) = 19.0064, $SS_{Res} = 11.3792$, $R(\beta_1|\beta_0) = 10.2070$ 和 $R(\beta_2|\beta_0, \beta_1) = 6.2856$ 。試求：

(一) R^2 值。(5 分)

(二)檢定上述迴歸模型是否顯著？(請寫出檢定統計量之分布和自由度。臨界值 = 3.24。)(10 分)

(三)若檢定 $\beta_3 = 0$ 不顯著，試求此時 σ^2 之估計量。(5 分)

(請接背面)

類 科：統計
科 目：迴歸分析

四、下表為一可能有 4 個解釋變數 x_1, x_2, x_3 和 x_4 之資料，分別配適具左方之解釋變數之迴歸模型時其餘變數之 partial F 值。(假設模型誤差為獨立同分佈的 $N(0, \sigma^2)$ 。)

模型中之 解釋變數	不在模型中之解釋變數的 partial F 值			
	x_1	x_2	x_3	x_4
—	12.60	21.96	4.40	22.80
x_1	—	208.58	0.31	159.30
x_2	146.52	—	11.82	0.42
x_3	5.81	36.68	—	100.36
x_4	108.22	0.17	40.29	—
x_1x_2	—	—	1.83	1.86
x_1x_3	—	220.55	—	208.24
x_1x_4	—	5.03	4.24	—
x_2x_3	68.72	—	—	41.65
x_2x_4	154.01	—	96.94	—
x_3x_4	22.11	12.43	—	—
$x_1x_2x_3$	—	—	—	0.04
$x_1x_2x_4$	—	—	0.02	—
$x_1x_3x_4$	—	0.5	—	—
$x_2x_3x_4$	4.34	—	—	—

考慮以 $F_{IN} = F_{OUT} = 4$ 試分別寫出下列方法所選出的解釋變數：

- (一) 逐步迴歸法 (Stepwise Regression)。(5 分)
- (二) 向前選擇法 (Forward Selection)。(5 分)
- (三) 後退消去法 (Backward Elimination)。(5 分)
- (四) 在下述各模型與所提供的訊息中選擇最佳的模型。(10 分)

解釋變數	R^2	R_{adj}^2 (adjusted R^2)	C_p	s
x_2	.666	.636	142.5	9.07
x_4	.675	.645	138.7	8.96
x_1x_2	.979	.974	2.7	2.41
x_1x_4	.972	.967	5.5	2.73
$x_1x_2x_4$.982	.976	3.0	2.31
$x_1x_3x_4$.981	.975	3.5	2.37