二十九、區別分析、判別分析
Discriminant Analysis

1. 區別分析是多元統計分析中用於判別樣品(受訪者)所屬類型(族群)的一種方法,與集群分析相同的勢將相似的樣本(受訪者)歸爲一類(族群),不同處卻在於集群分析預先不知道分類,而區別分析是在研究對象分類已知的情況下,根據樣本資料推導出一個或一組區別(判別)函數,同時指定一種判別規則,用於確定待判別樣本的所屬類別,使錯判率最小。

2. The basic purpose of discriminant analysis is to estimate the relationship between a single nonmetric (categorical) dependent variable and a set of metric independent variables.

3. Discriminant loadings: In recent years, loadings have increasingly been used as a basis for interpretation because of the deficiencies in utilizing weights. Discriminant loadings, referred to sometimes as structure correlations, measure the simple linear correlation between each independent variable and the discriminant function. The discriminant loadings reflect the variance that the independent variables share with the discriminant function and can be interpreted like factor loadings in assessing the relative contribution of each independent variable to the discriminant function.

4. After the discriminant function has been computed, the researcher must assess its level of significance. A number of different statistical criteria are available. The measures of Wilks' lambda, Hotelling's trace, and Pillai's criterion all evaluate the statistical significance of the discriminatory power of the discriminant functions(s).

5. 主要目的是運用於計算一組預測變數(自變數)包括知識、價值、態度、環保行爲的線性組合,對依變數(間斷變數)接受有機農產品更高售價之意願加以分類,並檢定其再分組的正確率。利用逐步區別分析針對願意願意與不願意支付更高售價購買有機農產品的兩個族群,以知識、價值、態度、環保行爲等變數進行區別化。進行逐步區別分析前數值先經過標準化(normality)和變異數同質性(homogeneity of variance/covariance matrices)程序之確認;在進行逐步區別分析前先將有效樣本抽取70%做爲係數估算分析使用,另外的30%有效問卷做爲結果的驗證。

分析方法

	自變數 X_1	因變數(依變數) Y_1
簡單迴歸分析	單一個自變數[預測變數(predictor variable)] Interval scale or ratio scale[連續變數(項)] Nominal scale or ordinal scale(轉化爲虛擬變數)	單一個依變數[效標變數(criterion variable)] Interval scale or ratio scale[連續變數(項)]
複迴歸分析 (logistic regression)	兩個(含)以上自變數[預測變數(predictor variable)] Interval scale or ratio scale[連續變數(項)] Nominal scale or ordinal scale(轉化爲虛擬變數)	單一個依變數[效標變數(criterion variable)] Interval scale or ratio scale[連續變數(項)]
邏輯迴歸分析	自變數[預測變數(predictor	單一個依變數

	自變數 X_1	因變數(依變數) Y_1
	variable)]數量不限 Interval scale or ratio scale[連續變數(項)] Nominal scale or ordinal scale(轉化為虛擬變數)	兩項式(二分)Nominal scale
區別分析、判別分析(discriminant analysis)	自變數 [預測變數 (predictor variable)]數量不限 Interval scale or ratio scale[連續變數(項)] Nominal scale or ordinal scale(轉化為虛擬變數)	單一個依變數(分組變數) 三項式以上 (三分以上)Nominal scale or ordinal scale

MANOVA 與 discriminant analysis 的差異：MANOVA 期望瞭解各集群(組)樣本在哪幾個依變數的平均值達到顯著水準。Discriminant analysis 透過得到自變數之線性組合方成函數，瞭解自變數(觀測值)在依變數(集群、組數)上分類的正確性，進而獲悉哪些自變數(預測變數)可以有效區分類別。

SPSS 操作方式

A.Analyze(分析) → Classify(分類) → Discriminant…(判別…)

B.在 Discriminant analysis 對話視窗中，將左邊小視窗中的類別變數選入右邊的 Grouping variable:中，按 Define Range…(定義範圍…)按鈕，會出現 Discriminant analysis: Define Range 對話小視窗，分別輸入類別變數的 Minimum:和 Maximum:數值按 Continue 回到 Discriminant analysis 對話視窗；左邊小視窗中的自變數選入右邊的 Independents:小視窗中。

C.在 Discriminant Analysis 視窗中勾選下面的 Classify… 按鈕，則會出現 Discriminant Analysis: Classification 次對話視窗中 Display 勾選 Summary table 選項，按 continue 按鈕回到 Discriminant Analysis 視窗。

D.在 Discriminant Analysis 視窗中勾選右側 OK 按鈕，即執行 Discriminant analysis。

結果
Discriminant

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		291	58.3
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	208	41.7
	Total	208	41.7
Total		499	100.0

Group Statistics

Cluster Number of Case			Mean	Std. Deviation	Valid N (listwise) Unweighted	Weighted
1	REGR factor score	1 for analysis 1	-1.0974	0.7679	63	63.000
	REGR factor score	2 for analysis 1	-0.6517	0.7107	63	63.000
	REGR factor score	3 for analysis 1	-0.4437	0.8218	63	63.000
	REGR factor score	4 for analysis 1	-0.4588	0.9881	63	63.000
2	REGR factor score	1 for analysis 1	-0.4627	0.7863	59	59.000
	REGR factor score	2 for analysis 1	0.5408	0.9833	59	59.000
	REGR factor score	3 for analysis 1	-0.5075	0.5897	59	59.000
	REGR factor score	4 for analysis 1	0.7173	0.8982	59	59.000
3	REGR factor score	1 for analysis 1	0.9560	0.5371	72	72.000
	REGR factor score	2 for analysis 1	-0.3140	1.0082	72	72.000
	REGR factor score	3 for analysis 1	-0.5681	0.7256	72	72.000
	REGR factor score	4 for analysis 1	-0.5122	0.8734	72	72.000
4	REGR factor score	1 for analysis 1	0.2846	0.6033	97	97.000
	REGR factor score	2 for analysis 1	0.3274	0.8473	97	97.000
	REGR factor score	3 for analysis 1	1.0186	0.6383	97	97.000
	REGR factor score	4 for analysis 1	0.2418	0.7928	97	97.000
Total	REGR factor score	1 for analysis 1	0.0000	1.0000	291	291.000
	REGR factor score	2 for analysis 1	0.0000	1.0000	291	291.000
	REGR factor score	3 for analysis 1	0.0000	1.0000	291	291.000
	REGR factor score	4 for analysis 1	0.0000	1.0000	291	291.000

Tests of Equality of Group Means

		Wilks' Lambda	F	df1	df2	Sig.
REGR factor score	1 for analysis 1	.441	121.357	3	287	.000
REGR factor score	2 for analysis 1	.788	25.757	3	287	.000
REGR factor score	3 for analysis 1	.478	104.602	3	287	.000
REGR factor score	4 for analysis 1	.765	29.405	3	287	.000

說明：各集群(組)平均值相等性檢定，即在各預測變數的在各集群(組)中之平均值差異驗證。當 F 值愈大，Wilks' Λ 值愈小，其平均值的差異性即愈大。上表結果顯示 $p < 0.05$，代表其差異性達到顯著性水準。

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Cluster Number of Case	Rank	Log Determinant
1	4	-2.133
2	4	-2.513
3	4	-2.390
4	4	-3.059
Pooled within-groups	4	-2.366

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M		62.222
F	Approx.	2.020
	df1	30
	df2	180863.475
	Sig.	.001

Tests null hypothesis of equal population covariance matrices.

說明：各集群(組)內共變異數矩陣(covariance matrices)相等性檢定。若p(Sig.) > 0.05時，表示接受虛無假設，各集群(組)內共變異數矩陣相等，符合區別分析的假設，可以使用聯合組內共變異數矩陣進行分析。若p(Sig.) < 0.05時，表示接受對立假設，各集群(組)內共變異數矩陣不相等，使用個別集群(組)內共變異數矩陣進行分析。

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.861(a)	65.8	65.8	.861
2	1.149(a)	26.4	92.2	.731
3	.339(a)	7.8	100.0	.503

a   First 3 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	.090	688.617	12	.000
2 through 3	.348	302.262	6	.000
3	.747	83.412	2	.000

說明：向度縮減分析，即驗證區別函數的顯著性。

Standardized Canonical Discriminant Function Coefficients

		Function		
		1	2	3
REGR factor score	1 for analysis 1	.652	-.797	.187
REGR factor score	2 for analysis 1	.683	.340	.508
REGR factor score	3 for analysis 1	.975	.256	-.518
REGR factor score	4 for analysis 1	.488	.569	.533

說明：標準化典型區別函數係數，代表各自變數(預測變數)在各區別函數(Functions)中的相對重要性(貢獻度)，係數值愈高，代表該自變數(預測變數)在此區別函數的重要性(貢獻度)愈高。

Structure Matrix

		Function		
		1	2	3
REGR factor score	1 for analysis 1	.383	-.842(*)	.317
REGR factor score	3 for analysis 1	.529	.249	-.810(*)
REGR factor score	4 for analysis 1	.165	.346	.520(*)

	Function		
	1	2	3
REGR factor score 2 for analysis 1	.225	.201	.481(*)

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions   Variables ordered by absolute size of correlation within function.

* Largest absolute correlation between each variable and any discriminant function

Functions at Group Centroids

Cluster Number of Case	Function		
	1	2	3
1	-2.449	.603	-.626
2	-.352	1.036	.991
3	-.388	-1.803	.202
4	2.093	.316	-.346

Unstandardized canonical discriminant functions evaluated at group means

說明：各集群(組)在各區別函數中的形心(centroid)。形心數值可由未標準化區別函數係
數×各集群(組)平均值獲得。在同一個區別函數(Function)中，形心數值差異愈大時，
表示該兩個集群(組)在區別函數(Function)的差異愈大。

Classification Results(a)

		Cluster Number of Case	Predicted Group Membership				
			1	2	3	4	Total
Original	Count	1	56	1	6	0	63
		2	2	54	0	3	59
		3	0	1	69	2	72
		4	0	1	1	95	97
	%	1	88.9	1.6	9.5	.0	100.0
		2	3.4	91.5	.0	5.1	100.0
		3	.0	1.4	95.8	2.8	100.0
		4	.0	1.0	1.0	97.9	100.0

a   94.2% of original grouped cases correctly classified.

表 6 逐步區別分析特徵函數

Eigenvalue	Canonical correlation	Wilks' Lambda	Chi square	Grouped cases correctly classified[a](%)
1.779	0.800	0.360	$\chi^2(6)=196.2$ $P < 0.001$	Willing respondent: 85.6 Unwilling respondent: 95.0 Total=90.8

Note: [a] Classification was based on a validation sample (30 % of the original sample)

表 7 t-檢定和逐步區別分析

因素構面[a]	t-檢定			區別分析	
	平均值		t 值	Wilks' Lambda	Loadings(L)
	願意多付	不願意多付			
有機優勢	-0.77	1.01	-15.83[c]	0.451[c]	0.88
自我肯定	-0.76	0.86	-12.70[c]	0.372[c]	0.69
永續生態	0.59	-0.81	-10.99[c]		-0.67
回收效益	-0.58	0.76	-10.29[c]		0.55

因素構面 [a]	t-檢定			區別分析	
	平均值		t 值	Wilks' Lambda	Loadings(L)
	願意多付	不願意多付			
資源回收	-0.50	0.84	-10.26[c]	0.372[c]	0.55
有機知識	4.78	5.97	-7.29[c]	0.368[c]	0.40
環保障礙	0.36	-0.51	6.73[c]		-0.33
購買危害	0.00	-0.41	2.85[b]	0.372[c]	-0.19
環境責任	0.13	-0.36	3.32[b]	0.373[c]	-0.18

Notes: [a] Numbers represent the factors' order of entry in the discriminant analysis($\alpha < 0.05$)

[b] Statistically significant ($\alpha < 0.01$)

[c] Statistically significant ($\alpha < 0.001$)

區別分析研讀報告

Obenour, W., Lengfelder, J., & Groves, D. (2005). The development of a destination through the image assessment of six geographic markets. *Journal of Vacation Marketing*, *11*(2), 107-119.

Yuksel, A., & Yuksel, F. (2002). Measurement of tourist satisfaction with restaurant services: a segment-based approach. *Journal of Vacation Marketing*, *9*(1), 52-68.

Laroche, M., Bergeron, J., & Barbaro-Forleo, G. (2001). Targeting consumers who are willing to pay more for environmentally friendly products. *The Journal of Consumer Marketing, 18*, 503-520.