

111年公務人員高等考試三級考試試題

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。

(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、下表為隨機抽取全臺 30 處茶園所生產的某一品種之茶葉資料，包括該茶園所在的區域（標示為 1、2、3）、茶葉每 10 公克價格（單位：新臺幣元）及由茶葉的特性綜合計算而得之兩種因素（甲、乙）分數。

茶園	因素甲 (X_1)	因素乙 (X_2)	區域 (X_3)	價格 (Y)
1	3.1	4.1	1	9.8
2	3.5	3.9	1	12.6
3	4.8	4.7	1	11.9
4	3.1	3.6	1	11.1
5	5.5	5.1	1	9.3
6	5.0	4.1	1	12.8
7	4.8	3.3	1	12.8
8	4.3	5.2	1	12.0
9	3.9	2.9	3	13.6
10	4.7	3.9	1	13.9
11	4.5	3.6	3	14.4
12	4.3	3.6	2	12.3
13	7.0	4.1	3	16.1
14	6.7	3.7	3	16.1
15	5.8	4.1	3	15.5
16	5.6	4.4	3	15.5
17	4.8	4.6	3	13.8
18	5.5	4.1	3	13.8
19	4.3	3.1	1	11.3
20	3.4	3.4	2	10.9
21	6.6	4.8	3	15.1
22	5.3	3.8	3	13.5
23	5.0	3.7	2	10.8
24	4.1	4.0	2	9.5
25	5.7	4.7	3	12.7
26	4.7	4.9	2	11.6
27	5.1	5.1	1	11.7
28	5.0	5.1	2	11.9
29	5.0	4.4	2	10.8
30	2.9	3.9	2	8.5
總和	144.00	123.90	60	375.60
總平方和	722.02	523.03	142	4821.30

「總和」與「總平方和」表示該變數之值的加總及取平方後之加總。

定義區域 X_3 之指標變數如下：

$$D_1 = \begin{cases} 1 & \text{若該茶園屬於區域 1} \\ 0 & \text{其他} \end{cases}, D_2 = \begin{cases} 1 & \text{若該茶園屬於區域 2} \\ 0 & \text{其他} \end{cases}, D_3 = \begin{cases} 1 & \text{若該茶園屬於區域 3} \\ 0 & \text{其他} \end{cases}$$

以茶葉每 10 公克之價格為反應變數，考慮迴歸模型如下：

$$Y = \beta_0 + \beta_1 X_3 + \varepsilon \quad \text{模型 A}$$

$$Y = \beta_0 + \beta_1 D_2 + \beta_2 D_3 + \varepsilon \quad \text{模型 B}$$

$$Y = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \varepsilon \quad \text{模型 C}$$

$$Y = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \varepsilon \quad \text{模型 D}$$

$$Y = \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \varepsilon \quad \text{模型 E}$$

ε 為隨機誤差項。

(一)說明模型 A、B、C、D 及 E 所表達的意義，並決定那些模型是較合適的。
(10 分)

(二)依據模型 B，得到以下估計結果：

	Estimate	Std Err
Intercept	11.7455	0.3789
D_2	-0.9580	0.5840
D_3	2.8091	0.5359

$$R^2 = 0.6410$$

試詳細說明係數 β_2 之估計值的意義，並計算此模型之迴歸標準誤的估計。(8 分)

(三)試寫出模型 C、D 及 E 之所有迴歸係數的估計值。(12 分)

二、承上題，令模型 F 為 $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 D_2 + \beta_4 D_3 + \varepsilon$ 。

依據模型 F，得到以下估計結果：

	Estimate	Std Err
Intercept	10.6039	1.4918
X_1	0.8881	0.2752
X_2	-0.6561	0.3686
D_2	-0.9285	0.5097
D_3	1.6658	0.5853

$$R^2 = 0.7473$$

(一)若有一茶園位於區域 1，且其因素甲與因素乙的分數皆為 4.5。該品種之茶葉包裝方式為 500 公克一盒，請估計一盒的售價。(4 分)

(二)若將反應變數的單位改為茶葉每公斤之價格，寫出此迴歸模型之係數估計結果及其標準誤（如表格中之 Estimate 與 Std Err）。(8 分)

(三)在顯著水準為 0.05 下，檢定 β_1 與 β_2 是否同時等於 0。(8 分)

三、若迴歸分析得到以下變異數分析（Analysis of variance, ANOVA）表：

Source of variation	Degrees of freedom	Sum of squares	Mean square	<i>F</i>
Regression	(1)	(6)	(8)	(13)
X_1	(2)	33.2	(9)	(14)
$X_2 X_1$	(3)	15.3	(10)	(15)
$X_3 X_1, X_2$	(4)	22.2	(11)	(16)
Error	(5)	(7)	(12)	
Total	29	106.3		

(一)請寫出 ANOVA 表中 (1) 至 (16) 的值。(16 分)

(二)在顯著水準為 0.05 下，詳述 ANOVA 表中 (13) 至 (16) 所提供之檢定的意義與結果。(14 分)

四、考慮 p 個解釋變數 ($X_1, X_2, \dots, X_p$)， n 個觀測值的複迴歸模型如下：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i, \quad i=1, 2, \dots, n$$

(一)試以上述模型詳述共線性的意義，當共線性現象發生時所可能導致之影響。(10 分)

(二)詳述變異膨脹因子 (variance inflation factor, VIF) 之定義及其判斷準則。(10 分)

【附表】F 分配表

F - Distribution ($\alpha = 0.05$ in the Right Tail)

df ₂	df ₁	Numerator Degrees of Freedom								
		1	2	3	4	5	6	7	8	9
Denominator Degrees of Freedom	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
	2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
	3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
	4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
	5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
	6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
	7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
	8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
	9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
	10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
	11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
	12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
	13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
	14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
	15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
	16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
	17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
	18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
	19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
	20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
	21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
	22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
	23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
	24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
	25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
	26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
	27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
	28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
	29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
	30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
	40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
	60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
	120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
	∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799