

5 實例分析

本章中將使用 Tanaka 在 1987 所提出的資料(如表格 1) 來驗證我們在前幾章所獲致的理論結果。資料中有三個實數自變數，一個模糊應變數及十筆觀察值。在此僅考慮應變數值為指數型模糊數。採用指數型隸屬度函數的優點是，只要適當選取 m 的值就可以充分表現應變數資料分布的特性。若應變數的值，落在所蒐集到的資料區間以外的可能性較大時，我們可選取較小的 m 值作為描述應變數資料的隸屬度函數，反之，若應變數的值落在資料區間以外的可能性較小，則選取較大的 m 值作為描述資料的隸屬度函數。

由於我們無法求得模型 III 最小平方估計的公式解，並且在本章中僅考慮指數型隸屬度函數，因此僅針對模型 I 與模型 II 進行資料分析。表格 2~ 表格 11 是在模型 I 與模型 II 的架構下，分別採用簡單距離公式與 Yang 和 Ko 的距離公式，以及選取不同 m 值所獲得的結果。各表中，除了包含參數的最小平方估計與殘差平方和外、還包含了偵測自變數部份離群值的槓桿值 h_{ii} 、偵測應變數部份離群值的殘差平方 $e_i^2, e_{(i)}^2$ 以及偵測具影響力觀察值的 Cook 距離 CD_i 值。在簡單距離公式下，無論模型 I 與模型 II 表格的內容都不會受 m 值改變的影響，因此僅列出 $m = 2$ 的情況(如表格 2 與表格 3)。

表格 1: Tanaka(1987) 資料

觀察項	自變數			模糊應變數
i	x_{i1}	x_{i2}	x_{i3}	$Y_i = (c_i, r_i)$
1	3	5	9	(96,42)
2	14	8	3	(120,47)
3	7	1	4	(52,33)
4	11	7	3	(106,45)
5	7	12	15	(189,79)
6	8	15	10	(194,65)
7	3	9	6	(107,42)
8	12	15	11	(216,78)
9	10	5	8	(108,52)
10	9	7	4	(103,44)

表格 2: 模型 I, $m = 2$, 簡單距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(93.20,44.62)	0.40	14.69	41.25	0.25
2	(120,47)	(122.48,49.13)	0.43	10.67	32.44	0.21
3	(52,33)	(49.36,32.11)	0.41	7.75	21.90	0.13
4	(106,45)	(104.82,43.01)	0.26	5.35	9.75	0.04
5	(189,79)	(191.79,76.71)	0.55*	13.06	63.57*	0.52*
6	(194,65)	(193.64,67.67)	0.39	7.25	19.38	0.11
7	(107,42)	(109.77,40.85)	0.60*	9.08	55.55*	0.50*
8	(216,78)	(211.65,77.08)	0.42	19.73	58.34*	0.37
9	(108,52)	(110.89,53.24)	0.37	9.91	25.12	0.14
10	(103,44)	(103.36,42.58)	0.18	2.14	3.22	0.01

$\hat{a} = (-1.39, 3.25, 7.92, 5.03)'$, $\hat{r} = (8.01, 1.64, 1.20, 2.85)'$, $\sum e_i^2 = 99.63$

表格 3: 模型 II, $m = 2$, 簡單距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(93.86,42.38)	0.40	4.71	11.08	0.36
2	(120,47)	(122.04,50.63)	0.43	17.34	31.44	0.38
3	(52,33)	(50.11,29.56)	0.41	15.41	92.87*	0.66*
4	(106,45)	(104.13,45.38)	0.26	3.65	6.55	0.35
5	(189,79)	(193.31,71.51)	0.55*	74.67*	90.38*	0.38
6	(194,65)	(192.58,71.30)	0.39	41.67	118.34*	0.63*
7	(107,42)	(108.12,46.55)	0.60*	21.97	38.00	0.39
8	(216,78)	(211.71,76.90)	0.42	19.64	78.55*	0.61*
9	(108,52)	(112.48,47.86)	0.37	37.44	82.31*	0.51
10	(103,44)	(102.67,44.96)	0.18	1.02	5.34	0.34

$\hat{a} = (-3.14, 3.43, 7.62, 5.40)'$, $\hat{b} = 0.29$, $\hat{d} = 14.88$, $\sum e_i^2 = 236.98$

表格 4: 模型 I, $m = 1.2$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(93.20,44.62)	0.40	35.61	100.02	0.24
2	(120,47)	(122.48,49.13)	0.43	26.43	80.41	0.20
3	(52,33)	(49.36,32.11)	0.41	22.26	62.94	0.15
4	(106,45)	(104.82,43.01)	0.26	11.17	31.59	0.03
5	(189,79)	(191.79,76.71)	0.55*	32.71	159.18*	0.51*
6	(194,65)	(193.64,67.67)	0.39	12.99	34.72	0.08
7	(107,42)	(109.77,40.85)	0.60*	25.61	156.77*	0.55*
8	(216,78)	(211.65,77.08)	0.42	58.16	171.95*	0.42
9	(108,52)	(110.89,53.24)	0.37	27.84	70.53	0.16
10	(103,44)	(103.36,42.58)	0.18	3.95	5.94	0.01

$\hat{a} = (-1.39, 3.25, 7.92, 5.03)'$, $\hat{r} = (8.01, 1.64, 1.20, 2.85)'$, $\sum e_i^2 = 256.73$

表格 5: 模型 II, $m = 1.2$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(94.91,37.94)	0.40	32.79	60.26	0.02
2	(120,47)	(122.32,49.77)	0.43	29.66	80.87	0.04
3	(52,33)	(52.54,19.65)	0.41	316.42*	360.47*	0.08
4	(106,45)	(105.01,42.30)	0.26	15.89	24.84	0.01
5	(189,79)	(191.09,79.46)	0.55*	13.48	63.15	0.04
6	(194,65)	(190.68,79.29)	0.39	394.20*	536.55*	0.07
7	(107,42)	(108.98,44.01)	0.60*	18.96	104.09	0.07
8	(216,78)	(209.07,87.22)	0.42	294.63*	498.27*	0.13*
9	(108,52)	(112.82,45.67)	0.37	140.58	207.55	0.05
10	(103,44)	(103.52,41.69)	0.18	10.51	13.72	0.002

$\hat{d} = -3.03$, $\hat{b} = 0.43$, $\hat{a} = (1.28, 3.30, 7.41, 5.19)'$, $\sum e_i^2 = 1267.13$

表格 6: 模型 I, $m = 2$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(93.20,44.62)	0.40	34.24	96.18	0.23
2	(120,47)	(122.48,49.13)	0.43	25.53	77.68	0.20
3	(52,33)	(49.36,32.11)	0.41	22.10	62.49	0.15
4	(106,45)	(104.82,43.01)	0.26	10.38	18.91	0.03
5	(189,79)	(191.79,76.71)	0.55*	31.66	154.08*	0.51*
6	(194,65)	(193.64,67.67)	0.39	11.57	30.93	0.07
7	(107,42)	(109.77,40.85)	0.60*	25.35	155.17*	0.55*
8	(216,78)	(211.65,77.08)	0.42	57.99	171.46*	0.43
9	(108,52)	(110.89,53.24)	0.37	27.53	69.75	0.16
10	(103,44)	(103.36,42.58)	0.18	3.55	5.34	0.01

$\hat{a} = (-1.39, 3.25, 7.92, 5.03)'$, $\hat{r} = (8.01, 1.64, 1.20, 2.85)'$, $\sum e_i^2 = 249.91$

表格 7: 模型 II, $m = 2$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(94.76,37.78)	0.40	32.73	61.86	0.02
2	(120,47)	(122.33,49.76)	0.43	28.30	77.95	0.04
3	(52,33)	(52.28,19.29)	0.41	295.52*	336.34*	0.09
4	(106,45)	(105.00,42.22)	0.26	15.12	23.74	0.01
5	(189,79)	(191.12,79.67)	0.55*	14.15	67.62	0.06
6	(194,65)	(190.93,79.59)	0.39	362.74*	493.27*	0.07
7	(107,42)	(109.07,43.99)	0.60*	19.09	107.17	0.08
8	(216,78)	(209.27,87.57)	0.42	279.81*	476.79*	0.13*
9	(108,52)	(112.65,45.54)	0.37	130.24	194.90	0.05
10	(103,44)	(103.59,41.60)	0.18	10.06	13.27	0.002

$\hat{d} = -3.45$, $\hat{b} = 0.43$, $\hat{a} = (1.11, 3.29, 7.45, 5.17)'$, $\sum e_i^2 = 1187.75$

表格 8: 模型 I, $m = 3$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(93.20,44.62)	0.40	34.41	96.65	0.23
2	(120,47)	(122.48,49.13)	0.43	25.64	78.01	0.20
3	(52,33)	(49.36,32.11)	0.41	22.12	62.54	0.15
4	(106,45)	(104.82,43.01)	0.26	10.48	19.08	0.03
5	(189,79)	(191.79,76.71)	0.55*	31.79	154.70*	0.51*
6	(194,65)	(193.64,67.67)	0.39	11.74	31.39	0.07
7	(107,42)	(109.77,40.85)	0.60*	25.38	155.36*	0.55*
8	(216,78)	(211.65,77.08)	0.42	58.01	171.52*	0.43
9	(108,52)	(110.89,53.24)	0.37	27.57	69.85	0.16
10	(103,44)	(103.36,42.58)	0.18	3.60	5.41	0.01

$\hat{a} = (-1.39, 3.25, 7.92, 5.03)'$, $\hat{r} = (8.01, 1.64, 1.20, 2.85)'$, $\sum e_i^2 = 250.74$

表格 9: 模型 II, $m = 3$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(94.78,37.79)	0.40	32.73	61.65	0.02
2	(120,47)	(122.33,49.76)	0.43	28.46	78.30	0.04
3	(52,33)	(52.31,19.34)	0.41	298.04*	339.26*	0.08
4	(106,45)	(105.00,42.23)	0.26	15.21	23.87	0.01
5	(189,79)	(191.11,79.64)	0.55*	14.06	67.06	0.05
6	(194,65)	(190.90,79.55)	0.39	366.55*	498.51*	0.07
7	(107,42)	(109.06,43.99)	0.60*	19.08	102.80	0.08
8	(216,78)	(209.24,87.52)	0.42	281.59*	479.11*	0.13*
9	(108,52)	(112.67,45.56)	0.37	131.49	196.43	0.05
10	(103,44)	(103.59,41.61)	0.18	10.11	13.32	0.002

$\hat{d} = -3.39$, $\hat{b} = 0.43$, $\hat{a} = (1.13, 3.29, 7.44, 5.17)'$, $\sum e_i^2 = 1197.32$

表格 10: 模型 I, $m = 10$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(93.20,44.62)	0.40	35.88	100.80	0.24
2	(120,47)	(122.48,49.13)	0.43	26.62	80.97	0.20
3	(52,33)	(49.36,32.11)	0.41	22.29	63.03	0.15
4	(106,45)	(104.82,43.01)	0.26	11.33	20.64	0.03
5	(189,79)	(191.79,76.71)	0.55*	32.92	160.21*	0.51*
6	(194,65)	(193.64,67.67)	0.39	13.28	35.49	0.08
7	(107,42)	(109.77,40.85)	0.60*	25.67	157.10*	0.54*
8	(216,78)	(211.65,77.08)	0.42	58.19	172.06*	0.42
9	(108,52)	(110.89,53.24)	0.37	27.90	70.69	0.15
10	(103,44)	(103.36,42.58)	0.18	4.02	6.06	0.01

$\hat{a} = (-1.39, 3.25, 7.92, 5.03)'$, $\hat{r} = (8.01, 1.64, 1.20, 2.85)'$, $\sum e_i^2 = 258.10$

表格 11: 模型 II, $m = 10$, Yang 和 Ko 距離公式下最小平方法結果

i	(c_i, s_i)	(\hat{c}_i, \hat{s}_i)	h_{ii}	e_i^2	$e_{(i)}^2$	CD_i
1	(96,42)	(94.93,37.97)	0.40	32.80	59.96	0.01
2	(120,47)	(122.31,49.77)	0.43	29.93	81.46	0.04
3	(52,33)	(52.59,19.71)	0.41	320.68*	365.36*	0.08
4	(106,45)	(105.01,42.31)	0.26	16.05	25.07	0.01
5	(189,79)	(191.09,79.43)	0.55*	13.36	62.29	0.04
6	(194,65)	(190.63,79.23)	0.39	400.57*	545.32*	0.07
7	(107,42)	(108.96,44.02)	0.60*	18.93	103.48	0.07
8	(216,78)	(209.03,87.16)	0.42	297.67*	502.71*	0.13*
9	(108,52)	(112.86,45.70)	0.37	142.67	210.11	0.05
10	(103,44)	(103.59,41.70)	0.18	10.61	13.82	0.002
$\hat{d} = -2.96, \hat{b} = 0.43, \hat{a} = (1.31, 3.30, 7.40, 5.19)'$, $\sum e_i^2 = 1283.29$						

由表格 2 與表格 3 可看出，根據模型 I 所獲得中心與分展度的預測值較模型 II 為佳。理論上，在採用 Yang 和 Ko 距離公式時模型 II 的中心與分展度預測值會受到 m 值的影響，但是由表格 5,7,9,11 卻發現， m 值改變對中心與分展度預測值影響不大。

理論上模型 I 的參數估計，不受距離公式與 m 值改變的影響，而模型 II 的參數估計是會受到影響的。由表格 3 與表格 5 可以看出，距離公式的不同會對模型 II 的參數估計造成較大的影響，又由表格 5 與表格 7 可看出， m 值的不同對模型 II 的參數估計影響較小。

由於第 5 筆與第 7 筆觀察值的槓桿值 h_{ii} 較大，因此有可能是自變數部份的離群值。對模型 I 而言，從殘差 e_i 來看，沒有特別突出的應變數部份離群值，但由表格 2,4,6,8 來看 $e_{(i)}$ 值，第 5,7,8 三筆資料為應變數部份離群值。對模型 II 而言，在簡單距離公式下，表格 3 顯示出可能的應變數離群值有第 3,5,6,8,9 這五筆資料；在 Yang 和 Ko 的距離公式下，表格 5,7,9,11 顯示出可能的應變數離群值僅有第 3,6,8 三筆資料。

對模型I而言，從表格2,4,6,8,10可以看出，只有第5,7筆的 CD_i 值較大，因此是具影響力的觀察值。對模型II而言，在簡單距離公式下，由表格3可以看出第3,6,8三筆資料的 CD_i 值較大，是具影響力觀察值；在Yang和Ko的距離公式下，由表格5,7,9,11可以看出僅第8筆資料的 CD_i 值較大，為具影響力觀察值。

在Yang和Ko的距離公式下，對模型II做模糊線性迴歸，若採用指數型隸屬度函數時，我們應如何選取 m 的值？一個可能簡單的法則是，選取使殘差平方和最小的 m 值。在本例中，由表格5,7,9,11可以看出，當 $m = 2$ 時殘差平方和是最小的，故常態隸屬度函數似乎是我們的最佳選擇。