# IBM SPSS Categories 19

Jacqueline J. Meulman

Willem J. Heiser

SPSS Inc.

Note: Before using this information and the product it supports, read the general information under Notices第 281 頁.

This document contains proprietary information of SPSS Inc, an IBM Company. It is provided under a license agreement and is protected by copyright law. The information contained in this publication does not include any product warranties, and any statements provided in this manual should not be interpreted as such.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

**© Copyright SPSS Inc. 1989, 2010.**

# 序

IBM® SPSS® Statistics為分析資料的強大系統。類別 的選用性附加模組能提供其他本手冊所說明的分析技術。類別 的附加模組必須與 SPSS Statistics Core 系統搭配使用，而且是完全整合到系統中。

## 關於 SPSS Inc.，是一家 IBM 公司

SPSS Inc.，是一家 IBM 公司，為全球領先的預測分析軟體和解決方案供應商。該公司完整的系列產品 — 資料收集、統計量、模型製造與部署 — 捕捉人們的態度和意見，預測客戶未來的互動結果，然後將分析融入業務程序，以依照所得見解採取行動。SPSS Inc. 解決方案藉由著重於收斂性分析、IT 架構和業務程序，以達成整個組織相互關聯的業務目標。全球商業、政府和學界客戶均仰賴 SPSS Inc. 技術為競爭優勢，以吸引、留住和增加客戶人數，同時減少欺詐並降低風險。SPSS Inc. 在 2009 年 10 月由 IBM 收購。如需詳細資訊，請造訪 http://www.spss.com。

## 技術支援

技術支援可提供客戶維護的服務。客戶可以電洽技術支援以取得 SPSS Inc. 產品在使用上的協助，或是支援硬體環境的安裝說明。如果要聯絡技術支援，請參閱 SPSS Inc. 網站（網址是 http://support.spss.com），或是透過網站（網址是 http://support.spss.com/default.asp?refpage=contactus.asp）尋找當地的辦事處。請求協助時，請準備好的您個人、組織和支援合約的相關資訊。

## 客戶服務

如果您對於自己的貨品或帳號有任何疑問，請聯絡您的當地辦公室，列示於網站上：http://www.spss.com/worldwide。請備妥您的序號以供識別。

## 訓練研討會

SPSS Inc. 同時提供公開與線上訓練研討會。所有的研討會皆以傳達工作群為其特色。研討會將定期在各主要城市舉辦。如需有關這些研討會的更多資訊，請聯絡您的當地辦公室，列示於網站上：http://www.spss.com/worldwide。

## 其他出版品

SPSS Statistics：Guide to Data Analysis（資料分析指南）、SPSS Statistics：Statistical Procedures Companion（統計程序指南）以及 SPSS Statistics：Advanced Statistical Procedures Companion（進階統計程序指南）是由 Marija Norušis 撰寫，

© Copyright SPSS Inc. 1989, 2010

由 Prentice Hall 發行，為推薦的輔助資料。這些出版品涵蓋 SPSS Statistics Base 模組、進階統計量模組和迴歸模組中的統計程序。不論您是資料分析的新手，還是已經準備使用高階應用程式，這些書籍都能幫助您善加利用 IBM® SPSS® Statistics 系列產品中的功能。如需其他資訊（包括出版品內容和章節樣本），請參閱作者的網站： http://www.norusis.com

## 致謝

IBM® SPSS® Statistics 中的最適尺度程序及其實作部份是由 Data Theory Scaling System Group (DTSS) 所開發，此機構為萊頓 (Leidon) 大學社會及行為科學學院的教育系和心理系的成員所組成。

Willem Heiser、Jacqueline Meulman、Gerda van den Berg 以及 Patrick Groenen 參與原始的 1990 程序。Jacqueline Meulman 和 Peter Neufeglise 參與類別迴歸、對應分析、類別主成分分析以及多元尺度之程序的開發。此外，Anita van der Kooij 對 CATREG、CORRESPONDENCE 和 CATPCA 特別有貢獻。Willem Heiser、Jacques Commandeur、Frank Busing、Gerda van den Berg 和 Patrick Groenen 都參與 PROXSCAL 程序的開發。Frank Busing、Willem Heiser、Patrick Groenen 和 Peter Neufeglise 都參與 PREFSCAL 程序的開發。

# 內容

## 部 I: 使用手冊


## 1 類別資料的最適尺度程序介紹 1

何謂最適尺度? . . . . . 1
為何使用最適尺度? . . . . . 1
最適尺度水準與測量水準 . . . . . 2
選擇最適尺度水準 . . . . . 2
轉換圖 . . . . . 3
類別代碼 . . . . . 3
應用上的最佳程序為何? . . . . . 5
類別迴歸 . . . . . 6
類別主成份分析 . . . . . 6
非線性典型相關分析 . . . . . 7
對應分析 . . . . . 7
多重對應分析 . . . . . 8
多元尺度方法 . . . . . 9
多元尺度展開 . . . . . 9
最適尺度圖表中的圖面比 . . . . . 10
閱讀資料推薦 . . . . . 10

## 2 類別迴歸 (CATREG) 12

在類別迴歸中定義尺度 . . . . . 13
類別迴歸類別化 . . . . . 14
類別迴歸遺漏值 . . . . . 15
類別迴歸選項 . . . . . 16
類別迴歸規則化 . . . . . 18
類別迴歸輸出 . . . . . 19
類別迴歸儲存 . . . . . 20
類別迴歸轉換圖 . . . . . 21
CATREG 指令的其他功能. . . . . 22


© Copyright SPSS Inc. 1989, 2010

3 類別主成分分析 (CATPCA) 23

在 CATPCA 中定義尺度與加權 . . . 25
類別主成分分析離散化 . . . 26
類別主成分分析遺漏值 . . . 26
類別主成分分析選項 . . . 28
類別主成分分析輸出 . . . 29
類別主成分分析儲存 . . . 31
類別主成分分析個體圖 . . . 31
類別主成分分析類別圖 . . . 32
類別主成分分析負荷量圖 . . . 33
CATPCA 指令的其他功能 . . . 34

4 非線性典型相關分析 (OVERALS) 35

定義範圍與尺度 . . . 38
定義範圍 . . . 38
非線性典型相關分析的選項 . . . 39
OVERALS 指令的其他功能 . . . 40

5 對應分析 41

定義對應分析中的列範圍 . . . 42
定義對應分析中的行範圍 . . . 43
對應分析模式 . . . 44
對應分析統計量 . . . 45
對應分析圖 . . . 46
CORRESPONDENCE 指令的其他功能 . . . 47

6 多重對應分析 48

在多重對應分析中定義變數加權 . . . 50
多重對應分析離散化 . . . 50
多重對應分析遺漏值 . . . 51
多重對應分析選項 . . . 52
多重對應分析輸出 . . . 53

多重對應分析儲存 . . . . . 55
多重對應分析個體圖 . . . . . 55
多重對應分析變數圖 . . . . . 56
MULTIPLE CORRESPONDENCE 指令的其他功能 . . . . . 57

7 多元尺度方法 (PROXSCAL) 58

近似性在直行的矩陣中 . . . . . 60
在單一行中的相似性 . . . . . 60
單一行中的相似性 . . . . . 61
建立資料的相似性 . . . . . 62
從資料建立測量法 . . . . . 63
定義多元尺度方法的模式 . . . . . 64
多元尺度方法的限制 . . . . . 65
多元尺度方法的選項 . . . . . 66
多元尺度方法的圖形 (版本 1) . . . . . 68
多元尺度方法的圖形 (版本 2) . . . . . 69
多元尺度方法的輸出 . . . . . 69
PROXSCAL 指令的其他功能 . . . . . 71

8 多維度展開 (PREFSCAL) 72

定義多維度展開模式 . . . . . 73
多維度展開限制 . . . . . 75
多維度展開選項 . . . . . 76
多維度展開圖 . . . . . 77
多維度展開輸出 . . . . . 79
PREFSCAL 指令的其他功能 . . . . . 80

部 II: 範例

9 類別迴歸 82

範例: 地毯清潔機資料 . . . . . 82
標準線性迴歸分析 . . . . . 82

類別迴歸分析 . . . . . 88
範例：臭氧資料 . . . . . 99
離散化變數 . . . . . 100
選擇轉換種類 . . . . . 100
量化的最佳值 . . . . . 111
轉換的作用 . . . . . 113
閱讀資料推薦 . . . . . 122

10 類別主成份分析 124

範例：檢驗社會系統的相互關係 . . . . . 124
執行分析 . . . . . 125
維度數 . . . . . 129
量化. . . . . 130
物件分數 . . . . . 131
成份負荷量 . . . . . 133
額外維度 . . . . . 134
範例：飲食失調症狀學 . . . . . 136
執行分析 . . . . . 137
轉換圖 . . . . . 149
模式摘要 . . . . . 151
成份負荷量 . . . . . 151
物件分數 . . . . . 153
檢驗病程結構 . . . . . 154
閱讀資料推薦 . . . . . 170

11 非線性典型相關分析 173

範例：調查結果分析 . . . . . 173
檢驗資料 . . . . . 174
說明集合之間的相似度 . . . . . 180
成份負荷量 . . . . . 183
轉換圖 . . . . . 184
單一類別對多重類別座標 . . . . . 187
重心和投影重心 . . . . . 187
其他分析 . . . . . 190
一般建議 . . . . . 194
閱讀資料推薦 . . . . . 194

12 對應分析 196

常態化 . . . 196
範例：咖啡品牌的感覺 . . . 197
執行分析 . . . 198
維度 . . . 201
貢獻 . . . 202
圖形 . . . 203
對稱常態化 . . . 204
閱讀資料推薦 . . . 206

13 多重對應分析 207

範例：硬體的特性 . . . 207
執行分析 . . . 207
模式摘要 . . . 210
物件分數 . . . 211
區別量數 . . . 212
類別量化 . . . 213
詳細查看物件分數 . . . 214
遺漏偏離值 . . . 216
閱讀資料推薦 . . . 220

14 多元尺度方法 222

範例：關係稱呼的檢查 . . . 222
選擇維度數 . . . 223
三維解 . . . 229
非預設轉換的三維解 . . . 235
討論 . . . 239
閱讀資料推薦 . . . 239

15 多維度展開 241

範例：喜愛的早餐食品 . . . 241
產生簡併解 . . . 241
量數 . . . 244
共同空間 . . . 245

執行非簡併分析 . . . . . 246
量數. . . . . 247
共同空間 . . . . . 248
範例：喜愛的早餐食品的三因子展開 . . . . . 248
執行分析 . . . . . 248
量數. . . . . 253
共同空間 . . . . . 254
個別空間 . . . . . 255
使用不同的起始組態 . . . . . 257
量數. . . . . 259
共同空間 . . . . . 260
個別空間 . . . . . 260
範例：檢查行為 - 情境適切性 . . . . . 262
執行分析 . . . . . 262
量數. . . . . 266
共同空間 . . . . . 267
相似性轉換 . . . . . 268
變更相似性轉換（次序） . . . . . 268
量數. . . . . 270
共同空間 . . . . . 271
相似性轉換 . . . . . 272
閱讀資料推薦 . . . . . 272

附錄

A 範例檔案 273

B Notices 281

參考書目 284

索引 292

# 部 I:
# 使用手冊

章節
1

# 類別資料的最適尺度程序介紹

類別程序使用最適尺度來分析標準統計程序難以分析或無法分析的資料。本章說明每個程序的作用為何、每個程序最合適的狀況、各程序之間的關係、以及這些程序及其對等的標準統計程序之間的關係。

注意：這些 IBM® SPSS® Statistics 中的程序及其實作部份是由 Data Theory Scaling System Group (DTSS) 所開發，此機構為 Leiden 大學社會及行為科學學院的教育系和心理系的成員所組成。

## 何謂最適尺度？

最適尺度的背景構想為指定數值量化給每個變數的類別，因而讓標準程序得以用在量化變數中求解。

指定最適尺度值給每一變數的類別是依照所使用的程序的最適條件來進行。有別於分析中的名義變數或次序變數的原始標記，這些尺度值都有矩陣的性質。

在大部份的類別程序中，每個尺度化變數的最適量化是透過一種稱為**改變最小平方**的疊代方法而獲得，此方法使用目前的量化來求解之後再使用此解來更新量化。然後再用更新後的量化來求得新解，再用新解來更新量化，如此持續進行直到提示需停止處理的某條件符合為止。

## 為何使用最適尺度？

市場研究、調查研究和社會及行為科學研究通常都會用到類別資料。事實上，許多研究機構幾乎都專攻類別資料的處理。

雖然專門用來分析類別資料的標準模式有很多種，通常卻未能妥善處理具有下列特性的資料集：

- 觀察值太少
- 變數太多
- 每一變數的值太多

藉由將類別加以量化，最適尺度技術得以避免上述情況的問題。甚至於另有適用的專門技術時，最適尺度仍非常有用。

最適尺度輸出通常是依照圖形顯示來解讀，而非解讀參數估計值。最適尺度技術提供了絕佳的勘查分析，同時亦可補其他 IBM® SPSS® Statistics 模式的不足。藉由縮小您的研究重點範圍，透過最適尺度將資料視覺化即可提供專注於解讀模式參數來進行分析的基礎。

© Copyright SPSS Inc. 1989, 2010

# 最適尺度水準與測量水準

當您首次使用類別程序時，在觀念上可能會非常模糊而混淆不清。指定水準時，您所指定的並非測量變數時的水準，而是將變數尺度化時的水準。其觀念在於量化的變數不論是用何種方式來測量，變數之間都可能會有非線性關係。

以類別的角度來看，共有三種基本的測量水準：

- **名義**水準表示，變數值代表未排列順序的類別。名義水準的變數範例包括地區、郵遞區號、宗教團體、以及複選類別。
- **次序**水準表示，變數值代表已排列順序的類別。範例包括代表滿意度或信心程度的態度尺度、以及偏好等級分數。
- **數值**水準表示，變數值代表含有實際意義矩陣的已排列順序類別，因此適合比較類別之間的距離。範例包括以年份表示的年齡和以千元為單位的收入。

例如，假設有三個變數，地區、工作和年齡，其編碼如下表所示。

**表格 1-1**
**地區、工作和年齡的編碼架構**

| 地區 | | 工作 | | 年齡 | |
|---|---|---|---|---|---|
| 1 | 北部 | 1 | 實習生 | 20 | 二十歲 |
| 2 | 南部 | 2 | 銷售代表 | 22 | 二十二歲 |
| 3 | 東部 | 3 | 管理員 | 25 | 二十五歲 |
| 4 | 西部 | | | 27 | 二十七歲 |

顯示的值代表每一變數的類別。地區即為名義變數。地區有四個個實質上並未排列的類別。變數值 1 到 4 只是代表四個類別；編碼架構完全是任意的。在另一方面，工作可假設為次序變數。原始類別會形成從實習生到管理員的進展。數值較高的代碼表示在企業生涯階梯的工作位階較高。但已知的資訊也僅限於排列的順序，完全沒有相鄰類別之間距離的相關資訊。反之，年齡則可假設為數值變數。在年齡這方面，數值之間的距離具有實質意義。20 和 22 之間的距離等於 25 和 27 之間的距離，22 和 25 之間的距離則大於上述兩者。

## 選擇最適尺度水準

請務必了解到，變數的任何實際性質都無法自動預先定義您應該指定何種最適尺度水準。當您在檢閱資料時，可以利用有意義且解讀上更容易的各種方式。例如，在次序水準上分析數值水準的變數，則使用非線性轉換或許能讓解的維度較少。

以下兩個範例說明，明顯的測量水準可能不是最佳的最適尺度水準。假設某變數將個體歸類到年齡群組中。雖然可將年齡尺度化成數值變數，事實卻有可能是，年齡小於 25 歲者的年齡與安全度維持正關係，年齡大於 60 歲者的年齡與安全度維持負關係。在這種情況下，將年齡視為名義變數可能比較恰當。

在另一個範例中，依照政治偏好對每個人進行歸類的變數則一定是名義變數。但如果您是依照從左派到右派來做政黨的排列，則可能會使用次序水準的分析而將政黨的量化對應到這種排列方式。

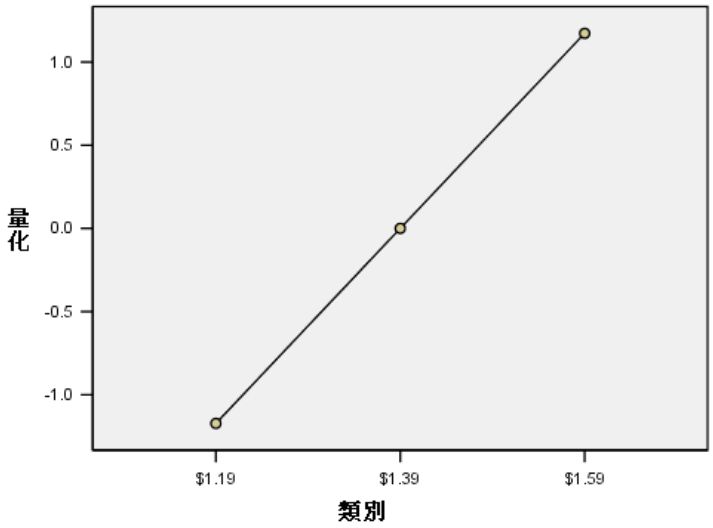
雖然變數沒有預先定義的性質可讓變數專屬於某一水準，新手還是可以藉由某些一般規則的協助來進行。使用單一名義量化，您通常並不知道類別的順序而必須由分析來套用一種順序。如果類別的順序已知，您應該嘗試使用次序量化。如果類別無法排序，您可以嘗試使用多重名義量化。

## 轉換圖

可尺度化的每個變數的不同水準都會對量化套用不同的限制。轉換圖說明了量化和選定的最適尺度水準所產生的原始類別兩者之間的關係。例如，將變數視為數值變數會產生線性轉換圖。將變數視為次序變數會產生非下降趨勢的轉換圖。將變數視為名義變數則產生顯示出二次關係的 U 形（或倒 U 形）轉換圖。名義變數亦可藉由完全變更類別的順序而產生沒有明顯趨勢的轉換圖。下圖顯示範例轉換圖。

轉換圖特別適合用來判斷選定的最適尺度水準的成效。如果許多類別都得到類似的量化，將這些類別合併成一個類別可能會更實際。或者，如果視為名義的變數所得到的量化顯示出上升趨勢，次序轉換可能也會得到類似的結果。若是顯示線性趨勢，則採用數值處理可能較為合適。但如果合併類別或變更尺度水準更加實際，分析上就不會有太大的改變。

**圖表 1-1**
**價格轉換圖（數值）**




## 類別代碼

對類別變數加以編碼時必須稍加注意，因為某些編碼架構可能會得到非預期的輸出或使分析不夠完整。工作的可能編碼架構，如以下表格所示。

表格 1-2
工作的另一個編碼架構

| | 架構 | | | |
|---|---|---|---|---|
| 類別 | A | B | C | D |
| 實習生 | 1 | 1 | 5 | 1 |
| 銷售代表 | 2 | 2 | 6 | 5 |
| 管理員 | 3 | 7 | 7 | 3 |

某些類別程序必須先定義所使用的每個變數的範圍。超出此範圍的所有數值都視為遺漏值。最小的類別值永遠為 1。最大的類別值則由使用者提供。此數值並非變數的類別個數—而是最大的類別值。例如 中，架構 A 的最大類別為 3 而架構 B 的最大類別值為 7，但兩種架構都對相同的三個類別加以編碼。

變數範圍決定了會從分析中刪除的類別為何。代碼超出定義範圍的類別都會從分析中刪除。此為刪除類別的簡便方式，但可能會造成非預期的分析。定義的最大類別若不正確，則可能將有效的類別從分析中刪除。例如，定義最大類別值為 3 的架構 B 指出，工作具有 1 到 3 的類別編碼；管理員類別即被視為遺漏值。由於所有的類別實際上皆未編碼為 3，分析中的第三個類別便不含觀察值。如果您要刪除所有的管理員類別，此種分析方式即可適用。但若欲包含管理員，則必須將最大類別定義為 7，且必須根據大於 7 或小於 1 的數值對遺漏值加以編碼。

對於視同名義或次序的變數而言，類別的範圍不會影響結果。若是名義變數，則有重要影響的是標記而非與標記有關的數值。若是次序變數，類別的順序便保留在量化中；類別值本身則並不重要。形成相同類別順序的所有編碼架構其結果將會完全相同。例如，若是在次序水準下分析工作，表格中的前三種架構便具有相同的作用。類別的順序在這些架構中完全相同。在另一方面，架構 D 將第二和第三個類別對調，並且產生異於其他架構的結果。

雖然變數的許多編碼架構在作用上大多相同，代碼之間差異較小的架構則較常被使用，因為代碼會對程序所產生的輸出量造成影響。依據介於 1 和使用者定義的最大值之間的數值加以編碼的所有類別都是有效的。其中若有任何空的類別，對應的量化即為系統遺漏值或 0，依程序而定。上述指定方式雖然都不會影響分析，但輸出則是針對這些類別而產生。因此在架構 B 中，工作有四個類別接收系統遺漏值。在架構 C 中則也有四個類別接收系統界定的遺漏指標。相反的，架構 A 並沒有系統界定的遺漏量化。使用連續整數作為名義變數或次序變數的代碼，即可產生較少的輸出而又不致影響結果。

在編碼架構的限制上，視為數值的變數比視為次序的變數更高。對於這些變數，連續類別之間的差異很重要。以下表格會針對年齡顯示三種編碼架構。

表格 1-3
年齡的可用編碼架構

| | 架構 | | |
|---|---|---|---|
| 類別 | A | B | C |
| 20 | 20 | 1 | 1 |
| 22 | 22 | 3 | 2 |
| 25 | 25 | 6 | 3 |
| 27 | 27 | 8 | 4 |

對數值變數重新編碼都必須保留類別之間的差異。使用原始數值是確保差異得以保留的一種方式。但這種方式可能會造成許多包含系統界定遺漏指標的類別。例如，架構 A 提供了原始觀察值。在「對應分析」以外的所有「類別」程序中，最大類別值為 27 而最小類別值則設為 1。最前面的 19 個類別為空的類別並且收到系統界定的遺漏指標。如果最大類別遠大於 1，便會迅速地產生多餘的輸出，並在 1 和最大值之間出現許多空的類別。

若要降低輸出量，您可以利用重新編碼。但若是在數值變數的情況下，則不應使用「自動重新編碼」。編碼為連續整數會造成 1 和所有連續類別之間的差異，而且所有量化之間的間隔都會相等。將變數視為數值時相當重要的矩陣特性會由於重新編碼為連續整數而遭到破壞。例如，表格中的架構 C 對應到自動重新編碼的 年齡。類別 22 和 25 之間的差異已從三變更為一，量化亦反映出後者的差異。

另一種可保留類別之間差異性的重新編碼架構為，將最小的類別值從每個類別中減去，每個差異再加入 1。架構 B 即這種轉換的結果。最小的類別值 20 已從每一類別中減去，再將 1 加入每個結果中。轉換後的代碼其最小值為 1，且所有的差異都和原始資料相同。現在，最大的類別值為 8，且第一個非零量化之前的所有零量化已全部消除。再者，對應到結構 B 所產生之每個類別的非零量化皆與結構 A 中的量化相同。

# 應用上的最佳程序為何？

四種程序（對應分析、多重對應分析、類別主成份分析、及非線性典型相關分係）的技術內涵同屬多變量資料分析的一般領域，稱為**維度縮減**。亦即變數之間的關係可用較少的維度來表示，通常使用—兩到三個—維度即已足夠。如此讓您可利用原本由於多樣且複雜而難以理解的關係來描述結構或樣式。在市場研究的應用中，這些技術可能是**知覺對應**的一種形式。這些程序的主要優點在於，它們能夠顯示各種最適尺度水準的資料。

類別迴歸描述了類別依變數和預測變數組合之間的關係。對應的迴歸加權值則描述每個預測變數對於依變數的影響。正如同其他的程序，各種最適尺度水準都可以用來分析資料。

「多元尺度方法」和「多維度展開」會使用個體之間的相似性來描述低維度空間中的個體之間的關係。

以下是各個程序的簡要指導原則：

- 使用「類別迴歸」，從類別自變數的組合來預測類別因變數的值。
- 使用「類別主成份分析」，藉由混合式最適尺度水準的一組變數來說明樣式的變化。
- 使用「非線性典型相關分析」來評估混合式最適尺度水準的兩組或多組變數相關時的量數。
- 使用「對應分析」來分析二因子列聯表，或是可表示成二因子表格的資料，例如品牌偏好或社調資料。
- 使用「多重對應分析」來分析類別多變量資料矩陣，前提是並未由於做出較強烈的假設而使所有的變數都在名義水準下進行分析。
- 使用「多元尺度方法」來分析近似性資料，以便在低維度空間中找出一組個體的最小平方表示法。
- 使用「多維度展開」來分析近似性資料，以便在低維度空間中找出二組個體的最小平方表示法。

## 類別迴歸

當您的分析目標是要從一組自變數（預測變數）中預測出依變數（反應值）時，使用「類別迴歸」最為恰當。正如同所有的最適尺度程序，尺度值是指定給每一變數的每個類別，因此這些值對於迴歸而言都是最適值。類別迴歸的解將轉換後的反應值和轉換後的加權預測變數之間的相關值平方最大化。

**和其他類別程序的關係。**最適尺度的類別迴歸和最適尺度典型相關分析，這兩個程序可加以比較。後者有兩個變數集，其中一個僅包含依變數。以後者的技術而言，變數集的相似性是藉由將每個變數集與位於所有變數集之間的某個未知變數集做比較而求得。在類別迴歸中，轉換後的依變數與轉換後的預測變數線性組合，兩者之間的相似性則是直接評估而求得。

**和標準技術的關係。** 在標準的線性迴歸中，可將類別變數重新編碼為指標變數，或是採用區間水準變數的方式來處理。以第一種方式而言，模式中包括每種類別變數的水準組合的個別截距和斜率。因此必須解讀大量參數。使用第二種方式則每個變數只需估計一個參數。但類別可以任意編碼的本質卻使得概化完全不可行。

如果有些變數不是連續的，即可使用替代分析。如果依變數是連續的而預測變數是類別的，則通常進行變異數分析。如果依變數是類別的而預測變數是連續的，則適合 Logistic 迴歸或判別分析。如果依變數和預測變數都是連續的，通常都是使用對數線性模式。

最適尺度的迴歸的每個變數各有三種尺度水準。這些水準的組合說明廣泛的非線性關係不適用於單一的「標準」方法因此，最適尺度比標準方式更具彈性，然而增加的複雜性卻最少。

此外，預測變數的非線性轉換通常可降低預測變數之間的相依性。如果您將預測變數相關矩陣的特徵值和最適尺度化預測變數相關矩陣的特徵值做比較，後者的變數個數通常少於前者。換言之，類別迴歸中的最適尺度會讓預測變數相關矩陣的較大特徵值變小，並讓較小的特徵值變大。

## 類別主成份分析

當您要在混合式最適尺度水準的一組變數中，說明變異數的樣式時，使用「類別主成份分析」最為合適。此項技術會儘可能地說明變異數，而同時嘗試降低該組變數的維度值。尺度值是指定給每一變數的每個類別，因此這些值對於主成份的解而言已最適化。分析中的個體依照量化的資料來接收成份分數。成份分數的圖形會顯示出分析個體之間的樣式，且亦可顯示出資料中的不尋常個體。類別主成份分析的解會在已指定成份（維度）數目的情況下，利用每個量化的變數將個體分數的相關性最大化。

類別主成份的一項重要應用在於檢查偏好資料，即受訪者依照偏好程度而對多個項目進行等級區分或評估。在一般的 IBM® SPSS® Statistics 資料組態中，橫列是個別的，直行是項目的測量，整列的分數是偏好分數（例如從 0 到 10 分），因此資料是依照橫列條件來做分析。對於偏好資料，您可能會想要將個別的資料視為變數。利用轉置程序，您就可以將資料轉置。評估者即成為變數，且所有的變數都會被宣告成次序的。使用的變數個數超過 CATPCA 中的個體個數並不會有所衝突。

**和其他類別程序的關係。**所有的變數如果都宣告成多重名義量數，類別主成份分析的分析方式便相當於在相同變數上執行的多重對應分析。因此，類別主成份分析可以視為多重對應分析的一種類型，而其中某些變數被宣告成次序的或數值的。

**和標準技術的關係。**如果所有變數的尺度都在數值水準上，則類別主成份分析相當於標準的主成份分析。

更通俗的說法為，類別主成份分析是替代方式，可用來計算非數值尺度之間的相關性，並使用標準的主成份或因子分析方式來分析相關性。單純將通用的 Pearson 相關係數視為次序資料的關連量數來使用可能導致估計上的相關性嚴重偏差。

## 非線性典型相關分析

「非線性典型相關分析」對於許多不同的應用而言是極為常見的程序。非線性典型相關分析的目標是要分析兩組或多組變數之間的關係，而不像主成份分析在於分析變數本身之間的關係。例如，您有兩組變數，其中一組只是提供給應答者作為圖例的背景項目，另一組則是一組態度項目的反應。分析中的尺度水準可以是名義、次序或數值的任意組合。最適尺度典型相關分析藉由同時比較每一組變數中的典型變數和指定給個體的一組折衷分數來判斷各組之間的相似性。

**和其他類別程序的關係。** 若兩組或多組變數中的每一組都只有一個變數，最適尺度典型相關分析相當於最適尺度主成份分析。若單組單變數分析（如上述）中的所有變數都是多重名義量數，最適尺度典型相關分析相當於重對應分析。若有兩組變數，其中一組只包含一個變數，最適尺度典型相關分析相當於最適尺度的類別迴歸。

**和標準技術的關係。** 標準的典型相關分析是一項統計技術，可找出兩者之間具有最大相關性的一組變數的線性組合和另一組變數的線性組合。有了一組線性組合，典型相關分析即可找出後續各組獨立的線性組合，稱為典型變數。其最大數量等於最小變數集的變數個數。

分析中若有兩組變數且變數全部定義為數值的，最適尺度典型相關分析相當於標準的典型相關分析。雖然 IBM® SPSS® Statistics 並沒有典型相關分析程序，但許多相關統計量可從多變量變異數分析中取得。

最適尺度典型相關分析還有其他不同的應用。若有兩組變數，其中一組包含宣告為單一名義量數的名義變數，則可使用類似於迴歸分析的方式來解讀最適尺度典型相關分析的結果。若上述變數為複名義量數，最適尺度分析便可替代判別分析。將變數分組為兩組以上即可使用各種不同的方式來分析資料。

## 對應分析

對應分析的目標是要製作對應表格的行列散佈圖。在對應表格中，是將橫列變數和直行變數假設為代表未排列順序的類別；因此永遠使用名義最適尺度水準。目的只在於檢閱這兩種變數的名義資訊。亦即，唯一的考量只是要知道有些個體位於同一類別，其他個體則不是。對於同一變數的各類別而言，其間的距離或順序完全未做假設。

對應分析的特殊作用在於分析二因子列聯表。如果表格有 r 個作用中橫列，c 個作用中直行，則對應分析中解的維度個數等於 r 的最小值加 1，或是 c 減 1，即從這兩值中取其小者。換言之，您可以將列聯表的列類別和行類別完整地呈現於維度的空間。但事實上，您可能想在低維度空間中呈現二因子表格的橫列和直行類別，例如兩個維度。因為二維圖形比多維空間表示法更容易理解。

當使用的維度少於最大可能維度時，分析所產生的統計量即說明了低維度表示法中，橫列和直行類別的分佈情形。再加上二維的解所呈現的品質極佳，因此您可以檢查列點數圖和行點數圖，以便了解橫列變數有哪些類別是相似的，直行變數有哪些類別是相似的，以及橫列類別和直行類別彼此之間的相似性。

**和其他類別程序的關係。**簡單的對應分析僅限於二因子表格。如果要探討兩個以上的變數，您可以將變數組合成交互變數。例如，如果是變數地區、工作和年齡，您可以將地區和工作組合，以建立含有 12 個類別的新變數重新工作，如下表所示。這個新變數構成包含年齡的二因子表格 (12 個橫列和 4 個直行)，可用對應分析來進行分析。

**表格 1-4**
**地區和工作的組合**

| 類別編碼 | 類別定義 | 類別編碼 | 類別定義 |
|---|---|---|---|
| 1 | 北部，實習生 | 7 | 東部，實習生 |
| 2 | 北部，銷售代表 | 8 | 東部，銷售代表 |
| 3 | 北部，管理員 | 9 | 東部，管理員 |
| 4 | 南部，實習生 | 10 | 西部，實習生 |
| 5 | 南部，銷售代表 | 11 | 西部，銷售代表 |
| 6 | 南部，管理員 | 12 | 西部，管理員 |

這種方式有一缺點，即任何一對變數皆可組合。我們可以組合工作和年齡而得到另一個包含 12 類別的變數。或者亦可組合地區和年齡而得到包含 16 個類別的新變數。這些交互變數都會個別構成一個包含其餘變數的二因子表格。這三個表格的對應分析並不會得到相同的結果，但每一種都是有效的分析方式。若是四個或更多變數時，甚至可以進一步建立二因子表格，用來進行各交互變數之間的比較。即使變數不多，但可用來分析的表格數目可能非常多。您可以選擇其中一個表格來做分析，或是分析全部。或者亦可使用「多重對應分析」程序來同時檢查所有的變數，而不需建立交互變數。

**和標準技術的關係。**「交叉表」程序也可以用來分析列聯表，即利用獨立性作為一般的分析重點。但即使是很小的表格，偵測偏離獨立性仍有困難。對應分析公用程式的作用在於顯示各種大小的二因子表格的此類樣式。如果列變數和行變數之間有相關性，亦即卡方值是有意義的，則對應分析可協助顯示出此關係的本質。

## 多重對應分析

多重對應分析會試著產生一種解，在圖形中，相同類別內的個體距離會比較近，而不同類別內的個體距離會比較遠。每一個物件會盡量接近包含該物件的類別的類別點。如此一來，即由類別將個體區分為同質性的次組別。將相同類別中的個體歸類到同一次組別的變數即為同質性變數。

在一維的解中，多重對應分析將最適尺度值（類別量化）指定給每個變數的每一類別，因此在整體上平均而言，類別具有最大的分佈。若是二維的解，多重對應分析會找出每個變數的第二組類別量化，而此量化與第一組量化無關，再試著將分佈最大化，並持續進行。由於變數的類別所收到的分數個數和維度個數相同，因此假設分析中的變數為最適尺度水準中的多重名義量數 。

多重對應分析亦指定分數給分析中的個體，因此類別量化亦即該類別中個體的個體分數平均數，或稱為重心。

**和其他類別程序的關係。**亦稱為同質性分析或是雙尺度。當變數只有兩個時，同質性分析會提供可比較但不相同的結果給對應分析。而由對應分析產生一個獨特的輸出，用以概述解的適合度和品質，包括穩定性資訊。因此，對應分析通常適用於雙變數情況下的多重對應分析。這兩個程序的另一差異在於，輸入到多重對應分析中的是資料矩陣，此時的列為個體而行為變數；但輸入到對應分析中的則可為相同的資料矩陣、一般的相似

性矩陣、或是聯合的列聯表，此為聚集矩陣，其中的列和行皆代表變數的類別。多重對應分析亦可視為多重名義水準下，尺度化資料的主成份分析。

**和標準技術的關係。**多重對應分析可視為多因子列聯表的分析。多因子列聯表也可以使用「交叉表」程序來分析，但交叉表所提供的是每個控制變數中，每一類別的個別摘要統計量。利用多重對應分析，即可藉由一張二維圖形來建立所有變數之間的關係摘要。多重對應分析的進階使用，是要以第一個維度中的最適尺度值來置換原始類別值，再進行次要的多變量分析。由於多重對應分析使用數值尺度值來置換類別標記，因此可在多重對應分析之後套用許多需要數值資料的不同程序。例如，「因子分析」程序所產生的第一個主成份相當於多重對應分析的第一個維度。第一個維度中的成份分數等於個體分數，且成份平方負荷量等於區別量數。但第二個多重對應分析維度卻不等於因子分析的第二個維度。

## 多元尺度方法

當您的分析目標是要找出一組個體或觀察值之間，一組距離量數的結構時，使用「多元尺度方法」最為合適。可藉著將觀察值指定到概念的低維度空間中特定位置來達成此目的，而空間中的點間距須儘可能符合規定的相似（相異）性。其結果為低維度空間中物件的最小平方表示，在許多例子中，這樣可協助您進一步瞭解您的資料。

**和其他類別程序的關係。**當您擁有多變量資料並使用它們來建立距離，然後再使用多元尺度方法來做分析時，其結果類似於使用包含個體主成份分析常態化的類別主成份分析來分析資料。這種 PCA 亦稱為「主成份座標分析」。

**和標準技術的關係。**「類別多元尺度程序」(PROXSCAL) 針對 Statistics Base 選項 (ALSCAL) 中可用的尺度程序提供了諸多改善。PROXSCAL 為某些模式提供了更快的演算法，並允許您在通用空間中加上限制。PROXSCAL 甚至嘗試將常態化原始壓力最大化，而非 S 壓力（亦稱為**應變**）。之所以偏好於常態化原始應力，是因為它是依照距離來測量，而 S 應力則是依照距離平方。

## 多元尺度展開

當您的分析目標是要找出二組個體之間（也稱為列或行物件），一組距離量數的結構時，使用「多維度展開」最為合適。可藉著將觀察值指定到概念的低維度空間中特定位置來達成此目的，而空間中的點間距須儘可能符合規定的相似（相異）性。其結果為低維度空間中列與行物件的最小平方表示，在許多例子中，這樣可協助您進一步瞭解您的資料。

**和其他類別程序的關係。**如果您的資料是由一組物件（平方，對稱矩陣）之間的距離構成，請使用「多元尺度方法」。

**和標準技術的關係。**「類別多維度展開」程序 (PREFSCAL) 針對 Statistics Base 選項 (ALSCAL) 中可用的展開功能提供了諸多改善。PREFSCAL 可以讓您為共同空間設限；甚至，PREFSCAL 會嘗試將懲罰應力係數測度降至最低，以避免解答變為無效（舊演算法有此傾向）。

# 最適尺度圖表中的圖面比

最適尺度圖形中的圖面比是等方性的。在二維圖形中，維度 1 中代表一個單位的距離等於維度 2 中代表一個單位的距離。若在二維圖形中變更某一維度的範圍，系統便會變更其他維度的大小，以便維持實際距離相等。最適尺度程序的等方性圖面比無法覆寫。

# 閱讀資料推薦

請參閱下文，以取得最適尺度技術的一般資訊：

Barlow, R. E., D. J. Bartholomew, D. J. Bremner, 和 H. D. Brunk. 1972. Statistical inference under order restrictions (次序限制下之統計推論). 紐約: John Wiley and Sons.

Benzécri, J. P. 1969. Statistical analysis as a tool to make patterns emerge from data (透過統計分析從資料中找出樣式). 於： Methodologies of Pattern Recognition (樣式辨識方法論), S. Watanabe，ed. 紐約: Academic Press (學術出版社).

Bishop, Y. M., S. E. Feinberg, 和 P. W. Holland. 1975. Discrete multivariate analysis: (離散多變量分析：)Theory and Practice (理論與實踐). Cambridge, Mass. (麻州劍橋): MIT Press (MIT 出版部).

De Leeuw, J. 1984. The Gifi system of nonlinear multivariate analysis (非線性多變量分析 Gifi 系統). 於： Data Analysis and Informatics III (資料分析與資訊學 III), E. Diday, et al. (Diday 與其他)，ed..

De Leeuw, J. 1990. Multivariate analysis with optimal scaling (最適尺度多變量分析). 於： Progress in Multivariate Analysis (多變量分析進度), S. Das Gupta, 和 J. Sethuraman，eds. Calcutta (加爾各答): Indian Statistical Institute (印度統計中心).

De Leeuw, J., 和 J. Van Rijckevorsel. 1980. HOMALS and PRINCALS—Some generalizations of principal components analysis (HOMALS 與 PRINCALS — 主成份分析泛論). 於： Data Analysis and Informatics (資料分析與資訊學), E. Diday, et al. (Diday 與其他)，ed. Amsterdam (阿姆斯特丹): North-Holland (北荷蘭).

De Leeuw, J., F. W. Young, 和 Y. Takane. 1976. Additive structure in qualitative data: (量化資料之附加結構)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 41, .

Gifi, A. 1990. Nonlinear multivariate analysis (非線性多變量分析). Chichester: John Wiley and Sons.

Heiser, W. J., 和 J. J. Meulman. 1995. Nonlinear methods for the analysis of homogeneity and heterogeneity (均齊性與異質性非線性分析). 於： Recent Advances in Descriptive Multivariate Analysis (敘述多變量分析最新發展), W. J. Krzanowski，ed. Oxford (牛津): Oxford University Press (牛津大學出版部).

Israëls, A. 1987. Eigenvalue techniques for qualitative data (性質資料的特徵值技巧). Leiden (萊頓): DSWO Press (DSWO 出版社).

Krzanowski, W. J., 和 F. H. C. Marriott. 1994. Multivariate analysis: (多變量分析)Part I, distributions, ordination and inference (第一部，分配、順序與推論). 倫敦: Edward Arnold.

Lebart, L., A. Morineau, 和 K. M. Warwick. 1984. Multivariate descriptive statistical analysis (多變量敘述統計分析). 紐約: John Wiley and Sons.

最大值, J. 1960. Quantizing for minimum distortion (最小誤差量化). Proceedings IEEE (Information Theory) (IEEE (資訊理論) 會議記錄), 6, .

Meulman, J. J. 1986. A distance approach to nonlinear multivariate analysis (非線性多變量分析距離法). Leiden (萊頓): DSWO Press (DSWO 出版社).

Meulman, J. J. 1992. The integration of multidimensional scaling and multivariate analysis with optimal transformations of the variables (以變數最適轉換整合多元尺度方法及多變量分析). Psychometrika (心理學計量報導), 57, .

Nishisato, S. 1980. Analysis of categorical data: (類別資料分析)Dual scaling and its applications (雙尺度分析及其應用). Toronto (多倫多): University of Toronto Press (多倫多大學出版部).

Nishisato, S. 1994. Elements of dual scaling: (雙尺度方法元素)An introduction to practical data analysis (實際資料分析簡介). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

Rao, C. R. 1973. Linear statistical inference and its applications (線性統計推論及其應用), 第二版 ed. 紐約: John Wiley and Sons.

Rao, C. R. 1980. Matrix approximations and reduction of dimensionality in multivariate statistical analysis (多變量統計分析之矩陣近似法與維度縮減). 於： Multivariate Analysis, Vol. V (多變量分析第五冊), P. R. Krishnaiah，ed. Amsterdam (阿姆斯特丹): North-Holland (北荷蘭).

Roskam, E. E. 1968. Metric analysis of ordinal data in psychology (心理學次序資料之距陣分析). Voorschoten: VAM.

Shepard, R. N. 1966. Metric structures in ordinal data (次序資料之距陣結構). Journal of Mathematical Psychology (數學心理學雜誌), 3, .

Wolter, K. M. 1985. Introduction to variance estimation (變異數估計簡介). 柏林: Springer-Verlag.

Young, F. W. 1981. Quantitative analysis of qualitative data (性質資料之量化分析). Psychometrika (心理學計量報導), 46, .

章節
2

# 類別迴歸 (CATREG)

**類別迴歸**乃是利用替類別指定數值的方式將類別資料量化，以便替轉換變數，產生最佳線性迴歸方程式。類別迴歸也稱為 CATREG (categorical regression)。

標準的線性迴歸分析，會將反應值 (依變數) 和加權預測 (自) 變數間差異的平方和最小化。一般而言，變數都是數值的，而 (名義的) 類別資料則編碼成二元變數，或者對比變數。結果就變成：類別變數被用來區分觀察值組別；而此方法為每一組別估計個別的參數集。其中，估計係數可反應出預測值的改變，對依變數所造成的影響。各種預測值組合皆有可能進行反應值的預測。

或者，還有別種方法，就是將反應值迴歸到類別預測變數值本身。因此，每個變數都會有一個估計係數。然而，對類別變數而言，類別值是任意的。以不同的方式將類別編碼會產生不同的係數，使得相同變數的分析難以比較。

CATREG 藉由同時製作名義、次序、和數值變數的尺度，來延伸其標準程序。此程序會將類別變數量化，以反映出原始類別的特性。此程序會將量化的類別變數，以與數值變數相同的方式處理。所以如果使用非線性轉換，就能以各種不同的水準來分析變數，以找出最合適的模式。

**範例。**類別迴歸可用來說明工作滿意程度，跟工作種類、地理區域和旅遊次數之間的相關性。您可能會發現，經理級跟旅遊次數少者，其工作滿意程度較高。而在這個範例下所產生的迴歸方程式，也可用來預測，這三種自變數之任意組合的工作滿意度。

**統計量與圖形。** 它包括：次數分配表、迴歸係數、ANOVA 表格、疊代過程、類別量化、未轉換預測值之間的相關、轉換後預測值之間的相關、殘差圖，以及轉換圖形。

**資料。** CATREG 用在類別指標變數。類別指標應為正整數。您可以使用「類別化」對話方塊，將分數值變數和字串變數轉換成正整數。

**假設。** 只能有一個反應值，但預測變數最多可以有 200 個。 資料最少必須包含三個有效的觀察值，而且有效觀察值的個數，也必須超過預測變數個數跟 1 的總和。

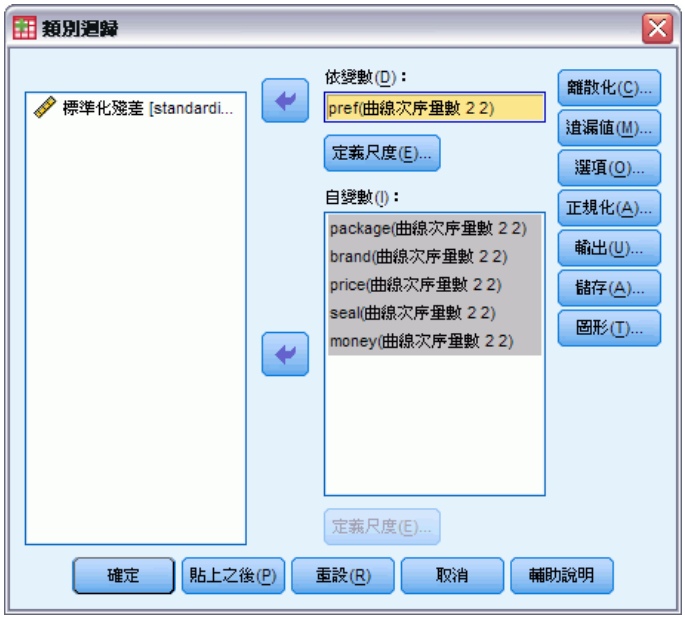
**相關程序。** CATREG 跟具最適尺度 (OVERALS) 的類別典型相關分析，這兩個程序是一樣的。 這個最適尺度有兩個變數集，其中一個僅包含一個變數。調整所有數值水準的變數，以對應至標準多重迴歸分析。

### 若要取得類別迴歸

▶ 從功能表選擇：
分析 > 迴歸方法 > 最適尺度 (CATREG)...

© Copyright SPSS Inc. 1989, 2010

圖表 2-1
「類別迴歸」對話方塊



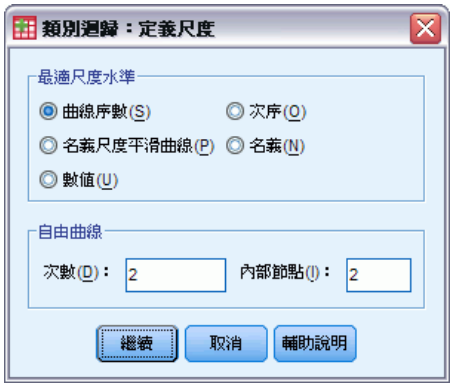

- 選取依變數和自變數。
- 按一下「確定」。

  您可以選擇性地變更每個變數的尺度水準。

# 在類別迴歸中定義尺度

您可以設定依變數及自變數的最適尺度水準。根據預設值，這些變數會調整為有兩個內部節點的二次單調曲線（次序）。此外，您也可以設定分析變數的權重。

圖表 2-2
「定義尺度」對話方塊




**最適尺度水準。** 您也可以選擇尺度水準來量化每一個變數。

- **順序尺度平滑曲線。**在最適尺度變數中，保留著已觀察變數的類別順序類別點將會位在通過原點的直線（向量）上。最後的轉換將會是所指定次數之平滑單調多項式的一部份。這一部份將由使用指定的數目、及程序決定的內部節點位置來指定。
- **名義尺度平滑曲線。**最適尺度變數中所保留之觀察變數的唯一資訊，便是類別中的個體分組。而系統並不保留觀察變數的類別順序。類別點將會位在通過原點的直線（向量）上。最後的轉換將會是所指定次數之平滑（可能為非單調）多項式的一部份。這一部份將由使用指定的數目、及程序決定的內部節點位置來指定。
- **次序。**在最適尺度變數中，保留著已觀察變數的類別順序類別點將會位在通過原點的直線（向量）上。最後的結果比順序尺度平滑曲線轉換更適合，但較不平滑。
- **名義。**最適尺度變數中所保留之觀察變數的唯一資訊，便是類別中的個體分組。而系統並不保留觀察變數的類別順序。類別點將會位在通過原點的直線（向量）上。最後的結果比名義尺度平滑曲線轉換更適合，但較不平滑。
- **數值。** 我們將類別依序排列，而且其間隔相等。類別的順序、以及觀察變數類別號碼間的相等距離，都保留在最適尺度變數中。類別點將會位在通過原點的直線（向量）上。當所有的變數都是數值水準時，分析就跟標準主成分分析很像。

# 類別迴歸類別化

您可以在「離散化」對話方塊中，選擇重新編碼變數的方法。除非您另行指定，否則系統將以接近常態分配的方式，將分數值的變數分成七種類別（如果這個數目少於 7，則視變數中不同值的數目決定）。並根據遞增的文數字順序來指定類別指標，將所有字串變數值轉換成正整數。字串變數的離散化適用這些整數。而根據預設值，其他變數則維持不變。離散化的變數稍後將用於分析中。

圖表 2-3
「離散化」對話方塊

**方法。** 選擇分組、分等級、和相乘。

- **分組。** 重新編碼為指定類別的數目、或依間隔來重新編碼。
- **分等級。** 變數將根據觀察值的等級來離散化。
- **相乘。** 變數目前的值將標準化、乘以 10、進位、並新增常數，使最低的類別化數值變成 1。

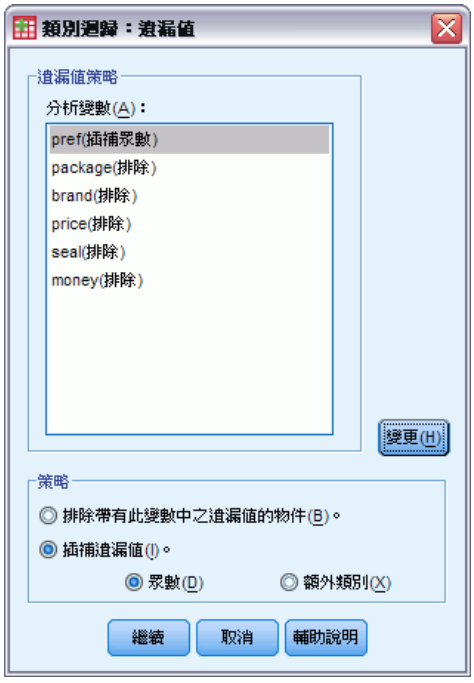
**分組。** 當根據分組離散化變數時，可以使用下列選項：

- **類別的個數。** 指定類別的數目，以及這些類別中的變數值是否應該遵循近似常態、或均勻分配。
- **等距。** 變數將重新編碼成由這些大小相等間隔所定義的類別。而您必須指定間隔的長度。

# 類別迴歸遺漏值

您可以在「遺漏值」對話方塊中，選擇處理分析變數及補充變數中遺漏值的方法。

圖表 2-4
「遺漏值」對話方塊



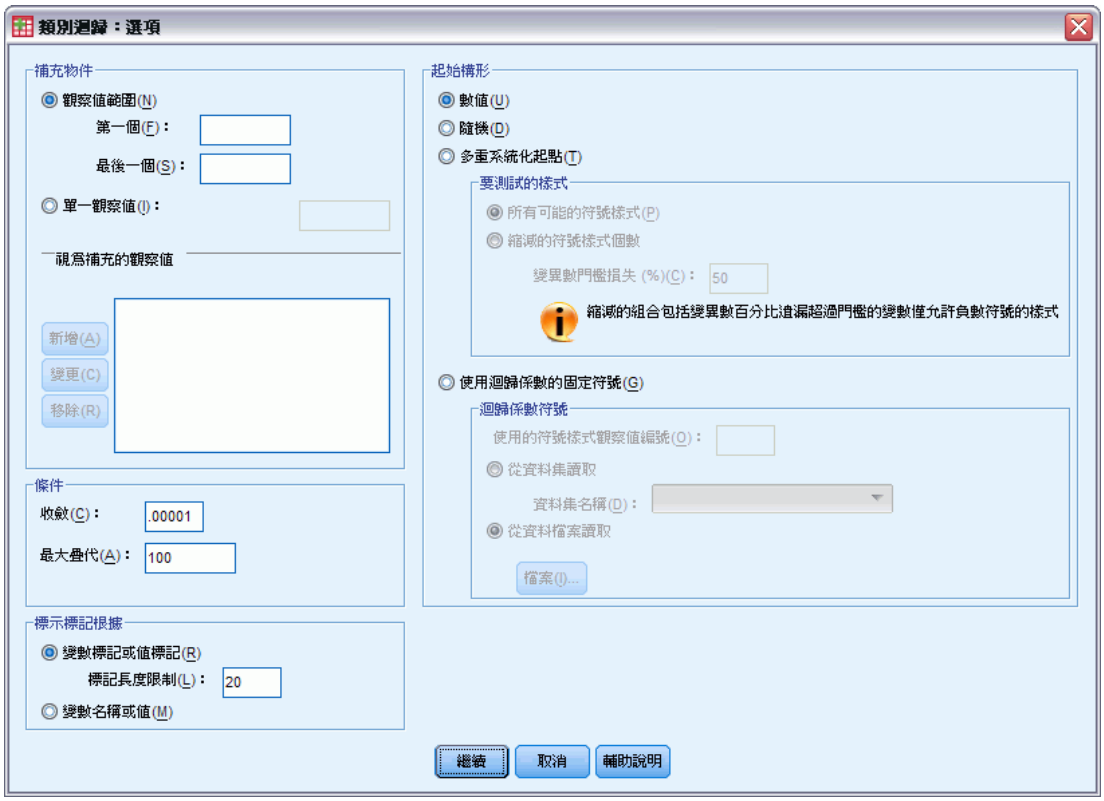

**策略。** 請選擇要排除有遺漏值的物件（刪除全部遺漏值），或是插補遺漏值（主動處理方式）。

- **排除帶有此變數中之遺漏值的物件。** 在選擇變數上有遺漏值的物件，將自分析中排除。這個策略無法用於補充變數。
- **插補遺漏值。** 系統會針對在選擇變數上有遺漏值的物件，插補其遺漏值。您可以選擇插補方法。選擇**眾數**可以次數最多的類別來取代遺漏值。當有多個眾數時，程式將使用類別指標最小的眾數。選擇**額外類別**可以額外類別的相同量化來取代遺漏值。這是表示，在這個變數上有遺漏值的物件，都將視為屬於相同的（額外）類別。

# 類別迴歸選項

「選項」對話方塊可供您選取起始組態類型，指定疊代與收斂條件，以及選擇補充性的分析標或設定圖形的標記。

圖表 2-5
「選項」對話方塊




**補充性的分析標。**此選項可讓您指定補充性的分析標。只要鍵入補充性的個體數字（或指定觀察值範圍），然後按一下「**新增**」。您不能加權補充性的分析標（會忽略指定的權重）。

**起始組態。** 如果沒有任何變數被視為名義變數，請選取**數值**組態。如果至少有一個變數被視為名義變數，請選取**隨機**組態。

此外，若有至少一個變數具有次序或順序尺度平滑曲線尺度水準，則一般的模式適合度演算法可能不會產生最適結果。選擇**多重系統起點**搭配所有可能的符號形式進行測試，始終可以獲得最佳的結果，但因為資料集中的次序或順序尺度平滑曲線變數增加，必要的處理時間也會快速地增加。您可以指定變異數門檻的適配不佳度百分比，以減少測試形式的數目，門檻越高，排除的符號形式越多。使用此選項不能保證取得最佳結果，但可降低不良結果的可能性。此外，若未找到最佳結果，也會降低不良結果和最佳結果迥異的可能性。要求多重系統起點時，每個起點的迴歸係數符號會寫入外部 IBM® SPSS® Statistics 資料檔或目前階段作業中的資料集。 如需詳細資訊，請參閱第 20 頁中的類別迴歸儲存.

前一次執行的結果和多重系統起點可讓您**針對迴歸係數使用固定符號**。符號（以 1 和 –1 表示）必須在指定資料集或檔案的列中。整數值的開始號碼是此檔案中列的觀察值號碼，該整數值則含有所要使用的符號。

**條件。** 您可以指定最大疊代數，以供計算迴歸時使用。另外，您也可以選擇收斂準則值。如果最後兩個疊代之間的總適合度差異，低於收斂數值，或者已經達到疊代的最大數目的話，迴歸就會停止疊代。

**依據...標記圖形。** 讓您指定是否要在圖形中使用變數與數值標記、或變數名稱與數值。您也可以指定標記的最大長度。

# 類別迴歸規則化

圖表 2-6
「規則化」對話方塊

**方法。** 規則化方法可將估計值縮小至趨向 0，藉以減少迴歸係數估計值差異，進而改進模式的預測錯誤。套索和彈性網絡會將部分係數估計值縮小為 0，因此可提供變數選擇格式。要求規則化方法時，每個懲罰係數值的規則化模式和係數會寫入外部 IBM® SPSS® Statistics 資料檔或目前階段作業的資料集中。 如需詳細資訊，請參閱第 20 頁中的類別迴歸儲存.

- **脊迴歸。**脊迴歸會導入等於平方係數乘以**懲罰係數**的懲罰條款，以縮小係數。此係數的範圍在 0（無懲罰）到 1之間；若您指定範圍和增量值，程序將會搜尋懲罰的最佳值。

- **套索。** 套索的懲罰條款是以絕對係數總和為基礎，懲罰係數的規格和脊迴歸的規格類似；不過，套索更需要進行大量計算。
- **彈性網絡。** 彈性網絡會將套索和脊迴歸懲罰合併，並在指定之數值網格中搜尋，以尋找最佳的套索和脊迴歸懲罰係數。對於指定的套索和脊迴歸懲罰，彈性網絡比起套索較不需要進行大量計算。

**顯示規則化圖。** 這些是規則化係數和規則化懲罰的比較圖。搜尋某範圍的值以取得最佳懲罰係數時，此比較圖可提供迴歸係數如何隨該範圍變更的檢視。

**彈性網絡圖。** 對於彈性網絡方法，脊迴歸懲罰的值會產生個別的規則化圖。**所有可能**圖會使用指定之最大和最小脊迴歸懲罰值所決定範圍中的每個值。**適用部分脊迴歸懲罰**可讓您指定最大值和最小值決定之範圍中的值的子集。只要鍵入懲罰值的數字（或指定值範圍），然後按一下**新增**。

# 類別迴歸輸出

「輸出」對話方塊，可讓您選擇在輸出中顯示的統計量。

圖表 2-7
「輸出」對話方塊

**表格。** 可產生下列項目的表格：

- **複相關係數 R.** 包含 $R^2$、調整的 $R^2$，以及列入考量的最適尺度之調整的 $R^2$。

- **ANOVA（變異數分析）。** 這個選項包括：迴歸和殘差平方和、平均數平方，和一個 F 檢定。會顯示兩個 變異數分析表：其中一個表格，其迴歸的自由度等於預測變數的個數，而另一個表格其中迴歸的自由度將最適尺度列入考慮。
- **係數。** 此選項會提供三個表格：一個包括 beta、beta 的標準誤、t 值和顯著性的係數表格；一個含有 beta 的標準誤（將最適尺度自由度列入考慮）的「係數最適尺度」表格；一個具有零階、部分、偏相關、Pratt’s 相對重要性測量，以及轉換前後之允差的表格。
- **疊代歷程。** 對於每個疊代，都會顯示演算法的起始值、複相關係數 R 及迴歸錯誤。複相關係數 R 中的增加自第一個疊代開始即列出。
- **原始變數的相關性。** 會顯示一個矩陣，其中顯示未轉換變數間的相關性。
- **轉換後變數的相關性。** 會顯示一個矩陣，其中顯示轉換後變數間的相關性。
- **規則化模式與係數。** 顯示每個規則化模式的懲罰值、R 平方和迴歸係數。若指定重新取樣方法或指定補充的物件（檢定觀察值），其也會顯示預測錯誤或檢定 MSE。

**重新取樣。** 重新取樣方法會提供您模式預測錯誤的估計值。

- **交叉驗證。** 叉驗證會將樣本分成次樣本數或折疊。類別迴歸模式接著會產生樹狀結構模式，然後從每個次樣本中排除資料。第一個模式是以第一個樣本折疊中以外的所有觀察值為基礎，第二個模式是以第二個範例折疊中以外的所有觀察值為基礎，依此類推。對於每一個模式，都會藉由將模式套用至在產生時被排除的次樣本來評估錯誤預測。
- **.632 自助法。** 使用自助法，會從資料中隨機抽取觀察值後放回，並重複此程序數次以取得多個自助法樣本。每個自助法樣本會有一個適配模式，每個模式的預測錯誤是由此適配模式所估計，然後套用至不在自助法樣本中的觀察值。

**類別變數的量化。**會出現顯示選取的變數之轉換值的表格。

**敘述統計。**會出現顯示選取變數之次數、遺漏值和眾數的表格。

# 類別迴歸儲存

「儲存」對話方塊可以讓您儲存預測值、殘差，以及將數值轉換為作用中的資料集，或/和將離散化資料、轉換值、規則化模式和係數，以及迴歸係數的符號儲存在外部 IBM® SPSS® Statistics 資料檔案或目前階段作業的資料集中。

- 可以在目前階段作業期間使用資料集，但是在後續的階段作業中則無法使用，除非您明確地將其儲存為資料檔案。資料集名稱必須符合變數命名規則。
- 檔案名稱或資料集名稱必須針對各個已儲存資料的類型而有所不同。

圖表 2-8
「儲存」對話方塊

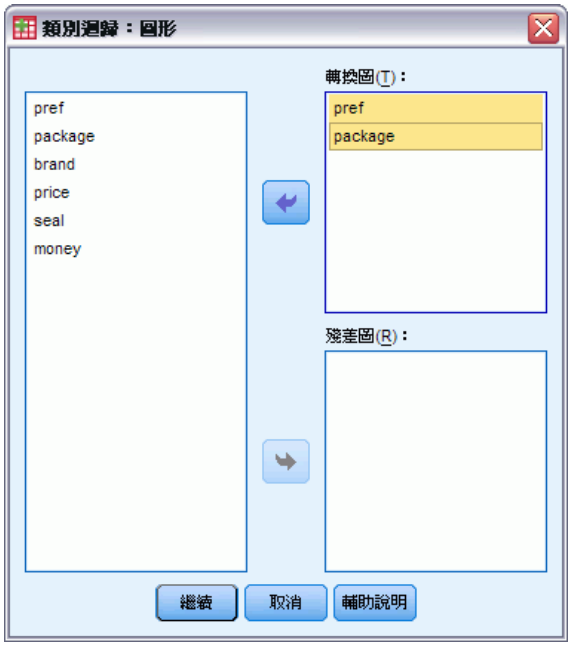
在「規則化」對話方塊中選取規則化方法時，會儲存規則化模式和係數。根據預設，此程序會建立唯一名稱的新資料集，但您可以自行指定名稱，或是寫入外部檔案。

使用多重系統起點做為「選項」對話方塊中的起始組態時，會儲存迴歸係數的符號。根據預設，此程序會建立唯一名稱的新資料集，但您可以自行指定名稱，或是寫入外部檔案。

# 類別迴歸轉換圖

「圖形」對話方塊可以讓你指定產生轉換及殘差圖形的變數。

圖表 2-9
「圖形」對話方塊




**轉換圖。** 這個功能會替每個變數，根據原始的類別數值，繪出類別量化的圖形。雖然空的類別也會出現在水平軸上，但不會影響到計算。如果需要識別這些類別，請透過連接量化之參考線中的分段，來加以識別。

**殘差圖。** 對於這些變數的每一個而言，殘差（針對所有預測變數中預測的依變數而計算，除了該預測變數本身以外）會根據類別指標而繪製圖形，而最適類別量化則根據類別指標來乘以 beta。

# CATREG 指令的其他功能

如果您將選取部分貼入語法視窗，並且編輯所產生的 CATREG 指令語法的話，就可以自訂類別迴歸。指令語法語言也可以讓您：

- 在將轉換變數存入作用中的資料集時，指定它的根名稱（利用 SAVE 次指令）。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

章節
3

# 類別主成分分析（CATPCA）

這個程序可在縮小資料的維度時，同時量化類別變數。類別主成分分析也可稱為 CATPCA，為 categorical principal components analysis 的縮寫。

主成分分析的目的，是要將原始的變數集，縮小為較小的不相關成分集，這些成分可代表在原始變數中所發現的大部分資訊。當使用大量的變數，而無法有效解釋個體（受試者和單位）之間的關係時，這個技巧十分有用。藉由縮減維度，您只需要解釋小部分元件，而不需解釋一大堆的變數。

在標準主成分分析中，系統假設數值變數之間為線性關係。而另一方面，最適尺度方法能讓變數調整為不同水準。在指定維度內的類別變數則進行最佳量化。如此一來，變數之間的非線性關係，就變成可以分析了。

**範例。** 利用類別主成分分析，可以用圖形的方式，顯示工作類別、工作部門、地區、旅遊次數（高、中、低）、和工作滿意度之間的關係。您可能會發現兩個維度便可以說明許多變異數。第一個維度會區隔工作類別與地區，而第二個維度則區隔工作部門與旅遊次數。您同時也會發現，高工作滿意度，是與中旅遊次數相關。

**統計量與圖形。** 次數、遺漏值、最適尺度水準、眾數、依重心座標說明的變異數、向量座標、每個變數和每個維度的總和、向量量化變數的成份負荷量、類別變數的量化和座標、疊代歷程、轉換後變數的相關性與相關矩陣的特徵值、原始變數的相關性與相關矩陣的特徵值、個體分數、類別圖、聯合類別圖、轉換圖、殘差圖、投影重心圖、個體圖、行列散佈圖、三角散佈圖、和成份負荷量圖。

**資料。** 字串變數值一定都會以遞增的文數字順序，轉換成正整數。使用者定義的遺漏值、系統遺漏值、和小於 1 的值都視為遺漏值；您可以重新編碼、或新增常數到變數中，讓小於 1 的變數成為非遺漏值。

**假設。** 資料最少必須含有三個有效觀察值。分析時須使用正整數資料。離散化選項將自動分類分數值的變數，也就是用接近「正常的」分配來分類其數值，並自動將字串變數的值轉換成正整數。您可以指定其他的離散化分法。

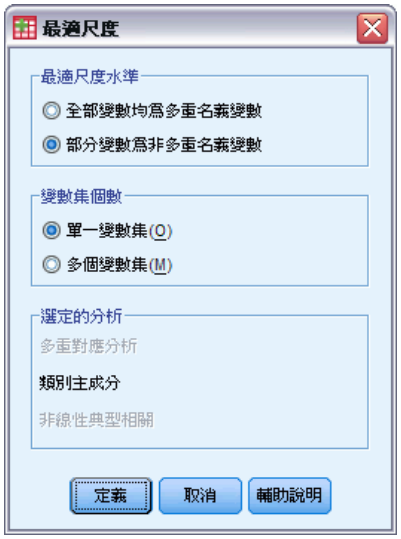
**相關程序。** 如果您是以數值水準，來製作所有變數的尺度的話，就可產生標準的主成分分析。若要使用其他繪製功能，可以在標準線性主成分分析中使用轉換變數。如果所有變數都有多重名義尺度水準，那麼類別主成分分析便與多重對應分析相同。如果您想要使用一組變數，那麼應該使用類別（非線性）典型相關分析。

### 若要取得類別主成分分析

▶ 從功能表選擇：
**分析 > 維度縮減 > 最適尺度...**

© Copyright SPSS Inc. 1989, 2010

圖表 3-1
「最適尺度」對話方塊




- 選取部份變數為非多重名義。
- 選擇「單一變數集」。
- 按一下「定義」。

圖表 3-2
「類別主成分」對話方塊

- 選擇兩個以上的分析變數，然後指定解答中的維度數。

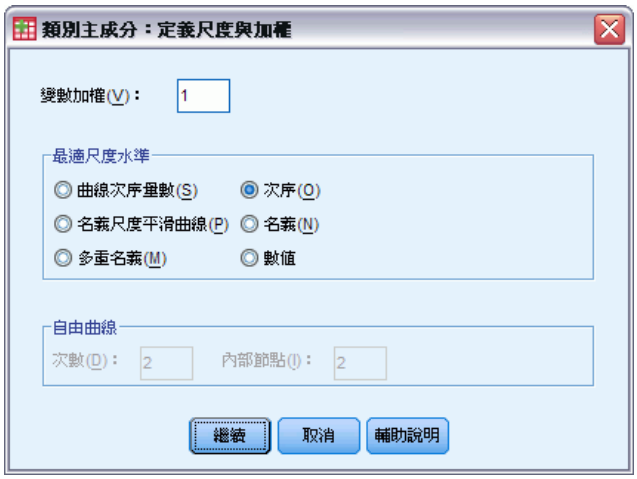
▶ 按一下「確定」。

您可以選擇指定補充變數，該變數可能適合所發現的解答、或標記圖形中的變數。

# 在 CATPCA 中定義尺度與加權

您可以設定分析變數及補充變數的最適尺度水準。根據預設值，這些變數會調整為有兩個內部節點的二次單調曲線（次序）。此外，您也可以設定分析變數的權重。

圖表 3-3
「定義尺度與加權」對話方塊




**變數加權。** 您可以選擇定義每個變數的加權。所指定的值必須為正整數。預設值為 1。

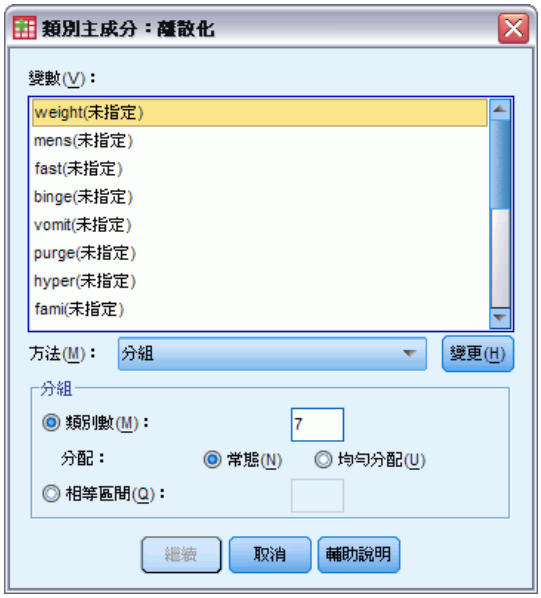
**最適尺度水準。** 此外，您還必須選取一個尺度水準，以便將每個變數加以量化。

- **順序尺度平滑曲線。**在最適尺度變數中，保留著已觀察變數的類別順序類別點將會位在通過原點的直線（向量）上。最後的轉換將會是所指定次數之平滑單調多項式的一部份。這一部份將由使用指定的數目、及程序決定的內部節點位置來指定。
- **名義尺度平滑曲線。**最適尺度變數中所保留之觀察變數的唯一資訊，便是類別中的個體分組。而系統並不保留觀察變數的類別順序。類別點將會位在通過原點的直線（向量）上。最後的轉換將會是所指定次數之平滑（可能為非單調）多項式的一部份。這一部份將由使用指定的數目、及程序決定的內部節點位置來指定。
- **複名義量數。**最適尺度變數中所保留之觀察變數的唯一資訊，便是類別中的個體分組。而系統並不保留觀察變數的類別順序。類別點將位在特定類別中個體的重心。Multiple 是表示從每個維度中取得不同組的量化。
- **次序。**在最適尺度變數中，保留著已觀察變數的類別順序類別點將會位在通過原點的直線（向量）上。最後的結果比順序尺度平滑曲線轉換更適合，但較不平滑。
- **名義。**最適尺度變數中所保留之觀察變數的唯一資訊，便是類別中的個體分組。而系統並不保留觀察變數的類別順序。類別點將會位在通過原點的直線（向量）上。最後的結果比名義尺度平滑曲線轉換更適合，但較不平滑。
- **數值。** 我們將類別依序排列，而且其間隔相等。類別的順序、以及觀察變數類別號碼間的相等距離，都保留在最適尺度變數中。類別點將會位在通過原點的直線（向量）上。當所有的變數都是數值水準時，分析就跟標準主成分分析很像。

# 類別主成分分析離散化

您可以在「離散化」對話方塊中，選擇重新編碼變數的方法。除非您另行指定，否則系統將以接近常態分配的方式，將分數值的變數分成七種類別（如果這個數目少於 7，則視變數中不同值的數目決定）。並根據遞增的文數字順序來指定類別指標，將所有字串變數值轉換成正整數。字串變數的離散化適用這些整數。而根據預設值，其他變數則維持不變。離散化的變數稍後將用於分析中。

圖表 3-4
「離散化」對話方塊




**方法。** 選擇分組、分等級、和相乘。

- **分組。** 重新編碼為指定類別的數目、或依間隔來重新編碼。
- **分等級。** 變數將根據觀察值的等級來離散化。
- **相乘。** 變數目前的值將標準化、乘以 10、進位、並新增常數，使最低的離散化數值變成 1。

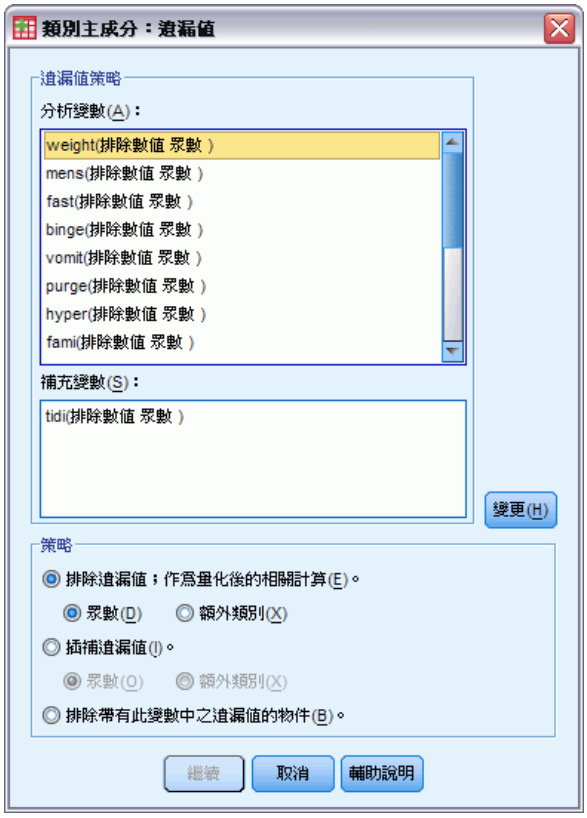
**分組。** 當您根據分組離散化變數時，可以使用下列選項：

- **類別的個數。** 指定類別的數目，以及這些類別中的變數值是否應該遵循近似常態、或均勻分配。
- **等距。** 變數將重新編碼成由這些大小相等間隔所定義的類別。而您必須指定間隔的長度。

# 類別主成分分析遺漏值

您可以在「遺漏值」對話方塊中，選擇處理分析變數及補充變數中遺漏值的方法。

圖表 3-5
「遺漏值」對話方塊




**策略。** 選擇排除遺漏值（被動處理方式）、插補遺漏值（主動處理方式）、或排除有遺漏值的物件（刪除全部遺漏值）。

- **排除遺漏值；作為量化後的相關插補。** 選擇變數中具遺漏值的物件，無利於這個變數的分析。如果對所有變數都施以被動處理方式，則所有變數中具有遺漏值的物件都會被視為補充性。如果「輸出」對話方塊中有指定相關，那麼（在分析後）系統將以變數中次數最多的類別、或眾數來插補遺漏值，以分析原始變數的相關。若是最適尺度變數的相關，您可以選擇插補的方法。選擇**眾數**可以最適尺度變數的眾數來取代遺漏值。選擇**額外類別**可以額外類別的量化來取代遺漏值。這是表示，在這個變數上有遺漏值的物件，都將視為屬於相同的（額外）類別。
- **插補遺漏值。** 系統會針對在選擇變數上有遺漏值的物件，插補其遺漏值。您可以選擇插補方法。選擇**眾數**可以次數最多的類別來取代遺漏值。當有多個眾數時，程式將使用類別指標最小的眾數。選擇**額外類別**可以額外類別的相同量化來取代遺漏值。這是表示，在這個變數上有遺漏值的物件，都將視為屬於相同的（額外）類別。
- **排除帶有此變數中之遺漏值的物件。** 在選擇變數上有遺漏值的物件，將自分析中排除。這個策略無法用於補充變數。

# 類別主成分分析選項

在「選項」對話方塊中，您可以選擇起始組態、指定疊代與收斂條件、選擇常態化方法、選擇標記圖形的方法、以及指定補充的物件。

圖表 3-6
「選項」對話方塊

**補充性的個體。** 請指定物件的觀察值號碼，或物件範圍的第一個和最後一個觀察值號碼，來將它們指定為補充的，然後按一下**新增**。繼續一直到您指定了所有的補充性的分析標為止。如果該物件已指定為補充物件，那麼程式會忽略該物件的觀察值加權。

**常態化方法。** 您可以指定五種選項中的其中一種，來常態化個體分數與變數。在指定分析中只能使用一種常態化方法。

- **變數主成份分析。** 這個選項可最佳化變數間的關係。個體空間中的變數座標，為成份負荷量（與主成份的相關，如維度與個體分數）。當您主要是對變數間的相關有興趣時，這個選項非常有用。
- **個體主成份分析。** 這個選項可最佳化個體間的距離。當您主要是對個體間的差異或相似性有興趣時，這個選項非常有用。
- **對稱的。** 如果您主要是對個體與變數間的關係有興趣時，請選取這個常態化選項。

- **自變數。** 如果您想要分別檢查個體間的距離、以及變數間的相關時，請使用這個常態化選項。
- **自訂。** 您可以指定封閉區間［-1, 1］之間的任何實數。數值 1 等於「個體主成份分析」法、數值 0 等於「對稱的」方法，而數值 -1 等於「變數主成份分析」法。利用指定大於 -1 及小於 1 的數值，您可以將特徵值散佈到個體與變數中。當您需要特別製作行列散佈圖時，這個方法非常有用。

**條件。** 您可以指定最大疊代數，以供計算迴歸時使用。另外，您也可以選擇收斂準則值。如果最後兩個疊代之間的總適合度差異，低於收斂數值，或者已經達到疊代的最大數目的話，迴歸就會停止疊代。

**依據...標記圖形。** 讓您指定是否要在圖形中使用變數與數值標記、或變數名稱與數值。您也可以指定標記的最大長度。

**繪製維度。** 讓您控制維度在輸出上的顯示方式。

- **顯示解答的所有維度。**解答中的所有維度都會顯示在散佈圖矩陣中。
- **限制維度的數目。**顯示的維度會限制為已繪製的成對資料。如果您限制維度，就必須選擇最低和最高維度來繪製。最低維度的範圍可從 1 到解答中維度的數字減 1，並且對較高的維度進行繪製。最高維度值的範圍可從 2 到解答中維度的數字，並且指出用來繪製維度成對資料的最高維度。 這個規格適用於所有要求的多維度圖。

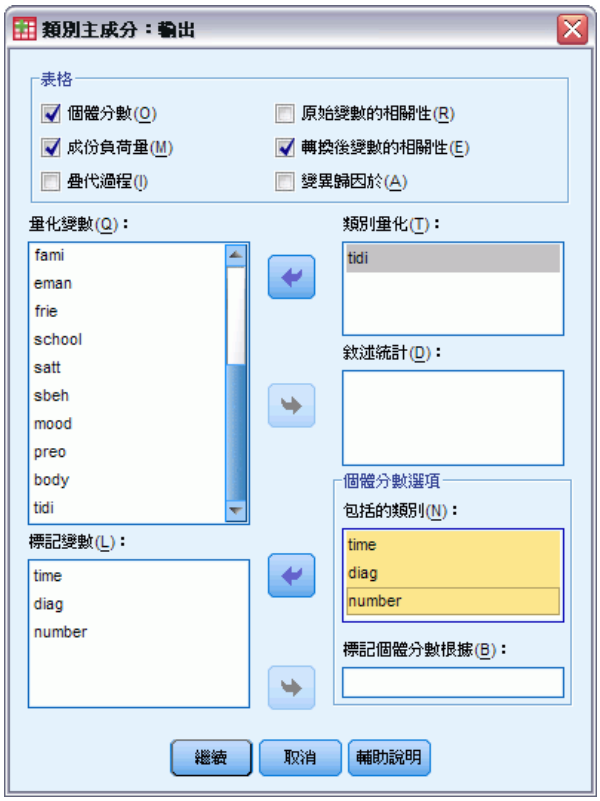
**組態。** 您可以從包含組態座標的檔案中讀取資料。在檔案中的第一個變數，必須包含第一維度的座標，而第二個變數則應該包含第二維的座標，依此類推。

- **初始。** 在指定檔案中的組態，將用來當作分析的開始點。
- **固定。** 在指定檔案中的組態，將用來配合變數。而可以配合的變數，必須選為分析變數，但因為組態是固定的，所以它們將被視為補充變數（所以它們並不需要選為補充變數）。

# 類別主成分分析輸出

在「輸出」對話方塊中，您可以建立個體分數、成分負荷量、疊代歷程、原始與轉換變數的相關、各變數與各維度所說明之變數、選取變數的類別變數量化、與選取變數之敘述統計的表格。

圖表 3-7
「輸出」對話方塊




**個體分數。** 顯示個體分數，並使用下列選項：

- **包括的類別。** 顯示所選擇分析變數的類別指標。
- **標記個體分數根據。** 您可以從指定為標記變數的變數清單中，選擇其中之一來標記物件。

**成分負荷量。** 顯示所有未指定多重名義尺度水準變數的成分負荷量。

**疊代歷程。** 針對每次疊代，顯示變異歸因於、適配不佳度、以及變異歸因於的增量。

**原始變數的相關性。** 顯示原始變數的相關矩陣，以及該矩陣的特徵值。

**轉換後變數的相關性。** 顯示轉換（最適尺度）變數的相關矩陣，以及該矩陣的特徵值。

**變異歸因於。** 顯示依每個變數與每個維度之重心座標、向量座標、與總和（重心與向量座標的組合）來說明的變異數量。

**類別變數的量化。** 指定選取變數各維度的類別變數的量化與座標。

**敘述統計。** 顯示選擇變數的次數、遺漏值數目、與眾數。

# 類別主成分分析儲存

「儲存」對話方塊可以讓您將離散化資訊、個體分數、轉換值，以及近似法儲存於外部 IBM® SPSS® Statistics 資料檔案或目前階段作業的資料集中。您也可以將轉換值、個體分數，和近似法儲存於作用中的資料集中。

- 可以在目前階段作業期間使用資料集，但是在後續的階段作業中則無法使用，除非您明確地將其儲存為資料檔案。資料集名稱必須符合變數命名規則。
- 檔案名稱或資料集名稱必須針對各個已儲存資料的類型而有所不同。
- 如果您儲存個體分數或將數值轉換為作用中的資料集，您可以指定多重名義維度的個數。

圖表 3-8
「儲存」對話方塊

# 類別主成分分析個體圖

在「個體與變數圖」對話方塊中，您可以指定所需的圖表類型，以及要用來建立圖形的變數。

圖表 3-9
「個體與變數圖」對話方塊

**個體點。** 顯示個體點的圖。

**個體與變數（行列散佈圖）。** 使用您選擇的變數座標來繪製個體點數 — 成分負荷量或變數重心。

**個體，負荷量與重心（三角散佈圖）。** 使用多重名義尺度水準變數的重心、與其他變數的成分負荷量來繪製個體點數。

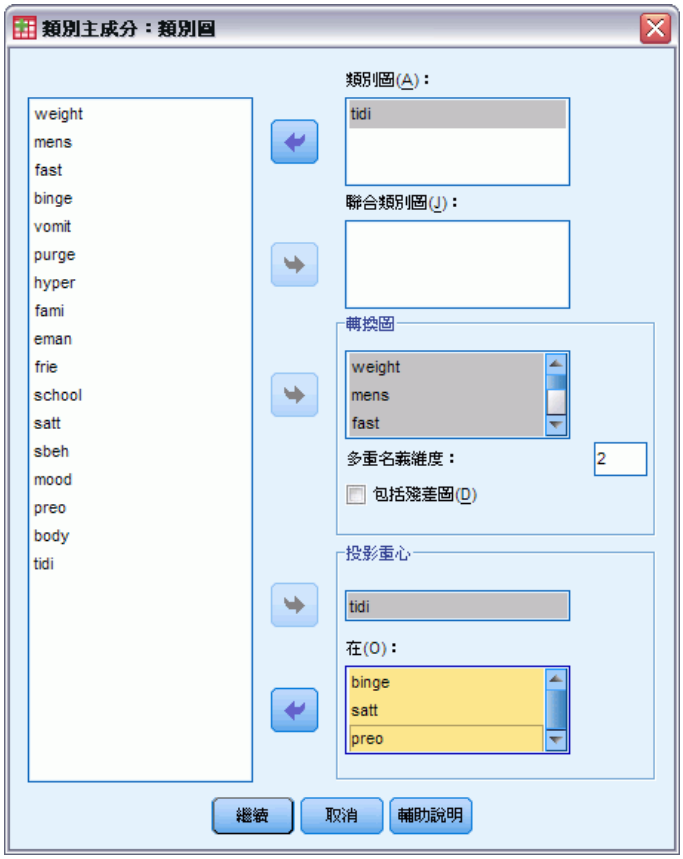
**行列散佈圖與三角散佈圖變數。** 您可以選擇使用行列散佈圖、三角散佈圖的所有變數或選取一個子集。

**標記個體。** 您可以選擇用選定變數的類別（您可以在「選項」對話方塊中選擇類別指標值或數值標記）、或其觀察值號碼來標記個體。如果您選擇**變數**，則每個變數會建立一個圖形。

# 類別主成分分析類別圖

使用「類別圖」對話方塊，可以指定所需的圖表類型，以及用來建立圖表的變數。

圖表 3-10
「類別圖」對話方塊




**類別圖。** 針對所選取的每個變數，繪製重心與向量座標圖。如果變數有多重名義尺度水準，則類別將會位在特定類別中個體的重心。如果為其他所有尺度水準，類別則會在通過原點的向量上。

**聯合類別圖。** 這是每個選取變數之單一重心圖與向量座標圖。

**轉換圖。** 顯示最佳類別變數量化對類別指標的圖形。您可以指定具有多重名義尺度水準之變數的維度數；系統會針對每個維度建立一個圖形。您也可以選擇顯示每個選定變數的殘差圖。

**投影重心。** 您可以選擇一個變數，然後將其重心投影在選擇的變數上。但不可以投影在有多重名義尺度水準的變數上。當要求這個圖形時，會同時顯示一個有投影重心之座標軸的表格。

# 類別主成分分析負荷量圖

「負荷量圖」對話方塊可讓您指定將包括在圖形中的變數，及圖形中是否要包括重心。

圖表 3-11
載入圖對話方塊

**顯示成份負荷量。**若選取此選項，會顯示成分負荷量的圖形。

**負荷量變數。** 您可以選擇使用成分負荷量圖的所有變數或選取子集。

**包括重心。** 有多重名義尺度水準的變數並沒有成分負荷量，但您可以選擇在圖形中加入那些變數的重心。您可以選擇使用所有複名義量數，或選擇一個子集。

# CATPCA 指令的其他功能

如果您將選用部分貼入語法視窗，並且編輯所產生的 CATPCA 指令語法，就可以自訂 類別主成分分析。指令語法語言也可以讓您：

- 在將轉換變數存入作用中的資料集時，指定它的根名稱（利用 SAVE 次指令）。
- 分別指定各圖形中標記的長度（使用 PLOT 次指令）。
- 指定不同的殘插圖變數清單（使用 PLOT 次指令）。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

章節
4

# 非線性典型相關分析（OVERALS）

下述兩種分析是互相對應的：非線性典型相關分析、具最適尺度的類別典型相關分析。這個程序的用意在於判斷類別變數集之間的相似程度。非線性典型相關分析，也稱為 OVERALS。

標準的典型相關分析，是從多重迴歸衍生出來的。但是，多重迴歸的第二個變數集，不只包含一個單一反應值，而是包含多個反應變數。因為它是想儘量地詳細說明，在低維度的空間中，兩個數字變數集之間，其關係上的變化。一開始時，每個變數集中的變數，都是以線性的方式組合起來，如此一來，線性組合才會產生最大相關。如果有這些組合，就可以判斷隨後的線性組合中，何者與之前的組合不相關，而何者可能會產生最大相關。

對於標準分析而言，最適尺度法已經在三個方面上，予以延伸。首先，OVERALS 可以使用兩個（或以上）變數集。其次，變數的尺度可以改成名義、次序，或是數值。如此一來，變數之間的非線性關係，就變成可以分析了。最後，變數集不會將變數集之間的關係最大化，而會跟未知折衷變數集互相比較，而且該變數集是利用個體分數來定義的。

**範例。** 您可將具最適尺度的類別典型相關分析以圖形表示，以顯示包含工作類別與教育年數之變數集、和另一個包含居住地和性別之變數集之間的關係。您可能會發現教育年數與居住地的判別力，優於其他剩餘的變數。此外，可能也會發現，判別能力最強者，落在教育年數的第一個維度上。

**統計量與圖形。** 它包括：次數分配表、重心、疊代歷程、個體分數、類別變數的量化、加權值、成份負荷量、單一及多重適合度、個體分數圖、類別座標圖、成份負荷量圖、類別重心圖、轉換圖。

**資料。** 當您替類別變數（名義或次序尺度水準）編碼時，請使用整數。若要將輸出最小化，請使用連續整數（從 1 開始），來替每個變數編碼。對於以數值水準為尺度的變數，則不能將其重新編碼成連續整數。若要將輸出最小化，對於以數值水準為尺度的每個變數而言，請將每個數值減去最小的觀察值，然後再加 1。 至於小數點後面的值，請予以捨去。

**假設。** 可將變數分類成二或多個變數集。分析中之變數區分為多重名義、單一名義、次序的或數值變數。程序中所使用的最大維度值，會隨變數的最適尺度水準而改變。如果將變數指定為次序、單一名義或數值時，維度的最大個數為以下兩個數值中之較小數值：觀察值個數減 1 或變數的總數。然而，如果您只定義兩個變數集的話，那麼維度的最大值就是：較小變數集的變數個數。如果部分變數是多重名義變數，則維度的最大值便是多重名義類別的總數，加上非多重名義變數的數目，再減去多重名義變數的數目。例如，如果分析跟五個變數有關，其中一個是多重名義量數，擁有四種類別的話，則維度的最大值就是 (4 + 4 - 1)，也就是 7。如果您指定的個數大於最大值，系統仍將使用最大值。

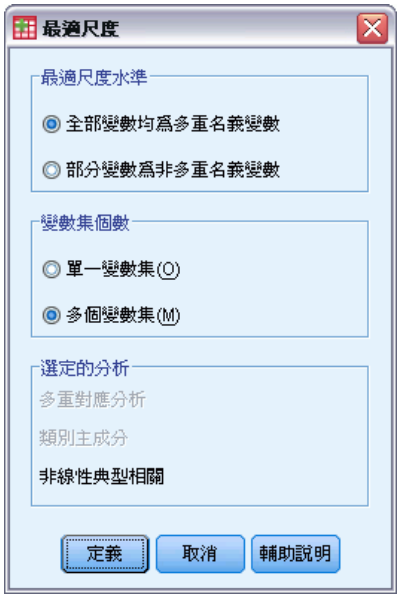
© Copyright SPSS Inc. 1989, 2010

**相關程序。** 如果每個變數集都只有一個變數，那麼非線性典型相關分析，就等於具最適尺度的主成份分析。如果每個變數都是多重名義量數，則這個分析就跟多重對應分析相對應。如果牽涉到兩個變數集，而其中一個變數集只有一個變數的話，那麼這個分析就等於具最適尺度的類別迴歸。

### 取得非線性典型相關分析

- 從功能表選擇：
  分析 > 維度縮減 > 最適尺度...

圖表 4-1
「最適尺度」對話方塊



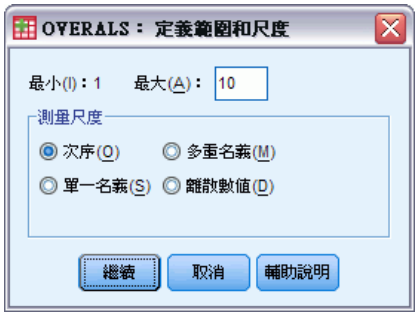

- 選擇「所有變數為多重名義量數」或「部分變數均為非多重名義量數」。
- 選擇「多個變數集」。
- 按一下「定義」。

圖表 4-2
「非線性典型相關分析 (OVERALS)」對話方塊

- 定義至少兩個變數集。選出想要加入第一個變數集中的變數。若要移至下一個變數集，請按一下「**下一個**」，並選擇第二個變數集中，所需包含的變數。此外，您還可以針對個人的需要，自行加入別的變數集。但是，如果您按一下 **前一個**，就會返回之前所定義的變數集。
- 替每個選用變數，定義數值範圍和測尺度度（最適尺度水準）。
- 按一下「**確定**」。
- 隨意：
  - 選擇一個（或多個）變數，以提供個體分數圖的點標記。每個變數會分別產生散佈圖，圖中的點會依該變數之數值來標記。您必須對這些散佈圖的每個標記變數，定義其範圍。如果您使用的是對話方塊，則對單一變數而言，進行分析時，您無法使用它，又把它當成標記變數。如果想要以分析中所用的變數，來標記個體分數圖的話，請使用「轉換」功能表上的「計算」功能，來建立該變數的副本。再使用那個新變數來標記圖形。或者使用指令語法。 指定所需之解答維度數目。
  - 指定解答中您所需要的維度數目。一般而言，會選擇最低必要維度，才能說明大部分的變化。如果分析牽涉到兩個以上的維度，則會產生前三個維度的立體圖形。至於其他的維度，您可以利用編輯圖表的方式予以顯示出來。

# 定義範圍與尺度

圖表 4-3
「定義範圍與尺度」對話方塊



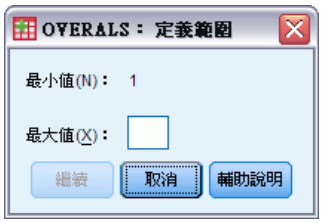

您必須定義每個變數的範圍。所指定的最大值必須是整數。而分數的資料值小數部分會於分析中捨去。而指定範圍外的類別數值則會於分析中被忽略。若要將輸出最小化，請使用「轉換」功能表上的「自動重新編碼功能，來為那些被視為名義或次序的變數，建立從 1 開始的連續類別。我們建議您，不要將透過數值來測量的變數，重新編碼成連續的整數。若要將數字變數的輸出最小化，請將每個變數的各數值減去其最小值後，再加 1。

此外，您還必須選擇一個尺度，以便將每個變數加以量化。

- **次序。** 量化變數中，保留了觀察變數的類別順序。
- **單一名義量數。** 在量化變數中，個體如果所屬類別相同，都會得到相同的分數。
- **多重名義量數。** 每個維度的量化情形，不需要相同。
- **離散數值。** 我們將類別依序排列，而且其間隔相等。類別個數間的差異，以及觀察變數的類別順序，都保留在量化的變數中。

# 定義範圍

圖表 4-4
「定義範圍」對話方塊



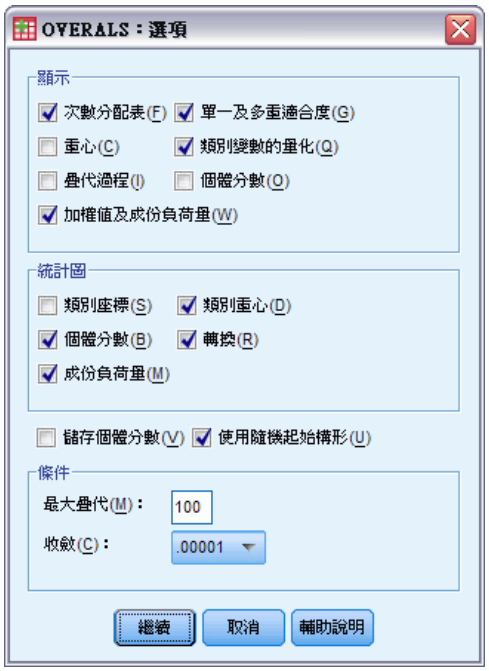

您必須定義每個變數的範圍。所指定的最大值必須是整數。而分數的資料值小數部分會於分析中捨去。而指定範圍外的類別數值則會於分析中被忽略。若要將輸出最小化，請使用「轉換」功能表上的「自動重新編碼」功能，來建立從 1 開始的連續類別。

您也必須分別定義每個用來標記個體分數圖的變數的範圍。但是如果類別標記的資料值，超出變數所定義範圍的話，還是會顯示在圖形上。

# 非線性典型相關分析的選項

您可以利用「選項」對話方塊，來選取選擇性的統計量或圖形、將個體分數儲存成作用中資料集中的新變數、指定疊代和收斂準則，以及指定分析的初始組態。

圖表 4-5
「選項」對話方塊




**顯示。** 您可以使用的統計量包括：邊際次數（個數）、重心、疊代過程、加權值及成份負荷量、類別變數的量化、個體分數，以及單一及多重適合度統計量。

- **重心。** 所屬變數類別相同，且包含於每個集合之中的個體（觀察值）類別量化，以及投影與實際個體分數平均值。
- **加權和成份負荷量。** 各　組中每個量化變數的各個維度的迴歸係數，物件分數會在該組之量化變數以及個體空間中量化變數的投影中迴歸。提供指示說明每個變數於每組中對維度的貢獻度。
- **單一和多重適合度。** 有關物件的單一和多重類別之座標/類別量化的適合度量數。
- **類別量化。** 指定給變數類別的最適尺度值。
- **個體分數。** 在特定維度指派給個體（觀察值）的最適分數。

**圖形。** 您可以繪製類別座標、個體分數、成份負荷量、類別重心，以及轉換圖。

**儲存個體分數。** 您可以在作用中資料集中，將個體分數儲存為新的變數。個體分數會根據您在主對話方塊中所指定的維度個數加以存檔。

**使用隨機起始組態。** 如果所有（或部分）變數是單一名義量數的話，則應使用隨機初始組態。如果沒有選擇這個選項的話，系統將會使用巢狀初始組態。

**條件。** 您可以指定最大疊代次數，以便在計算中，使用非線性典型相關分析。另外，您也可以選擇收斂準則值。如果最後兩疊代之間的總適合度差異低於收斂值，或者已經達到最大的疊代次數的話，分析就會停止疊代。

# OVERALS 指令的其他功能

如果您將選用部分貼入語法視窗，並且編輯所產生的 OVERALS 指令語法的話，就可以自訂非線性典型相關分析。指令語法語言也可以讓您：

- 指定所需繪製之維度成對資料，如此就不需將所有選用維度，都繪製出來（使用 PLOT 次指令上的 NDIM 關鍵字）。
- 指定數值標記字元的個數，以便標記圖形上的點（使用 PLOT 次指令）。
- 指定五個以上的變數，當作物件分數圖的標記變數（使用 PLOT 次指令）。
- 指定分析中所使用的變數，當作物件分數圖的標記變數（使用 PLOT 次指令）。
- 選擇變數以提供量化分數圖的點標記（使用 PLOT 次指令）。
- 如果您不想使用作用中資料集中，每一個觀察值的話，請指定分析中，所需包含的觀察值個數（使用 NOBSERVATIONS 次指令）。
- 在儲存個體分數時，指定所需建立的變數根名稱（使用 SAVE 次指令）。
- 指定所要存檔的維度數目即可，而不用將所有選擇的維度都儲存起來（使用 SAVE 次指令）。
- 將類別變數量化值寫入矩陣檔案中（使用 MATRIX 次指令）。
- 產生可能比一般高解析度圖形更易於閱讀的低解析度圖形（使用 SET 指令）。
- 只替指定變數，產生重心圖和轉換圖（使用 PLOT 次指令）。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

章節
5

# 對應分析

對應分析的主要目的之一，便是說明低維度空間中，在對應表格內的兩個名義變數之間的關係，同時並描述各變數類別之間的關係。對每個變數而言，圖形中，數個類別點之間的距離，會反應出類別之間的關係，而且跟這些類別很相似的其他類別，會畫在它們的旁邊。變數如果在向量上通過原點，則它跟其他變數的類別點所產生出來的投影點，會說明這些變數之間的關係。

一般而言，列聯表分析應包括檢查列和行剖面圖，並經由卡方統計量，來檢定其獨立性。但是，剖面圖的數量可能會非常龐大，而且卡方檢定也無法透露其依賴方式。所以，雖然「交叉表」程序，會提供數種關聯測量和關聯檢定，但是它無法以圖形的方式，來表示變數之間的任何關係。

因子分析是一種標準技術，用來說明低維度空間中，變數彼此之間的關係。但是，因子分析需要間隔資料，而且觀察值個數應該是變數個數的五倍。而且從另一個角度來看，對應分析除了假設名義變數之外，還可以說明變數類別彼此之間的關係，以及變數彼此之間的關係。此外，您也可以利用對應分析，來分析任何正對應測量的表格。

**範例**。我們可以利用「對應分析」，以圖形的方式來顯示職員類別和抽煙習慣之間的關係。以抽煙來說，您可能會發現，年輕的經理與秘書不同，但秘書跟年長的經理是一樣的。您也可能會發現，年輕的經理煙癮較重，而秘書的煙癮較輕。

**統計量與圖形**。它包括：對應測量、列和行剖面圖、奇異值、列和行分數、慣性、大量、列和行分數信賴統計量、奇異值信賴統計量、轉換圖、列點數圖、行點數圖，以及行列散佈圖。

**資料**。您需要分析的類別變數，是以名義方式來測量的。如果是整合資料，或是非次數的對應測量，請使用具正相似數值的加權變數。此外，如果是表格資料，請透過語法加以讀取。

**假設**。程序中所使用的最大維度值，乃是根據作用中的列個數、行類別個數，以及相等性限制個數來決定的。如果沒有使用任何相等性限制，而且所有類別都是作用中的話，則最大維度值就是：具最少類別之變數的類別個數，再減掉 1。例如，如果某變數有五種類別，而其他變數有四種，則最大維度值就是三。此處，補充性類別是不算在內的。例如，如果某變數有五種類別，其中有兩個是補充的，而其他變數有四種類別，則最大維度值便是 2。不然，就是將限制成必須相等的類別組，都視成同一個類別。舉例來說，如果一個變數有五種類別，其中三種限制為相等的，當決定最大維度時，該變數應該被視為有三種類別。兩種類別沒有限制，同時第三種類別相當於三種限制的類別。如果您指定的維度值大於最大值，系統仍將使用最大值。

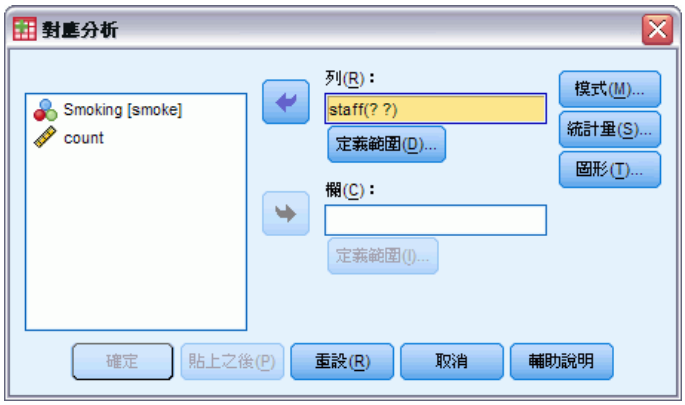
**相關程序**。如果牽涉到兩種以上的變數，請使用多重對應分析。如果變數應該以次序方式測量的話，請使用類別主成分分析。

© Copyright SPSS Inc. 1989, 2010

### 若要取得對應分析

- 從功能表選擇：
  分析 > 維度縮減 > 對應分析...

圖表 5-1
「對應分析」對話方塊。

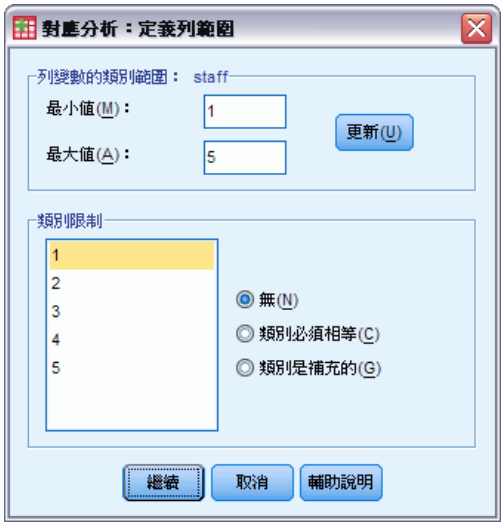



- 選取一個列變數。
- 選取一個行變數。
- 定義變數的範圍。
- 按一下「確定」。

# 定義對應分析中的列範圍

您必須定義行變數的範圍，指定的最小值和最大值必須是整數。而分數的資料值小數部分會於分析中捨去。而指定範圍外的類別數值則會於分析中被忽略。

圖表 5-2
「定義列範圍」對話方塊



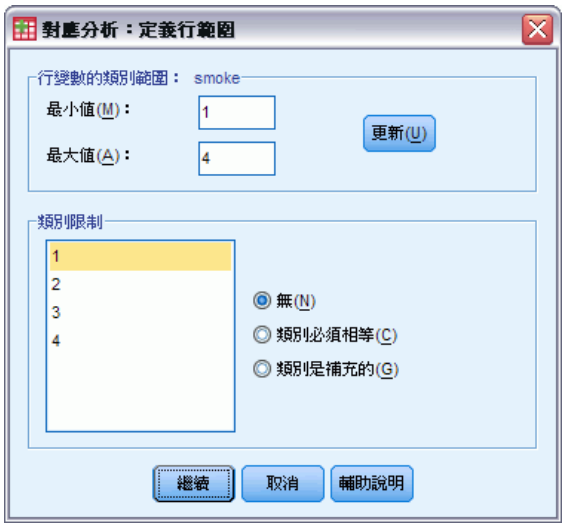
一開始時，所有的類別都沒有受到任何限制，而且都是作用中的。因此，您可以限制某些列類別，必須等於其他列類別，或者將列類別定義成補充性類別。

- **類別必須相等。**類別的分數必須相等。如果所取得的類別次序，並非您想要的，請使用相等限制。至於您所能限制為相等的最大列類別數，算法為作用列類別的總數減 1。 若要將多種相等限制，套用到類別組上，此時請透過語法加以執行。例如，您可以使用語法，限制類別 1 等於類別 2、類別 3 等於類別 4。
- **類別為補充的。** 補充類別並不會影響分析，但是卻會出現在由作用類別所定義的空間中。補充類別在定義維度上並沒有任何作用。補充列類別的最大值是列類別的總數值減 2。

# 定義對應分析中的行範圍

您必須定義行變數的範圍。指定的最小值和最大值必須是整數。而分數的資料值小數部分會於分析中捨去。而指定範圍外的類別數值則會於分析中被忽略。

圖表 5-3
「定義行範圍」對話方塊




一開始時，所有的類別都沒有受到任何限制，而且都是作用中的。因此，您可以限制某些行類別，必須等於其他行類別，或者將行類別定義成補充性類別。

- **類別必須相等。**類別的分數必須相等。如果所取得的類別次序，並非您想要的，請使用相等限制。至於您所能限制為相等的最大行類別數，算法為作用行類別的總數減 1。若要將多種相等限制，套用到類別組上，此時請透過語法加以執行。例如，您可以使用語法，限制類別 1 等於類別 2、類別 3 等於類別 4。
- **類別為補充的。** 補充類別並不會影響分析，但是卻會出現在由作用類別所定義的空間中。補充類別在定義維度上並沒有任何作用。補充行類別的最大值是行類別的總數值減 2。

# 對應分析模式

您可以利用「模式」對話方塊，來指定維度的數目、測量距離、標準化方法、常態化方法。

圖表 5-4
「模式」對話方塊

**解的維度。**指定維度的數目。一般而言，會選擇最低必要維度，才能說明大部分的變化。最大的維度值，乃是根據分析和相等性限制中，所使用的作用類別數目來決定的。維度的最大值為較小的：

- 作用中列類別的數目，減去限制為相等的列類別數目，再加上受限制之列類別集合的數目。
- 作用中行類別的數目，減去限制為相等的行類別數目，再加上受限制之行類別集合的數目。

**距離量數。**您可以選擇在對應的表格中列及行間的距離測量。請選擇下列其中一個選項：

- **卡方分配。**使用已加權的剖面距離，此處的加權值是指列或行的總數。如果是標準式對應分析中，就必須使用這種測量。
- **歐基里得。** 將列組和行組之間距離的平方和，再開平方，所得出的平方根。

**標準化方法。**請選擇下列其中一個選項：

- **刪除列與行平均數。** 讓列與行都中心化。如果是標準對應分析，就必須使用這種方法。
- **刪除列平均數。**僅列中心化。
- **刪除行平均數。** 僅行中心化。

- **列總和均等化並刪除平均數。** 在將列中心化之前，先將列邊際均等化。
- **行總和均等化並刪除平均數。**在將行中心化之前，先將行邊際均等化。

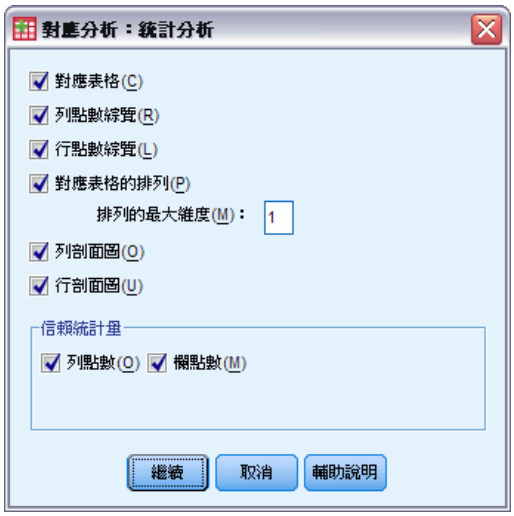
**常態化方法。**請選擇下列其中一個選項：

- **對稱的。**對每個維度而言，列分數指的是：行分數的加權平均數，再除以符合奇異值；而行分數指的是：列分數的加權平均數，再除以符合奇異值。若要檢查兩變數中、所屬類別彼此之間的差異或相似性的話，請使用此方法。
- **主成份分析。**列點數之間的距離，和行點數之間的距離，會跟對應表格中的距離很接近，而這個表格是根據選用的距離測量，所選擇出來的。如果想要檢查任何一個變數，或者兩變數中、所屬類別彼此之間的差異，而不是兩變數之間差異的話，請使用這個方法。
- **列主成分分析。** 列點數之間的距離，會跟對應表格中的距離很接近，而這個表格是根據選用距離，所測量出來的。而且列分數，就是行分數的加權平均數。若要檢查列變數中、所屬類別彼此之間的差異或相似性，請使用此方法。
- **行主成分分析。** 行點數間的距離，會跟對應表格中的距離很接近，而這個表格是根據選用距離，所測量出來的。而且行分數，就是列分數的加權平均數。若要檢查行變數中、所屬類別彼此之間的差異或相似性，請使用此方法。
- **自訂。** 您必須指定一個介於 -1 和 1 之間的值。 如果值是 -1，則會跟行主成份分析相對應。如果值是 1，則會跟列主成分分析互相對應。但是如果值是 0 時，代表它是對稱的。所有其他的數值，將慣性地以不同程度散佈在列與行分數中。當您需要特別製作行列散佈圖時，這個方法非常有用。

# 對應分析統計量

您可以利用「統計量」對話方塊，來指定所產生的數值輸出。

圖表 5-5
「統計量」對話方塊




**對應表格。** 一種交叉表列，其上有輸入變數，以及列與行的邊際總和。

**列點數綜覽。** 對每個列類別而言，它會顯示維度慣性中，分數、大量、慣性所佔有的比例，以及維度在點慣性中，所佔有的比例。

**行點數綜覽。** 對每個行類別而言，它會顯示維度慣性中，分數、大量、慣性所佔有的比例，以及維度在點慣性中，所佔有的比例。

**列剖面圖。** 說明某個行變數中，每個列類別橫跨多少個行變數類別。

**行剖面圖。** 說明某個列變數中，每個行類別橫跨多少個列變數類別。

**對應表格的排列。** 重新組織交叉表列，因此列與行，會根據第一個維度上面的分數，以遞增的順序排列。另外，你可以隨意指定產生的變更表格的最大維度數。產生每一維度從一到指定數字的變更表格。

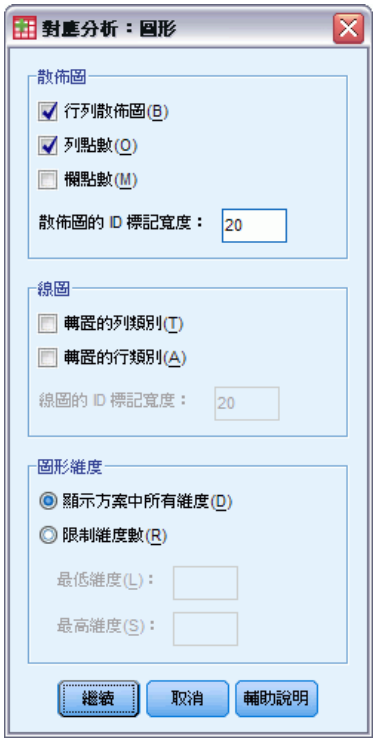
**列點數信賴統計量。** 這個選項包含：所有非補充列點數的標準差和相關值。

**行點數信賴統計量。** 這個選項包含：所有非補充行點數的標準差和相關值。

# 對應分析圖

您可以利用「圖形」對話方塊，來指定要產生哪一種圖形。

圖表 5-6
「圖形」對話方塊




**散佈圖。** 利用維度所有成對圖形，產生一個矩陣。您可以使用的散佈圖，包括：

- **行列散佈圖。** 利用列、行點數的聯合圖形，產生一個矩陣。如果您選取的是主成份分析常態化，則無法使用行列散佈圖。
- **列點數。** 利用列點數圖形，產生一個矩陣。
- **行點數。** 利用行點數圖形，產生一個矩陣。

或者，在標記點時，您可以指定所需使用之數值標記字元個數。但是，此數值必須是正整數，而且小於或等於 20。

**線形圖。** 替選用變數的每個維度，都產生一個圖形。您可以使用的線形圖，包括：

- **轉換後的列類別。** 根據列分數 (與原始列類別值互相對應)，所產生的圖形。
- **轉換後的行類別。** 根據列分數 (與原始行類別值互相對應)，所產生的圖形。

或者，在標記點時，您可以指定所需使用之數值標記字元個數。但是，此數值必須是正整數，而且小於或等於 20。

**繪製維度。** 讓您控制維度在輸出上的顯示方式。

- **顯示解答的所有維度。**解答中的所有維度都會顯示在散佈圖矩陣中。
- **限制維度的數目。**顯示的維度會限制為已繪製的成對資料。如果您限制維度，就必須選擇最低和最高維度來繪製。最低維度的範圍可從 1 到解答中維度的數字減 1，並且對較高的維度進行繪製。最高維度值的範圍可從 2 到解答中維度的數字，並且指出用來繪製維度成對資料的最高維度。 這個規格適用於所有要求的多維度圖。

# CORRESPONDENCE 指令的其他功能

如果您將選用部分貼入語法視窗，並且編輯所產生的 CORRESPONDENCE 指令語法，就可以自訂對應分析。指令語法語言也可以讓您：

- 將表格資料，指定成輸入，而不是使用全部觀察值的資料 (使用 TABLE = ALL 次指令)。
- 指定數值標記的字元個數，以便標記各種散佈圖矩陣、或是行列散佈圖矩陣上面的小點 (使用 PLOT 次指令)。
- 指定數值標記的字元個數，以便標記各種線形圖上面的小點 (使用 PLOT 次指令)。
- 將列與行分數的矩陣，寫入矩陣資料檔中 (使用 OUTFILE 次指令)。
- 將奇異值和分數的信賴統計量矩陣 (變異數和共變異數)，寫入矩陣資料檔中 (使用 OUTFILE 次指令)。
- 指定多組類別，必須相等 (使用 EQUAL 次指令)。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

章節
6

# 多重對應分析

「多重對應分析」會透過指定觀察值（個體）和類別的數值來量化名義（類別）資料，因此相同類別的個體會緊密在一起，而不同類別的個體則會分開。每一個物件會盡量接近包含該物件的類別的類別點。如此一來，即由類別將個體區分為同質性的次組別。將相同類別中的個體歸類到同一次組別的變數即為同質性變數。

**範例。**我們利用「多重對應分析」，以圖形的方式，來顯示工作類別、少數民族、性別三者之間的關係。您可能會發現，若要從人群中，區分少數民族和性別是很容易的，但工作類別就很困難。您可能也會發現，拉丁和非裔美人這兩個少數民族，彼此非常類似。

**統計量與圖形。**個體分數、區別量數、疊代歷程、原始與轉換後變數的相關性、類別量化、敘述統計、個體點圖、行列散佈圖、類別圖、聯合類別圖、轉換圖與區別量數圖。

**資料。** 字串變數值一定都會以遞增的文數字順序，轉換成正整數。使用者定義的遺漏值、系統遺漏值、和小於 1 的值都視為遺漏值；您可以重新編碼、或新增常數到變數中，讓小於 1 的變數成為非遺漏值。

**假設。** 所有的變數都有多重名義尺度水準。資料最少必須含有三個有效觀察值。分析時須使用正整數資料。離散化選項將自動分類分數值的變數，也就是用接近「正常的」分配來分類其數值，並自動將字串變數的值轉換成正整數。您可以指定其他的離散化分法。

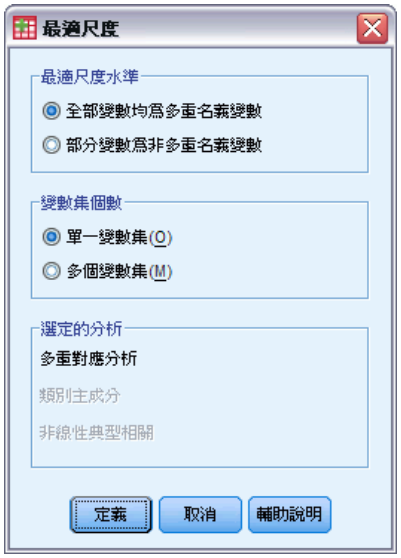
**相關程序。** 如果有兩個變數，「多重對應分析」就會跟「對應分析」很像。如果您確信變數包含次序或數值性質的話，那麼就應使用「類別主成份分析」。如果您想要使用一組變數，那麼應該使用「非線性典型相關分析」。

### 若要取得多重對應分析

▶ 從功能表選擇：
分析(A) > 維度縮減(D) > 最適尺度...

© Copyright SPSS Inc. 1989, 2010

圖表 6-1
「最適尺度」對話方塊




- 選擇「全部變數多重名義量數」。
- 選擇「單一變數集」。
- 按一下「定義」。

圖表 6-2
「多重對應分析」對話方塊

- 選擇兩個以上的分析變數，然後指定解答中的維度數。

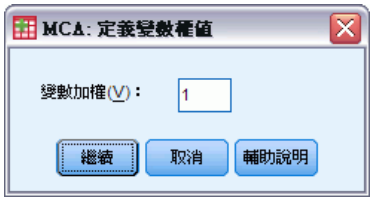
▶ 按一下「確定」。

您可以選擇指定補充變數，該變數可能適合所發現的解答、或標記圖形中的變數。

# 在多重對應分析中定義變數加權

此外，您也可以設定分析變數的權重。

圖表 6-3
「定義變數加權」對話方塊

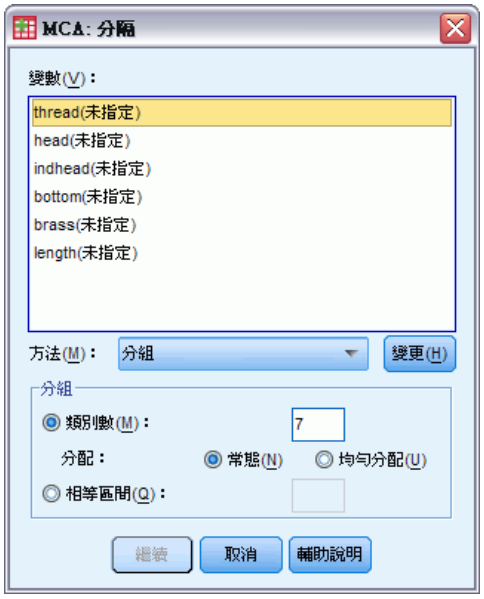



**變數加權。** 您可以選擇定義每個變數的加權。所指定的值必須為正整數。預設值為 1。

# 多重對應分析離散化

您可以在「離散化」對話方塊中，選擇重新編碼變數的方法。除非您另行指定，否則系統將以接近常態分配的方式，將分數值的變數分成七種類別（如果這個數目少於 7，則視變數中不同值的數目決定）。並根據遞增的文數字順序來指定類別指標，將所有字串變數值轉換成正整數。字串變數的離散化適用這些整數。而根據預設值，其他變數則維持不變。離散化的變數稍後將用於分析中。

圖表 6-4
「離散化」對話方塊




**方法。** 選擇分組、分等級、和相乘。

- **分組。** 重新編碼為指定類別的數目、或依間隔來重新編碼。
- **分等級。** 變數將根據觀察值的等級來離散化。
- **相乘。** 變數目前的值將標準化、乘以 10、進位、並新增常數，使最低的類別化數值變成 1。

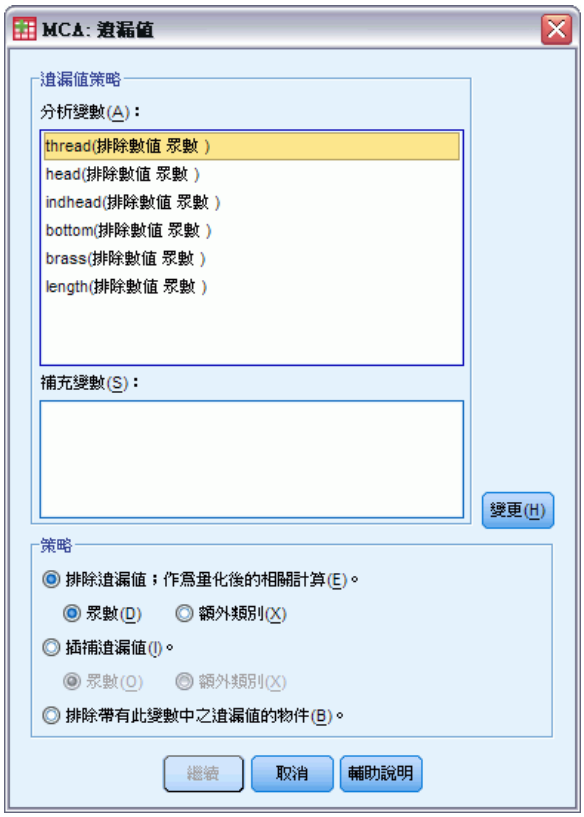
**分組。** 當根據分組離散化變數時，可以使用下列選項：

- **類別的個數。** 指定類別的數目，以及這些類別中的變數值是否應該遵循近似常態、或均勻分配。
- **等距。** 變數將重新編碼成由這些大小相等間隔所定義的類別。而您必須指定間隔的長度。

# 多重對應分析遺漏值

您可以在「遺漏值」對話方塊中，選擇處理分析變數及補充變數中遺漏值的方法。

圖表 6-5
「遺漏值」對話方塊




**遺漏值策略。** 選擇排除遺漏值（被動處理方式）、插補遺漏值（主動處理方式）、或排除有遺漏值的物件（刪除全部遺漏值）。

- **排除遺漏值；作為量化後的相關計算(E)。** 選擇變數中具遺漏值的物件，無利於這個變數的分析。如果對所有變數都施以被動處理方式，則所有變數中具有遺漏值的物件都會被視為補充性。如果「輸出」對話方塊中有指定相關，那麼（在分析後）系統將

以變數中次數最多的類別、或眾數來插補遺漏值，以分析原始變數的相關。若是最適尺度變數的相關，您可以選擇插補的方法。選擇**眾數**可以最適尺度變數的眾數來取代遺漏值。選擇**額外類別**可以額外類別的量化來取代遺漏值。這是表示，在這個變數上有遺漏值的物件，都將視為屬於相同的（額外）類別。

- **插補遺漏值(I)。** 系統會針對在選擇變數上有遺漏值的物件，插補其遺漏值。您可以選擇插補方法。選擇**眾數**可以次數最多的類別來取代遺漏值。當有多個眾數時，程式將使用類別指標最小的眾數。選擇**額外類別**可以額外類別的相同量化來取代遺漏值。這是表示，在這個變數上有遺漏值的物件，都將視為屬於相同的（額外）類別。
- **排除帶有此變數中之遺漏值的物件(B)。** 在選擇變數上有遺漏值的物件，將自分析中排除。這個策略無法用於補充變數。

# 多重對應分析選項

在「選項」對話方塊中，您可以選擇起始組態、指定疊代與收斂條件、選擇常態化方法、選擇標記圖形的方法、以及指定補充的物件。

圖表 6-6
「選項」對話方塊

**補充性的個體。** 請指定物件的觀察值號碼（或物件範圍的第一個和最後一個觀察值號碼），來將它們指定為補充的，然後按一下「**新增**」。繼續一直到您指定了所有的補充性的分析標為止。如果該物件已指定為補充物件，那麼程式會忽略該物件的觀察值加權。

**常態化方法。** 您可以指定五種選項中的其中一種，來常態化個體分數與變數。在指定分析中只能使用一種常態化方法。

- **變數主成份分析。** 這個選項可最佳化變數間的關係。個體空間中的變數座標，為成份負荷量（與主成份的相關，如維度與個體分數）。當您對變數間的相關有興趣時，這個選項非常有用。
- **個體主成份分析。** 這個選項可最佳化個體間的距離。當您對個體間的差異或相似性有興趣時，這個選項非常有用。
- **對稱的。** 如果您對個體與變數間的關係有興趣時，請選擇這個常態化選項。
- **自變數。** 如果您想要分別檢查個體間的距離、以及變數間的相關時，請使用這個常態化選項。
- **自訂。** 您可以指定封閉區間［-1, 1］之間的任何實數。數值 1 等於「個體主成份分析」法、數值 0 等於「對稱的」方法，而數值 -1 等於「變數主成份分析」法。利用指定大於 -1 及小於 1 的數值，您可以將特徵值散佈到個體與變數中。當您需要特別製作行列散佈圖時，這個方法非常有用。

**條件。** 您可以指定最大疊代數，以供計算迴歸時使用。另外，您也可以選擇收斂準則值。如果最後兩個疊代之間的總適合度差異，低於收斂數值，或者已經達到疊代的最大數目的話，迴歸就會停止疊代。

**依據...標記圖形。** 讓您指定是否要在圖形中使用變數與數值標記、或變數名稱與數值。您也可以指定標記的最大長度。

**繪製維度。** 讓您控制維度在輸出上的顯示方式。

- **顯示解答的所有維度。**解答中的所有維度都會顯示在散佈圖矩陣中。
- **限制維度的數目。**顯示的維度會限制為已繪製的成對資料。如果您限制維度，就必須選擇最低和最高維度來繪製。最低維度的範圍可從 1 到解答中維度的數字減 1，並且對較高的維度進行繪製。最高維度值的範圍可從 2 到解答中維度的數字，並且指出用來繪製維度成對資料的最高維度。 這個規格適用於所有要求的多維度圖。

**組態。** 您可以從包含組態座標的檔案中讀取資料。在檔案中的第一個變數，必須包含第一維度的座標，而第二個變數則應該包含第二維的座標，依此類推。

- **初始。** 在指定檔案中的組態，將用來當作分析的開始點。
- **固定。** 在指定檔案中的組態，將用來配合變數。而可以配合的變數，必須選為分析變數，但因為組態是固定的，所以它們將被視為補充變數（所以它們並不需要選為補充變數）。

# 多重對應分析輸出

在「輸出」對話方塊中，您可以建立個體分數、區別量數、疊代歷程、原始與轉換變數的相關、選擇變數的類別變數量化、與選擇變數之敘述統計的表格。

圖表 6-7
「輸出」對話方塊

**個體分數。**顯示個體分數，包括大量、慣性以及貢獻，而且含有以下選項：

- **包括的類別。** 顯示所選擇分析變數的類別指標。
- **標記個體分數根據。** 您可以從指定為標記變數的變數清單中，選擇其中之一來標記物件。

**區別量數。** 顯示每個變數與每個維度的區別量數。

**疊代歷程。** 針對每次疊代，顯示變異歸因於、適配不佳度、以及變異歸因於的增量。

**原始變數的相關性。** 顯示原始變數的相關矩陣，以及該矩陣的特徵值。

**轉換後變數的相關性。** 顯示轉換（最適尺度）變數的相關矩陣，以及該矩陣的特徵值。

**類別量化與貢獻。** 指定所選變數每一個維度的類別量化（座標），包括大量、慣性以及貢獻。

注意：座標與貢獻（包含質量與慣量）會顯示在樞軸表輸出結果中不同的階層中，預設會顯示座標。若要顯示貢獻，請啟動表格（連按兩下），並選取「階層」下拉清單中的「貢獻」。

**敘述統計。** 顯示選擇變數的次數、遺漏值數目、與眾數。

# 多重對應分析儲存

「儲存」對話方塊可以讓您將離散化資訊、個體分數以及轉換值儲存於外部 IBM® SPSS® Statistics 資料檔案或目前階段作業的資料集中。您也可以將轉換值和個體分數儲存於作用中資料集中。

- 可以在目前階段作業期間使用資料集，但是在後續的階段作業中則無法使用，除非您明確地將其儲存為資料檔案。資料集名稱必須符合變數命名規則。
- 檔案名稱或資料集名稱必須針對各個已儲存資料的類型而有所不同。
- 如果您儲存個體分數或將數值轉換為作用中的資料集，您可以指定多重名義維度的個數。

圖表 6-8
「儲存」對話方塊

# 多重對應分析個體圖

使用「個體圖」對話方塊，可以指定所需的圖表類型，以及用來繪製圖表的變數。

圖表 6-9
「個體圖」對話方塊

**個體點。** 顯示個體點的圖。

**個體與重心（行列散佈圖）。**使用變數重心來繪製個體點。

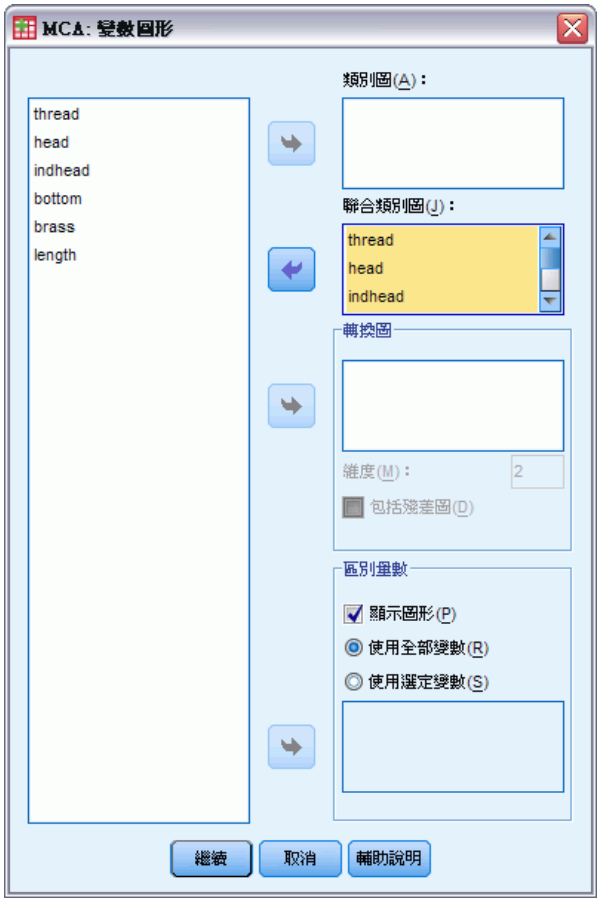
**行列散佈圖變數。** 您可以選擇使用行列散佈圖的所有變數或選擇一個子集。

**標記個體。** 您可以選擇用選定變數的類別（您可以在「選項」對話方塊中選擇類別指標值或數值標記）、或其觀察值號碼來標記個體。如果您選擇**變數**，則每個變數會建立一個圖形。

# 多重對應分析變數圖

使用「變數圖」對話方塊，可以指定所需的圖表類型，以及用來繪製圖表的變數。

圖表 6-10
「變數圖」對話方塊




**類別圖。** 針對所選擇的每個變數，繪製重心座標圖。類別位於特殊類別中個體的重心中。

**聯合類別圖。** 這是每個選擇之變數的單一重心座標圖。

**轉換圖。** 顯示最佳類別變數量化對類別指標的圖形。您可以指定想要的維度數目，每一個維度都會產生一個圖。您也可以選擇顯示每個選定變數的殘差圖。

**區別量數。**產生所選變數之區別量數的單一圖。

# MULTIPLE CORRESPONDENCE 指令的其他功能

如果您將選用部分貼入語法視窗，並且編輯所產生的 MULTIPLE CORRESPONDENCE 指令語法，就可以自訂「多重對應分析」。指令語法語言也可以讓您：

- 在將轉換變數存入作用中的資料集時，指定它的根名稱（利用 SAVE 次指令）。
- 分別指定各圖形中標記的長度（使用 PLOT 次指令）。
- 指定不同的殘插圖變數清單（使用 PLOT 次指令）。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

章節
7

# 多元尺度方法（PROXSCAL）

多元尺度方法試圖尋找在一組物件間相似性量測的結構。可藉著將觀察值指定到概念的低維度空間中特定位置來達成此目的，而空間中的點間距須儘可能符合規定的相似（相異）性。其結果為低維度空間中物件的最小平方表示，在許多例子中，這樣可協助您進一步瞭解您的資料。

**範例。** 多元尺度方法在確定概念上的關係時很有用。例如，當考慮您的產品形象時，您就可以進行一項調查，以便取得說明您的產品與競爭產品的認知相似性（或相似性）的資料集。使用這些相似性及自變數（如價格），您可以嘗試確定哪些變數對人們如何看待這些產品是很重要的，並可隨之調整您的形象。

**統計量與圖形。** 共有疊代歷程、壓力係數測量、壓力係數分解、共同空間座標、最終組態內的物件距離、個別空間加權、個別空間、轉換後的相似性、轉換後的自變數、壓力圖、共同空間散佈圖、個別空間加權散佈圖、個別空間散佈圖、轉換圖、Shepard 殘差圖及自變數轉換圖。

**資料。** 資料可依相似性矩陣或轉換為相似性矩陣的變數形式來提供。這些矩陣可能會格式成單一行或跨行形式，相似性可能會依比例的、間隔的、次序的或尺度水準來處理。

**假設。**至少必須指定三項變數，但維度數不能超過物件個數減 1。如果與多重隨機起點結合在一起，維度縮減就會被省略。如果只有指定一個來源，那所有模式都會相等於單位模式，所以，分析係預設為單位模式。

**相關程序。** 如果您是以數值水準來製作所有變數的尺度的話，就會產生標準的多元尺度分析。

© Copyright SPSS Inc. 1989, 2010

### 取得多元尺度方法

▶ 從功能表選擇：

分析 > 尺度 > 多元尺度方法（PROXSCAL）...

此時「資料格式」對話方塊也隨之開啟。

圖表 7-1
資料格式對話方塊

▶ 指定資料格式：

**資料格式。** 可指定您的資料是由相似性矩陣構成，或是想要建立資料的相似性。

**來源數。** 如果您的資料是相似性，請指定您的相似性量測來源是單一或多重的。

**單一來源。** 如果有一個相似性來源，可指定您的資料集格式是相似性在行的矩陣中或是在單一行中有兩個不同變數，以確定每個相似性的列及行。

- **跨行的近似性矩陣。** 近似性矩陣所跨的行數等於物件的個數。此選項會進入「跨行矩陣的近似性」對話方塊。
- **單一行中的近似性。** 近似性矩陣會收合在單一行或單一變數中必須有另外兩個變數，分別辨識每個儲存格的列與行。此選項會進入「單一行中的近似性」對話方塊。

**多重來源。** 如果有多重相似性來源，請指定資料集格式是相似性在行的堆積矩陣中、相似性在每個行有一個來源的多個行中或是相似性堆積在單一行中。

- **跨行堆疊矩陣中的近似性。** 近似性矩陣分跨多行中，行數等於物件的個數；並一個個堆疊在列上，列數等於物件個數乘以來源個數。此選項會進入「跨行矩陣的近似性」對話方塊。

- **多行中的近似性，每行有一個來源。** 近似性矩陣會收合在多行或多個變數中。必須有另外兩個變數，分別辨識每個儲存格的列與行。此選項會進入「多行中的近似性」對話方塊。
- **堆疊在單一行中的相似性。** 相似性矩陣會收合在單一行或變數中。必須有另外三個變數，分別辨識每個儲存格的列、行和來源。此選項會進入「單一行中的近似性」對話方塊。

▶ 按一下「定義」。

# 近似性在直行的矩陣中

如果您在「資料格式」對話方塊中，選擇單一來源或多重來源的相似性在矩陣資料模式中的話，主要的對話方塊將會顯示如下：

圖表 7-2
「在行的矩陣中的相似性」對話方塊

▶ 選擇三個以上的相似性變數(請確定清單中變數的順序符合相似性行的順序)。

▶ 隨意選擇等於相似性變數個數的加權變數。(確定加權值的順序符合他們所加權的相似性的順序)。

▶ 如果有多重來源，請選擇性地選取一個來源變數。(每個相似性變數中觀察值個數應該等於相似性變數個數乘上來源數)。

此外，您還可以定義多元尺度方法的模式，定出共同空間上的限制，設定收斂條件，指定所要使用的起始組態，以及選擇圖形與輸出。

# 在單一行中的相似性

如果您在「資料格式」對話方塊中，選擇多重來源的多個行模式的話，主要的對話方塊將會顯示如下：

圖表 7-3
「單一行中的相似性」對話方塊

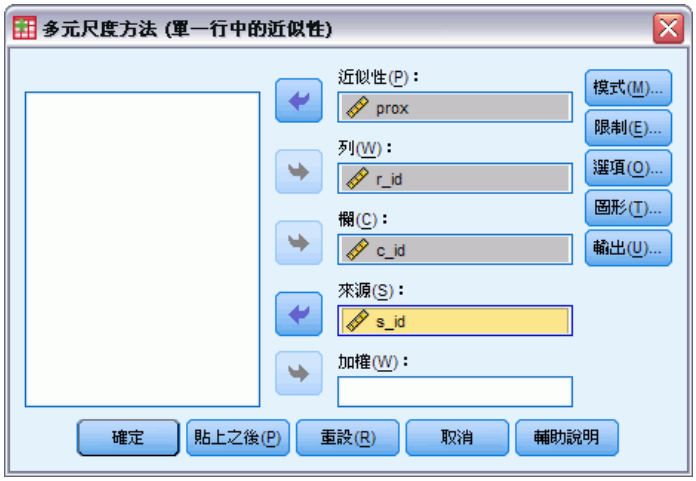
- 選擇兩個以上的相似性變數(每個變數係假設為來自不同來源的相似性矩陣)。
- 選擇一個列變數來定義各個相似性變數中的相似性之列的位置。
- 選擇一個行變數來定義各個相似性變數中相似性之行的位置（未指定列/行的相似性矩陣儲存格會被視為遺漏)。
- 隨意選擇等於相似性變數個數的加權變數。

此外，您還可以定義多元尺度方法的模式，定出共同空間上的限制，設定收斂條件，指定所要使用的起始組態，以及選擇圖形與輸出。

# 單一行中的相似性

如果您在「資料格式」對話方塊中，選擇單一來源或多重來源的單一行模式的話，主要的對話方塊將會顯示如下：

圖表 7-4
「單一行中的相似性」對話方塊




- 選擇一個相似性變數(它係假設為一或多個相似性矩陣)。
- 選擇一個列變數來定義相似性變數中列位置的相似性。
- 選擇一個行變數來定義相似性變數中行位置的相似性。
- 如果有多重來源，請選擇一個來源變數。(對每個來源而言，未指定列/行的相似性矩陣儲存格會被視為遺漏)。
- 隨意地選擇一個加權變數。

  此外，您還可以定義多元尺度方法的模式，定出共同空間上的限制，設定收斂條件，指定所要使用的起始組態，以及選擇圖形與輸出。

# 建立資料的相似性

如果您在「資料格式」對話方塊中，選取建立資料的相似性的話，主要的對話方塊將會顯示如下：

圖表 7-5
「建立資料的相似性」對話方塊

- 如果您建立變數之間距離（請參閱「從資料建立測量法」對話方塊），請至少選擇三個變數，這些變數將用來建立相似性矩陣（如果有多重來源的話，則為多個矩陣）。如果您是建立觀察值之間距離，就只需要一個變數。
- 如果有多重來源，請選擇一個來源變數。
- 隨意地選取一個建立相似性的測量法。

  此外，您還可以定義多元尺度方法的模式，定出共同空間上的限制，設定收斂條件，指定所要使用的起始組態，以及選擇圖形與輸出。

# 從資料建立測量法

圖表 7-6
從資料建立測量法對話方塊

多元尺度方法使用相異性資料來建立尺度法解答。如果您的資料是多變量資料（已測量變數的數值），您就必須建立相異性資料，以便計算多元尺度方法的解答。您也可以指定從資料建立相異性量數的詳細內容。

**量數。**可讓您指定分析所需的相異性量數，從對應您的資料類型的「測量」組中，選擇一個選項，然後從對應那項測量類型的下拉式清單中選擇一種量測。可用的選項有：

- **區間**。共有歐基里得直線距離、歐基里得直線距離平方、Chebychev、區塊、Minkowski 或自訂式。
- **個數**。卡方量數或 Phi 平方量數。
- **二進位**。共有歐基里得直線距離、歐基里得直線距離平方、大小差異、形式差異、變異數或 Lance 與 Williams。

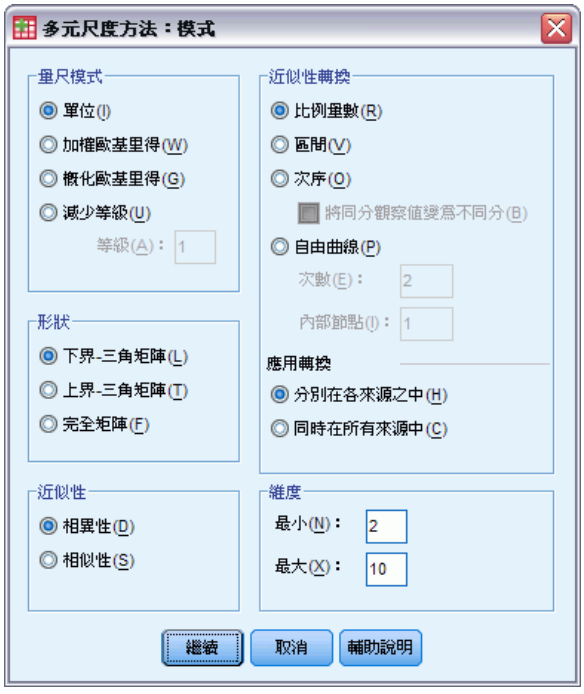
**距離矩陣的計算法。**可讓您選擇分析的單位，選項共有「變數之間」或「觀察值之間」。

**轉換值。**在某些例子中，例如，當變數是以差異甚大之尺度來測量時，您會想在計算相似性（無法應用在二進位資料）之前先將數值標準化。從「標準化」下拉式清單中，選擇一種標準化方法（如果無須標準化，請選擇「無」）。

# 定義多元尺度方法的模式

「模式」對話方塊可讓您指定尺度模式、它的維度數最小值與最大值、相似性矩陣的架構、用在相似性的轉換，以及相似性是個別在每個來源內轉換，或是無條件地在所有來源中轉換。

圖表 7-7
「模式」對話方塊




**尺度模式。** 從下列選項選擇：

- **單位。** 所有來源都有相同的組態。
- **加權歐基里得。** 此模式為個別差異模式。每個來源都有個別空間，其中對共同空間的每個維度依差別而加權。
- **概化歐基里得。** 此模式為個別差異模式。每個來源都有相當於共同空間旋轉後再對維度依差別而加權的個別空間。
- **減少等級。** 此為一種概化的歐基里得模式，可供您指定個別空間的等級。您必須指定大於或等於 1，以及小於維度數最大值的等級。

**形狀。** 可指定相似性應該是從相似性矩陣的下界-三角部分或從上界-三角部分來取得。您可能會指定要使用完全矩陣，那麼在此例中，將會分析上界-三角部分及下界-三角部分的加權總和。無論如何，雖然僅會用到指定的部分，但還是應該指定完整的矩陣（包含對角線）。

**相似性。** 可指定您的相似性矩陣是包括相似性或相異性量測。

**相似性轉換。** 從下列選項選擇：

- **比例量數。** 轉換後的相似性是與原始相似性成比例，但此轉換只容許正值的相似性。
- **間隔。** 轉換後的相似性是與原始相似性成比例，再加上一個截距項，而此截距可確保所有轉換後的相似性都是正數。
- **次序。** 轉換後的相似性與原始相似性的順序一樣，您可以指定同分相似性應該是維持同分或容許變為不同分。
- **自由曲線。** 轉換後的相似性是原始相似性的平滑非遞減成段的多項式轉換。您可以指定多項式的次數及內部節點數。

**應用轉換。** 可指定只在每個來源內彼此比較相似性，或是否無條件地在所有來源中進行比較。

**維度。** 根據預設，解答是在兩個維度 (最小值 = 2，最大值 = 2) 中計算的。只要最小值小於或等於最大值，您就可以選擇一個整數最小值及最大值，範圍從 1 到個體個數減 1。此程序會先依最大維度計算解答，然後再逐步減少維度，直到達到最小維度。

# 多元尺度方法的限制

「限制」對話方塊可供您設定共同空間的限制。

圖表 7-8
限制對話方塊

**共同空間的限制。** 可指定想要的限制類型。

- **無限制。** 不會定限制在共同空間上。
- **某些座標固定。** 第一個選擇的變數含有第一個維度上之個體的座標，第二個變數則對應到第二維度的座標，以此類推。一個遺漏值則表示一個維度的座標是空白的。所選擇變數的個數必須等於所要求的最大維度數。
- **自變數的線性結合。** 共同空間會受限於僅能是所選擇變數的線性結合。

**限制變數。** 可選擇定義共同空間限制的變數。如果您已指定一項線性結合，就可以指定線性變數的間隔、名義、次序的或自由曲線轉換。不論指定何者，每個變數的觀察值個數都必須等於物件個數。

# 多元尺度方法的選項

「選項」對話方塊可供您選擇起始組態類型，指定疊代與收斂條件，以及選擇標準或隨意更新。

圖表 7-9
「選項」對話方塊

**起始組態。**請選擇下列其中一個選項：

- **簡算法。** 物件會在最大維度上，依彼此間相同距離來放置。進行一步疊代以改進高維度組態，隨之進行維度縮減以取得具最大維度數的起始組態，而此最大維度數係於「模式」對話方塊中所指定。
- **Torgerson。** 會以一種典型尺度解答當作起始組態。
- **單一隨機起點。** 會隨機選擇一種組態。
- **多重隨機起點。**會隨機選取一些組態，而且常態化列壓力最低的那個組態將當作起始組態。
- **自訂。** 您可以選擇包含您自己之起始組態座標的變數。選擇的變數個數應該等於指定的最大維度數，以第一個變數對應第一個維度的座標，第二個變數對應第二個維度的座標，並依此類推。而每個變數中的觀察值個數應該等於物件個數。

**疊代準則。** 可指定疊代準則的數值。

- **壓力係數收斂。** 當連續的常態化列壓力值之間差異小於此處所指定的數字（必須介於 0.0 到 1.0 之間）時，演算法就會停止疊代。
- **最小壓力係數。** 當常態化列壓力低於此處所指定的數字（必須介於 0.0 到 1.0 之間）時，演算法就會停止疊代。
- **最大疊代。** 除非上列準則的其中之一已先達成，否則演算法將會執行指定的疊代次數。
- **使用隨意更新。** 隨意更新可以加速演算法的速度；但這些更新無法與單位模式以外的模式或限制一起使用。

# 多元尺度方法的圖形（版本 1）

「圖形」對話方塊可供您指定要產生哪些圖形。如果您有相似性在行中的資料格式，下列「圖性」對話方塊就會顯示出來。對**個別空間加權**、**原始對轉換後的相似性**及**轉換後的相似性對距離**圖而言，您可以指定應該產生圖形的來源。可用的來源清單就是主要對話方塊中的相似性變數清單。

圖表 7-10
圖對話方塊，第 1 版

**壓力。** 會產生常態化列壓力對維度的圖形，只有當最大維度數大於最小維度數時，才會產生這個圖形。

**共同空間。** 會顯示共同空間的座標散佈圖矩陣。

**個別空間。** 對每個來源而言，個別空間的座標會顯示在散佈圖矩陣中。但這只有在「模式」對話方塊中，指定了個別差異模式的其中之一的話，才有可能做得到。

**個別空間加權。** 散佈圖會產生個別空間加權。但這只有在「模式」對話方塊中，指定了個別差異模式的其中之一的話，才有可能做得到。對加權的歐基里得模式而言，權數會列出在每個軸為一個維度的圖形中。對概化的歐基里得模式而言，每個維度會產生一個圖形，包含該項維度的旋轉及加權。減少等級模式會產生與歐基里得模式一樣的圖形，但會縮減個別空間的維度數。

**原始相似性對轉換後的相似性。** 會依原始相似性對轉換後的相似性來產生圖形。

**轉換後的相似性 vs. 距離。** 會依轉換後的相似性對距離來產生圖形。

**轉換後的自變數。** 會產生自變數的轉換圖。

**變數與維度之相關。** 會顯示自變數及共同空間維度之間的相關圖。

# 多元尺度方法的圖形（版本 2）

「圖形」對話方塊可供您指定要產生哪些圖形。如果您的資料格式是行中相似性之外的任何格式，下列「圖形」對話方塊就會顯示出來。對**個別空間加權**、**原始對轉換後的相似性**及**轉換後的相似性對距離**圖而言，您可以指定應該產生圖形的來源。輸入的來源數必須是主要對話方塊中所指定的來源變數值，且其範圍必須可從 1 到來源數。

圖表 7-11
圖對話方塊，第 2 版

# 多元尺度方法的輸出

「輸出」對話方塊可供您控制顯示輸出的數量，並將某些輸出儲存至不同的檔案。

圖表 7-12
「輸出」對話方塊

**顯示。**選擇下列一或多個顯示的項目：

- **共同空間座標。** 會顯示共同空間的座標。
- **個別空間座標。** 只有當模式不是單位模式時，才會顯示個別空間的座標。
- **個別空間加權。** 只有當個別差異模式的其中之一被指定時，才會顯示個別空間加權。依模式而定，空間加權會分解成旋轉加權及維度加權，而這些也都會顯示出來。
- **距離。** 會顯示組態中物件之間的距離。
- **轉換後的相似性。** 會顯示組態中物件之間轉換後的相似性。
- **輸入資料。** 包含原始相似性，而且如果有的話，也會包含資料加權、起始組態，以及自變數個數的固定座標。
- **隨機起點的壓力係數。** 會顯示隨機種子及每個隨機起點的常態化列壓力值。
- **疊代歷程。** 會顯示主要演算法的疊代歷程。
- **多重壓力係數測度。** 會顯示不同的壓力值，表格中包含常態化列壓力、應壓力-I、壓力-II、S 應力、分散情形歸因於 (DAF) 及適用的 Tucker’s 係數。
- **壓力係數分解。** 會顯示最終常態化列壓力的物件及來源分解，包含每個物件平均值及每個來源平均值。
- **轉換後的自變數。** 如果已選擇線性結合限制的話，轉換後的自變數及對應的迴歸加權就會顯示出來。
- **變數與維度之相關。** 如果已選擇線性結合限制的話，自變數及共同空間維度之間的相關就會顯示出來。

**儲存至新檔案。** 您可將共同空間座標、個別空間加權、距離、轉換後的相似性及轉換後的自變數儲存至不同的 IBM® SPSS® Statistics 資料檔。

# PROXSCAL 指令的其他功能

如果您將選取部分貼入語法視窗，並且編輯所產生的 PROXSCAL 指令語法的話，就可以自訂相似性分析的多元尺度方法。指令語法語言讓您也可以：

- 指定轉換及殘差圖的不同變數清單（使用 PLOT 次指令）。
- 指定個別空間加權、轉換及殘差圖的不同來源清單（使用 PLOT 次指令）。
- 指定所要顯示的自變數轉換圖子集（使用 PLOT 次指令）。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

章節
8

# 多維度展開 (PREFSCAL)

「多維度展開」程序會嘗試搜尋一般數值尺度，讓您以目視的方式檢查兩組個體之間的關係。

**範例。** 要求 21 個人將 15 個早餐項目依照自己的喜好分級為 1 到 15，使用「多維度展開」時，您可以有兩個主要方式來決定個案對於早餐項目之間的喜好差別：軟麵包和硬麵包之間，以及有脂肪與無脂肪項目之間。

或者，您已經要求一群駕駛員為 26 種汽車模型進行評等，評等依據為 10 種屬性以及從 1=「非常不正確」，到 6=「非常正確」的 6 點尺度範圍。平均值超過個別值，值會被視為相似性。使用「多維度展開」，您就可以尋找相似模式的集群，以及這些集群最為密切相關的屬性。

**統計量與圖形。**「多維度展開」程序可以產生疊代歷程、壓力係數測度、壓力係數分解、共同空間座標、最終組態內的個體距離、個別空間加權、個別空間、轉換後的相似性、壓力圖、共同空間散佈圖、個別空間加權散佈圖、個別空間散佈圖、轉換圖以及 Shepard 殘差圖。

**資料。** 資料是以矩形相似性矩陣的形式提供。每一個行都會被視為個別的行個體，而相似性矩陣的每一個列都會被視為個別的列個體。當有多個相似性來源時，就會將矩陣加以堆疊。

**假設。** 至少必須指定兩項變數，但是解答中的維度數不能超過個體數減 1 之後的數量。如果只有指定一個來源，那所有模式都會相等於單位模式，所以，分析係預設為單位模式。

### 取得多維度展開

▶ 從功能表選擇：
分析 > 尺度 > 多維度展開 (PREFSCAL)...

© Copyright SPSS Inc. 1989, 2010

圖表 8-1
「多維度展開」主對話方塊

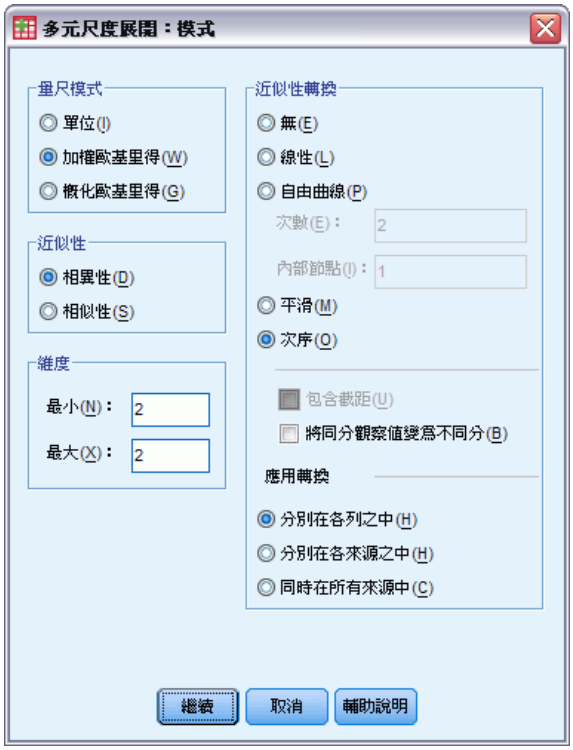
- 選擇兩個以上可以識別矩形相似性矩陣中之行的變數。每一個變數都是表示個別的行個體。
- 或是選擇等於行個體變數的加權變數數量。加權變數的順序必須符合變數所加權之行個體的順序。
- 選擇一個列變數。這個變數的值（或數值標記）是用來標記輸出中的列物件。
- 如果有多重來源，請選擇一個來源變數。資料檔案中的觀察值個數應該等於列個體乘上來源數之後的數量。

此外，您還可以定義多維度展開的模式、為共同空間設定限制、設定收斂條件、指定所要使用的起始組態，以及選擇圖形與輸出。

## 定義多維度展開模式

「模式」對話方塊可以讓您指定尺度模式、它的最小與最大維度數、相似性矩陣的架構、用在相似性上的轉換，以及指定相似性是否可以在列上進行有條件轉換，或是有條件或無條件地在來源中轉換。

圖表 8-2
「模式」對話方塊




**尺度模式。** 從下列選項選擇：

- **單位。** 所有來源都有相同的組態。
- **加權歐基里得。** 此模式為個別差異模式。每個來源都有個別空間，其中對共同空間的每個維度依差別而加權。
- **概化歐基里得。** 此模式為個別差異模式。每個來源都有相當於共同空間旋轉後再對維度依差別而加權的個別空間。

**相似性。** 可指定您的相似性矩陣是包括相似性或相異性量測。

**維度。** 根據預設，解答是在兩個維度（最小值 = 2，最大值 = 2）中計算的。只要最小值小於或等於最大值，您就可以選擇一個整數的最小值及最大值，範圍從 1 到個體個數減 1。此程序會先依最大維度計算解答，然後再逐步減少維度，直到達到最小維度。

**相似性轉換。** 從下列選項選擇：

- **無。** 不轉換任何相似性。您可以選擇「**包含截距**」，這樣相似性就可以由常數項進行轉換。
- **線性。** 轉換後的相似性會與原始相似性成比例，也就是說，轉換函數會估計斜度，而且截距會固定在 0。這也稱為比例轉換。您可以選擇「**包含截距**」，這樣相似性也就可以由常數項進行轉換。這也稱為間隔轉換。
- **自由曲線。** 轉換後的相似性是原始相似性的平滑非遞減成段的多項式轉換。您可以指定多項式的自由度及內部節點數。您可以選擇「**包含截距**」，這樣相似性也就可以由常數項進行轉換。

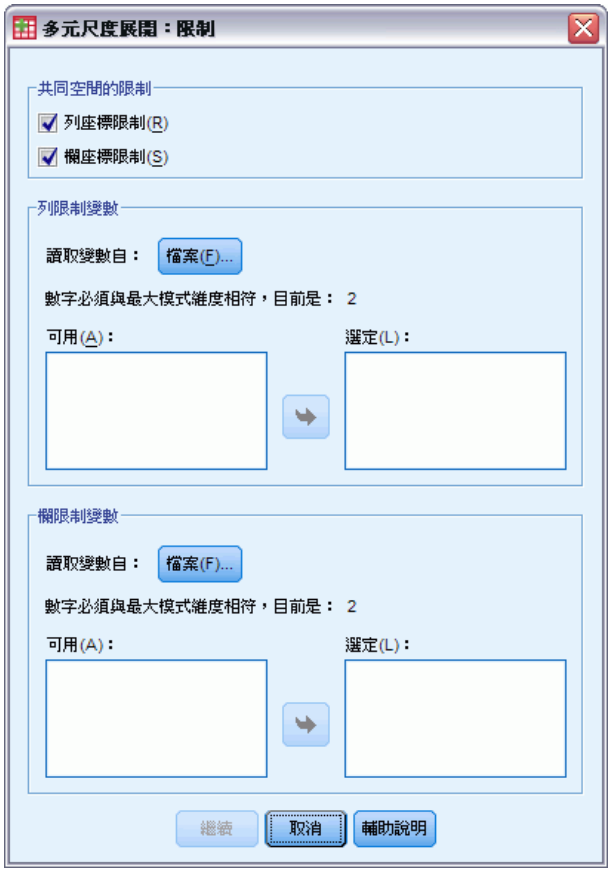
- **平滑。** 轉換後的相似性具有與原始相似性一樣的順序，包括採用了後續值之間差異的限制。結果就是「平滑次序」轉換。您可以指定同分相似性應該是維持同分或容許變為不同分。
- **次序。** 轉換後的相似性與原始相似性的順序一樣，您可以指定同分相似性應該是維持同分或容許變為不同分。

**應用轉換。**指定是否只要互相比較每一列中的相似性，或是只要互相比較各個來源中的相似性，或是無條件地在各個列與來源上比較，也就是說，執行相似性的比較時是每一列或每一個來源地執行，或是一次整個地執行。

# 多維度展開限制

「限制」對話方塊可供您設定共同空間的限制。

圖表 8-3
限制對話方塊




**共同空間的限制。** 您可以選擇修改共同空間中之列和/或行物件的座標。

**列/行限制變數。** 選取包含限制的檔案，以及選擇定義共同空間上之限制的變數。第一個選擇的變數含有第一個維度上之個體的座標，第二個變數則對應到第二維度的座標，以此類推。一個遺漏值則表示一個維度的座標是空白的。所選擇變數的個數必須等於所要求的最大維度數。而每個變數中的觀察值個數應該等於物件個數。

# 多維度展開選項

「選項」對話方塊可供您選取起始組態類型，指定疊代與收斂條件，以及設定壓力的懲罰條款。

圖表 8-4
「選項」對話方塊

**起始組態。**請選擇下列其中一個選項：

- **典型。** 矩形相似性矩陣是用來補充整個對稱 MDS 矩陣的內部區塊（列之間與行之間的數值）。整個矩陣形成之後，會使用典型尺度解答來做為起始組態。內部區塊可以透過插補，使用三角不等式或 Spearman 距離來填入。
- **Ross-Cliff。** Ross-Cliff 起點會使用在雙中心與平方相似性矩陣上之奇異值分解的結果，來做為列與行物件的開始值。
- **對應。** 對應的起點會搭配列與行分數的對稱常態化，使用在相反資料（是相似性而不是相異性）上之對應分析的結果。
- **重心。** 透過使用特徵值分解，將列物件放置在組態中來啟動程序。然後行物件會放置在指定選擇的重心。如果是選擇的數目，請指定 1 至相似性變數的數目之間的正整數。

- **多重隨機起點。** 會為隨機選擇的初始組態計算解答，並且將最低懲罰壓力會顯示為最佳解答。
- **自訂。** 您可以選擇包含您自己之起始組態座標的變數。選擇的變數個數應該等於指定的最大維度數，以第一個變數對應第一個維度的座標，第二個變數對應第二個維度的座標，並依此類推。而每個變數中的觀察值個數應該等於列與行物件之合起來的個數。列與行座標應該是堆疊的，行座標接著列座標。

**疊代準則。** 可指定疊代準則的數值。

- **壓力係數收斂。** 當連續的懲罰應力值之間的相對差異小於此處所指定的數字（必須為非負數）時，演算法就會停止疊代。
- **最小壓力係數。** 當懲罰應力低於這裡所指定的數目（必須為非負數）時，演算法就會停止。
- **最大疊代。** 除非上列準則的其中之一已先達成，否則演算法將會執行此處所指定的疊代次數。

**懲罰條款。** 演算法會根據轉換的相似性的變異係數，嘗試將懲罰應力、相當於 Kruskal’s Stress-I 之產品的適合度測量，以及懲罰條款最小化。這些控制可以讓您設定懲罰條款的範圍與強度。

- **強度。** 強度參數的數值越小，表示懲罰越大。指定 0.0 和 1.0 之間的數值。
- **範圍。** 這個參數會在懲罰啟動時設定。如果設定為 0.0，懲罰則在停用狀態。增加數值，會造成演算法在轉換值中以較大的變異數來搜尋解答。指定一個非負數的數值。

# 多維度展開圖

「圖形」對話方塊可供您指定要產生哪些圖形。

圖表 8-5
「圖形」對話方塊

**圖形。** 您可以使用的圖形如下：

- **多重起點。** 所顯示懲罰應力的堆疊直方圖，會同時顯示應力與懲罰。
- **初始共同空間。** 顯示初始共同空間之座標的散佈圖矩陣。
- **每一個維度的壓力。** 產生懲罰壓力的線形圖與維度範圍的比較。只有當最大維度數大於最小維度數時，才會產生這個圖形。
- **最終共同空間。** 會顯示共同空間的座標散佈圖矩陣。
- **空間加權。** 散佈圖會產生個別空間加權。但這只有在「模式」對話方塊中，指定了個別差異模式的其中之一的話，才有可能做得到。對加權的歐基里得模式而言，所有來源的權數會顯示在每個軸為一個維度的圖形中。對概化的歐基里得模式而言，每個維度會產生一個圖形，包含各個來源之維度的旋轉及加權。

- **個別空間。** 會顯示各個來源之個別空間的座標散佈圖矩陣。但這只有在「模式」對話方塊中，指定了個別差異模式的其中之一的話，才有可能做得到。
- **轉換圖。** 會依原始相似性對轉換後的相似性來產生散佈圖。依據如何套用轉換，會將不同的顏色指定給各列或各個來源。無條件轉換會產生單一顏色。
- **Shepard 圖。** 原始相似性與轉換後的相似性和距離兩者的比較。利用點來指出距離，而利用線來指出轉換後的相似性。依據如何套用轉換，不同的直線是為各列與各個來源而產生。無條件轉換會產生直線。
- **適合度的散佈圖。** 會顯示轉換後之相似性的散佈圖與距離的比較。如果指定多個來源時，會將不同的顏色會指定給各個來源。
- **殘差圖。** 會顯示轉換後之相似性的散佈圖與殘差（轉換後的相似性減距離）的比較。如果指定多個來源時，會將不同的顏色會指定給各個來源。

**列物件樣式。** 這些會讓您進一步控制圖中的列物件。選擇性的顏色變數之數值可用來循環切換所有顏色。選擇性的標記變數之數值可用來循環切換所有可能的標記。

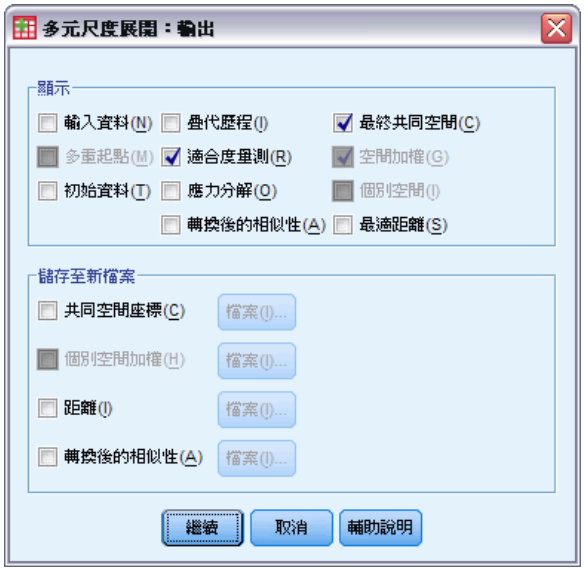
**來源圖。**如果是個別空間、適合度的散佈圖和殘差圖—而且如果轉換是由來源所套用，則對於轉換圖和 Shepard 圖—您可以指定要為其產生圖的來源。輸入的來源數必須是主要對話方塊中所指定的來源變數值，且其範圍可從 1 到來源數。

**列圖。** 如果轉換是由列所套用的話，則對於轉換圖與 Shepard 圖，您就可以指定要為其產生圖形的列。輸入的列號碼範圍必須從 1 到列數。

# 多維度展開輸出

「輸出」對話方塊可供您控制顯示輸出的數量，並將某些輸出儲存至不同的檔案。

圖表 8-6
「輸出」對話方塊




**顯示。** 選擇下列一或多個顯示的項目：

- **輸入資料。** 包括原始相似性，而且如果有的話，也會包含資料加權、起始組態以及固定座標。
- **多重起點。** 會顯示亂數種子及每個隨機起點的懲罰壓力值。
- **起始資料。** 顯示初始共同空間的座標。
- **疊代歷程。** 會顯示主要演算法的疊代歷程。
- **適合量數。** 顯示不同的量數。表格含有幾項適合度、不適合度、相關、變異以及非退化量數。
- **壓力係數分解。** 會顯示物件、行以及懲罰壓力的來源分解，包括列、行以及來源平均數與標準差。
- **轉換後的相似性。** 會顯示轉換後的相似性。
- **最終共同空間。** 會顯示共同空間的座標。
- **空間加權。** 會顯示個別空間加權。只有指定其中一個個別差異模式時，才能使用這個選項。依模式而定，空間加權會分解成旋轉加權及維度加權，而這些也都會顯示出來。
- **個別空間。** 會顯示個別空間的座標。只有指定其中一個個別差異模式時，才能使用這個選項。
- **適合距離。** 會顯示組態中物件之間的距離。

**儲存至新檔案。** 您可將共同空間座標、個別空間加權、距離、轉換後的相似性儲存至不同的 IBM® SPSS® Statistics 資料檔。

# PREFSCAL 指令的其他功能

如果您將選取部分貼入語法視窗，並且編輯所產生的 PREFSCAL 指令語法的話，就可以自訂相似性分析的「多維度展開」。指令語法語言讓您也可以：

- 當多個來源為可用時（使用 PLOT 次指令），為「個別」空間指定多個來源清單、適合度的散佈圖，以及殘差圖—矩陣條件式轉換的情況下，針對轉換圖和 Shepard 圖時。
- 在列條式件轉換的情況下，為轉換圖和 Shepard 圖指定多列的清單（使用 PLOT 次指令）。
- 指定列的數目而不是列的 ID 變數（使用 INPUT 次指令）。
- 指定來源的數目而不是來源的 ID 變數（使用 INPUT 次指令）。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

# 部 II:
# 範例

章節
9

# 類別迴歸

最適尺度類別迴歸的目的是描述回應值及一組預側值之間的關係。而您可以量化這個關係，來預測各種預測值組合的回應值。

本章將舉兩個範例來說明最適尺度迴歸的分析。第一個範例將使用一個小型資料集來說明基本概念。第二個範例將在實際範例中使用較大的變數集及觀察值。

## 範例：地毯清潔機資料

在一個普遍的範例中(Green 和 Wind, 1973)，計劃銷售全新地毯清潔機的公司想要檢驗影響消費者偏好的五個因子，包括—包裝設計、品牌名稱、價格、「優秀家用品」獎章及退費保證。包裝設計有三個因子水準，每個水準中的清潔刷位置都不相同；三個品牌名稱 (K2R、Glory、及 Bissell)；三個價格水準；且最後兩個因子各有兩個水準 (無論無或有)。下表顯示地毯清潔機研究中所用的變數、其變數標記及數值。

**表格 9-1**
**地毯清潔機研究的說明變數**

| 變數名稱 | 變數標記(L) | 數值標記 |
|---|---|---|
| 包裝 | 包裝設計 | A*, B*, C* |
| 品牌 | 品牌名稱 | K2R、Glory、Bissell |
| 價格 | Price | $1.19, $1.39, $1.59 |
| 獎章 | 「優秀家用品」獎章 | 無、有 |
| 費用 | 退費保證 | 無、有 |

十名消費者將這些因子所定義的 22 種組合分級。「偏好」變數包含每個組合平均排名的等級。排名數值較小者會對應高偏好程度。這個變數反映每個組合偏好的整體量數。您可以使用類別迴歸，探索這五個因子與偏好的相關性。本資料集存放在 carpet.sav。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

### 標準線性迴歸分析

▶ 若要產生標準線性迴歸輸出，在功能表上選擇：
**分析(A) > 迴歸 > 線性...**

注意：本功能需要 Statistics Base 選項。

© Copyright SPSS Inc. 1989, 2010

圖表 9-1
「線性迴歸」對話方塊

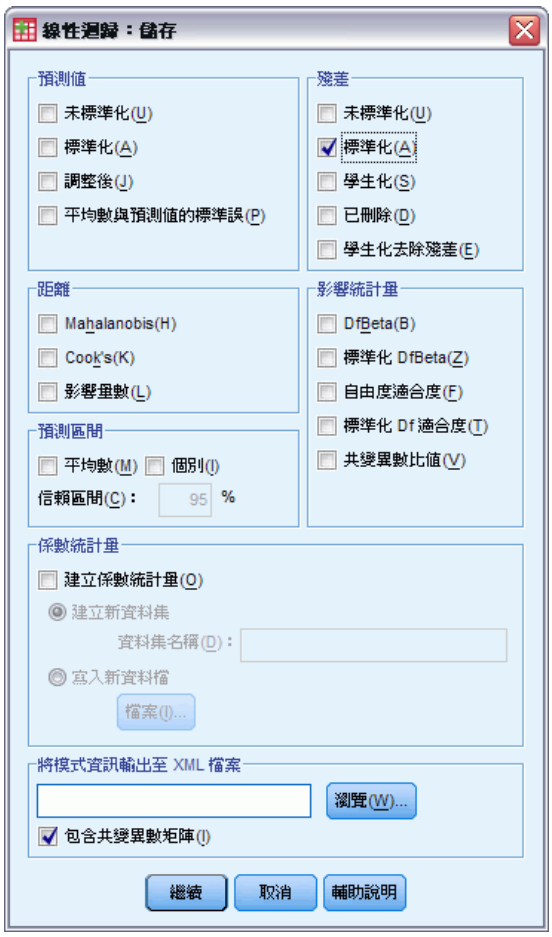
- 選取「偏好」作為依變數。
- 選取「退費保證」的「包裝設計」作為自變數。
- 按一下「圖形」。

圖表 9-2
「圖形」對話方塊

- 選取 *ZRESID 作為 y 軸的變數。
- 選取 *ZPRED 作為 x 軸的變數。
- 按一下「繼續」。

▶ 在「線性迴歸」對話方塊中，按一下「儲存」。

圖表 9-3
「儲存」對話方塊




▶ 選取「殘差」組別中的「標準化」。

▶ 按一下「繼續」。

▶ 在「線性迴歸」對話方塊中，按一下「確定」。

## 模式摘要

圖表 9-4
標準線性迴歸的模式摘要

| 模式 | R | R 平方 | 調過後的 R 平方 | 估計的標準誤 |
|---|---|---|---|---|
| 1 | .841[a] | .707 | .615 | 3.99810 |

a. 預測變數:(常數), 退費保證, 價格級別, 好家務封印, 品名, 套裝設計

線性迴歸是在這個問題中描述關係的標準方式。測量迴歸模式是否完全配適資料最常見的方式是 $R^2$ 值。這個統計量代表預測值的加權組合說明了回應值內的變異數個數。若 $R^2$ 值愈接近 1，模式就愈配適。五個預測值的迴歸「偏好」結果為 0.707 的 $R^2$ 值，顯示線性迴歸中的預測值變數說明了偏好等級內約 71% 的變異數。

## 係數(0)

標準化係數顯示在以下表格中。係數符號指出預測值增加時，無論預測回應增加或減少，其他所有的預測值皆保持不變。 至於類別資料，類別編碼將決定預測值增加時所代表的意義。例如，增加「退費保證」、「包裝設計」，或「優秀家用品獎章」等因子，預測偏好的排名會往前進。「退費保證」的編碼如果是 1，則代表沒有退費保證，如果是 2，則代表有退費保證。「退費保證」增加會與新增退費保證對應。因此，新增退費保證會使預測偏好排名往前進，而與預測偏好增加對應。

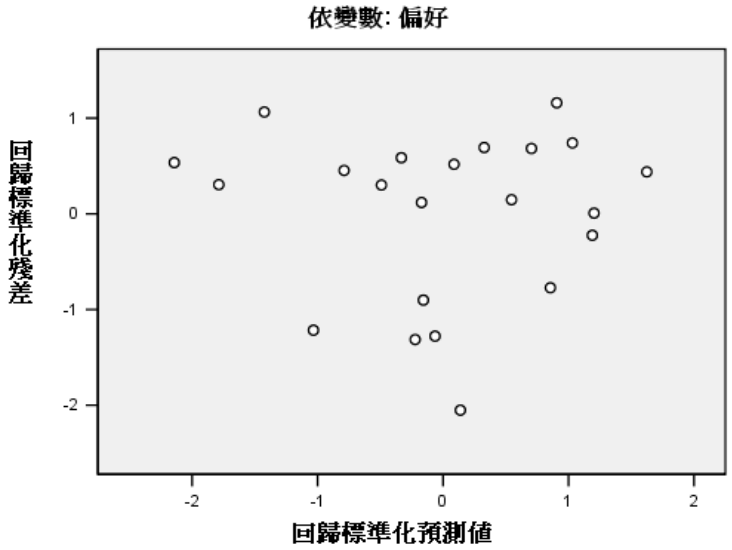
圖表 9-5
迴歸係數

| 模式 | | 未標準化係數 | | 標準化係數 | 未標準化係數 | |
|---|---|---|---|---|---|---|
| | | B 之估計值 | 標準誤差 | Beta 分配 | t | 顯著性 |
| 1 | (常數) | 22.529 | 5.177 | | 4.352 | .000 |
| | 套裝設計 | -4.159 | 1.036 | -.560 | -4.015 | .001 |
| | 品名 | .429 | 1.054 | .056 | .407 | .689 |
| | 價格級別 | 2.703 | 1.009 | .366 | 2.681 | .016 |
| | 好家務封印 | -4.314 | 1.780 | -.330 | -2.423 | .028 |
| | 退費保證 | -2.779 | 1.921 | -.197 | -1.447 | .167 |

係數值會反映預測偏好排名中的變更數。使用標準化係數時，係數解讀將以變數的標準差為基礎。每個係數會指出因預測值內一個標準差變更而導致預測回應變更的標準差數量，其他所有預測值皆維持不變。例如，「品牌名稱」的標準差變更會使預測偏好的標準差提升 0.056。「偏好」的標準差為 6.44，因此「偏好」會增加 $0.056 \times 6.44 = 0.361$。「包裝設計」變更會使預測偏好改變最大。

## 殘差散佈圖

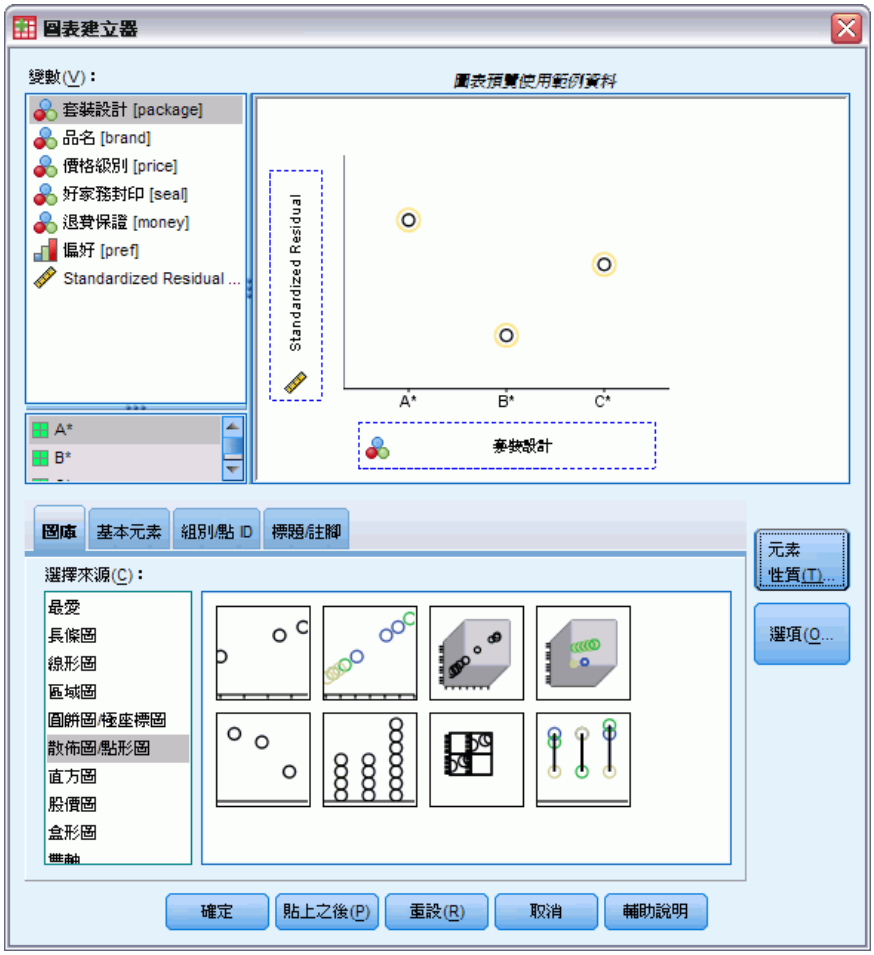
圖表 9-6
殘差對預測值




繪製標準化殘差，將其比對標準化預測值。如果模式配適，應該不會出現任何樣式。這時您可以看到一個 U 型圖，其中低及高標準化預測值都有正殘差。接近 0 的標準化預測值傾向有負殘差。

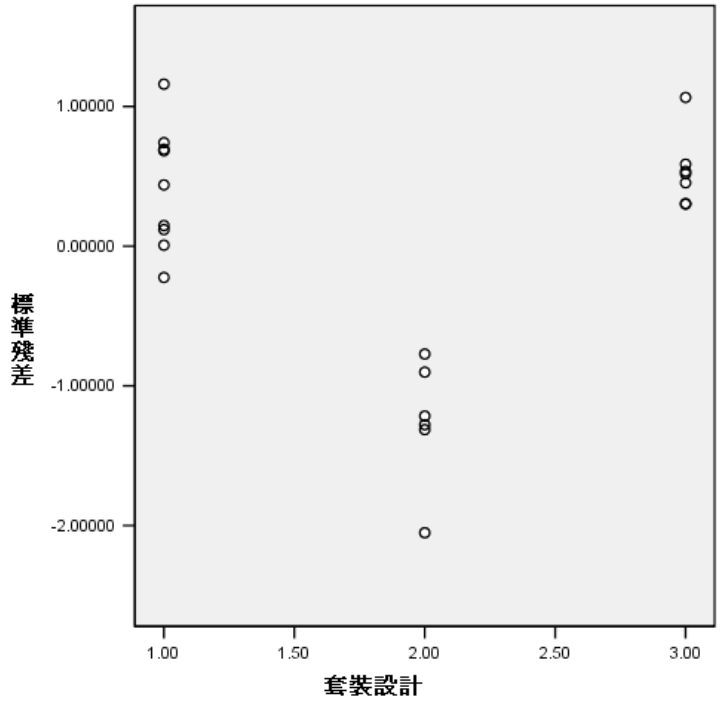
▶ 若要依「包裝設計」預測值而產生散佈圖，從功能表中選擇：
**統計圖(G) > 圖表建立器(C)...**

圖表 9-7
圖表建立器




- 選取「**散佈圖/點形圖**」圖庫並選擇「簡單散佈圖」。
- 選取「標準化殘差」作為 y 軸變數及「包裝設計」作為 x 軸變數。
- 按一下「**確定**」。

圖表 9-8
殘差對包裝設計




與包裝相比，標準化殘差圖內的 U 型圖較明顯。「設計 B」* 的每個殘差都是負數，而另外兩個設計中只有一個殘差是負數。由於線性迴歸模式符合每個變數中的一個參數，因此標準方式無法擷取這個關係。

## 類別迴歸分析

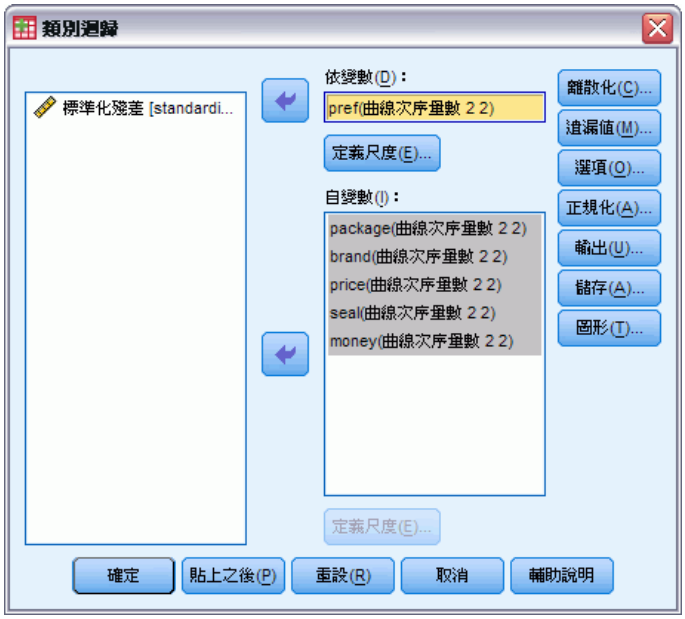
「偏好」及「包裝設計」之間變數的類別性質及非線性關係表示最適分數的迴歸執行結果可能比標準迴歸更佳。U 型的殘差圖指出應該使用「包裝設計」的名義處理。其他所有預測值都將以數值尺度水準處理。

回應變數為特殊考量的根據。您想要預測「偏好」的數值。因此，必須儘可能在量化中復原這個類別的性質。利用次序或名義尺度水準會忽略回應類別之間的差異。但是，以線性方式轉換回應類別會保存類別差異。因此，一般而言以數值方式調整回應值較受歡迎，也將用於此處。

### 執行分析

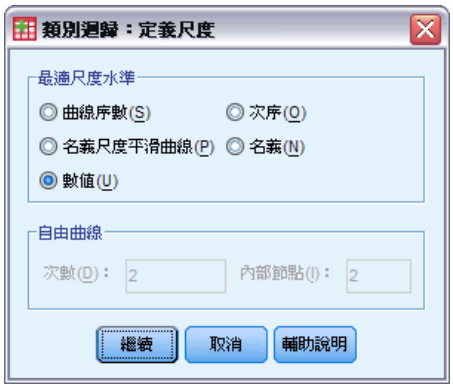
▶ 若要執行「類別迴歸」分析，請從功能表選擇：
**分析(A) > 迴歸 > 最適尺度 (CATREG)...**

圖表 9-9
「類別迴歸」對話方塊



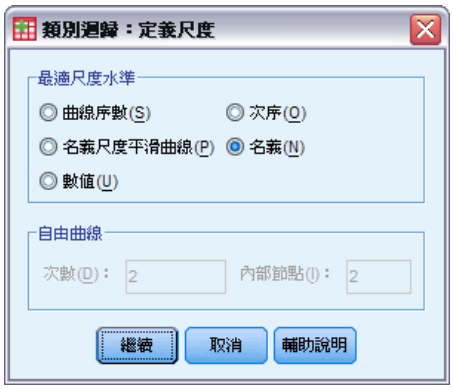

- 選取「偏好」作為依變數。
- 選取「退費保證」的「包裝設計」作為自變數。
- 選取「偏好」並按一下「**定義尺度**」。

圖表 9-10
「定義尺度」對話方塊



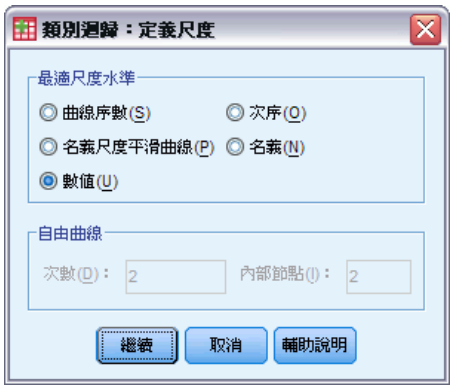

- 選取「**數值**」作為最適尺度水準。
- 按一下「繼續」。
- 選取「包裝設計」，並按一下「類別迴歸」對話方塊中的「**定義尺度**」。

圖表 9-11
「定義尺度」對話方塊




- 選取「名義」作為最適尺度水準。
- 按一下「繼續」。
- 透過「退費保證」選取「品牌名稱」，並按一下「類別迴歸」對話方塊中的「定義尺度」。

圖表 9-12
「定義尺度」對話方塊




- 選取「數值」作為最適尺度水準。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「輸出」。

圖表 9-13
「輸出」對話方塊

- 選取「原始變數的相關性」及「轉換後變數的相關性」。
- 取消選取 ANOVA。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「儲存」。

圖表 9-14
「儲存」對話方塊

- 選取「將殘差儲存到作用中資料集」。
- 在「轉換變數」組別中，選取「將殘差儲存到作用中資料集」。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「圖形」。

圖表 9-15
「圖形」對話方塊

- 選擇建立「包裝設計」及「價格」的轉換圖。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「確定」。

## 相關性

對於識別迴歸中的多重共線性而言，預測值之間的相關性是很有用的。高度相關的變數將產生不穩定的迴歸估計值。但是，由於其高度相關，從模式中略過其中一個僅會對預測值造成些微的影響。可由遺漏變數說明之回應值內的變異數仍然由其餘相關變數來說明。但是，零階相關對偏離值很敏感，且因預測值及其他預測值組合之間的高度相關而無法識別多重共線性。

圖表 9-16
原始預測值的相關性

| | Package design | Brand name | Price | Good Housekeeping seal | Money-back guarantee |
|---|---|---|---|---|---|
| Package design | 1.000 | -.189 | -.126 | .081 | .066 |
| Brand name | -.189 | 1.000 | .065 | -.042 | -.034 |
| Price | -.126 | .065 | 1.000 | .000 | .000 |
| Good Housekeeping seal | .081 | -.042 | .000 | 1.000 | -.039 |
| Money-back guarantee | .066 | -.034 | .000 | -.039 | 1.000 |
| 維度 | 1 | 2 | 3 | 4 | 5 |
| 適合度 | 1.291 | 1.038 | .980 | .905 | .785 |

圖表 9-17
轉換後預測值的相關性

| | Package design | Brand name | Price | Good Housekeeping seal | Money-back guarantee |
|---|---|---|---|---|---|
| Package design | 1.000 | -.156 | -.089 | .032 | .102 |
| Brand name | -.156 | 1.000 | .065 | -.042 | -.034 |
| Price | -.089 | .065 | 1.000 | .000 | .000 |
| Good Housekeeping seal | .032 | -.042 | .000 | 1.000 | -.039 |
| Money-back guarantee | .102 | -.034 | .000 | -.039 | 1.000 |
| 維度 | 1 | 2 | 3 | 4 | 5 |
| 適合度 | 1.248 | 1.043 | .983 | .905 | .821 |

顯示未轉換及轉換後預測值的相關性。所有數值都接近 0，顯示個別變數之間的多重共線性不是問題。

請注意唯一變更的相關性包括「包裝設計」。由於其他所有預測值皆以數值方式處理，因此保留這些變數的類別及類別順序的差異。因此，相關性無法變更。

## 模式適合度及係數

「類別迴歸」程序會產生 0.948 的 $R^2$ 值，顯示最佳轉換後預測值的迴歸說明了轉換偏好排名中 95% 的變異。轉換預測值可改善標準方式的配適度。

圖表 9-18
類別迴歸的模式摘要

| 多 R | R 方 | 調整 R 方 |
|---|---|---|
| .974 | .948 | .927 |

自變數: Preference
預測變數: Package design Brand name Price Good Housekeeping seal Money-back guarantee

下表顯示標準化迴歸係數。類別迴歸會將變數標準化，所以只報告標準化係數。這些數值會除以其對應標準誤，產生每個變數的 F 檢定。但是，每個變數的檢定都視模式中其他存在的預測值而定。換句話說，檢定將判斷模式中略過一個預測變數後，其餘所有預測值是否會顯著降低模式中的預測能力。不應使用這些數值一次略過後續模式中的多個變數。此外，改變最小平方會最佳化量化，這代表這些檢定必須謹慎解讀。

圖表 9-19
轉換後預測值的標準化係數

| | 標準系數 | | df | F | 顯著性 |
|---|---|---|---|---|---|
| | Beta | 標準誤 | | | |
| Package design | -.748 | .060 | 2 | 155.289 | .000 |
| Brand name | .045 | .060 | 1 | .578 | .459 |
| Price | .371 | .059 | 1 | 39.312 | .000 |
| Good Housekeeping seal | -.350 | .059 | 1 | 35.299 | .000 |
| Money-back guarantee | -.159 | .059 | 1 | 7.175 | .017 |

自變數: Preference

最大係數會落在「包裝設計」。「包裝設計」的標準差提升會使預測偏好排名中的標準差降低 0.748。但是，「包裝設計」是以名義方式處理，所以量化的提升不需要相應提升原始類別代碼。

標準化係數通常解讀為用來反映每個預測值的重要性。但是，迴歸係數無法完全描述預測值的影響或預測值之間的關係。其他統計量必須與標準化係數一起使用，以便充分探索預測值的作用。

## 相關性及重要性

若要解讀預測值對迴歸的貢獻，只有檢閱迴歸係數是不夠的。此外，應該檢閱相關、偏相關和部分相關。下表包括每個變數的相關量數。

零階相關是轉換後預測值及轉換後回應值之間的相關性。對這項資料而言，最高的相關性落在「包裝設計」。但是，如果您可以解釋預測值或回應值內的部分差異，您將可利用更好的方式呈現預測值的表現。

**圖表 9-20**
**零階、部分、和偏相關（轉換變數）。**

| | 相關 | | | | 容差 | |
|---|---|---|---|---|---|---|
| | 零排序 | 淨排序 | 部分排序 | 重要性 | 轉換后 | 轉換前 |
| Package design | -.816 | -.955 | -.733 | .644 | .959 | .942 |
| Brand name | .206 | .193 | .045 | .010 | .971 | .961 |
| Price | .440 | .851 | .369 | .172 | .989 | .982 |
| Good Housekeeping seal | -.370 | -.838 | -.349 | .137 | .996 | .991 |
| Money-back guarantee | -.223 | -.569 | -.158 | .037 | .987 | .993 |

自變數: Preference

模式中的其他變數可以混淆某個預測值在預測回應值的結果。偏相關係數可自預測值及回應值中移除其他預測值的線性作用。此量數就是在其他預測值上迴歸得來的殘差與回應值之間的相關性。 偏相關平方對應移除其他變數後，說明的變異數相對於剩餘的回應值殘差變異比例。例如，「包裝設計」的偏相關為 -0.955。移除其他變數的作用後，「包裝設計」會說明 $(-0.955)^2 = 0.91$ = 偏好排名中 91% 的變異數。如果移除其他變數，「價格」及「優秀家用品」獎章也都會說明大部分的變異數。

自回應值及預測值內移除變數作用的另一個作法是只從預測值內移除作用。回應值及在另一個預測值上迴歸預測值所得的殘差之間的相關性為部分相關。這個數值的平方會產生已說明比例量數，代表變異數相對於回應值整體變異數的比例。如果您移除「品牌名稱」、「優秀家用品」獎章、「退費保證」及「價格」的效應，將上述自「包裝設計」中移除，則「包裝設計」的其餘部分會說明 $(-0.733)^2 = 0.54$ = 偏好排名中 54% 的變異數。

## 重要性

除了迴歸係數及相關性之外，Pratt' 的相對重要性量數(Pratt, 1987)可協助解讀預測值對迴歸的貢獻。與其他重要性相對的大型個別重要性會對應到迴歸重要的預測值。此外，如果一個變數的係數與重要預測值大小類似，則這個變數的低重要性會提示抑制變數的存在。

相較於迴歸係數，這個量數以相加方式定義預測值的重要性—，也就是一組預測值的重要性為預測值個別重要性的總和。Pratt’ 的量數等於迴歸係數與預測值零階相關的乘積。這些乘積加至 $R^2$ 值，因此除以 $R^2$ 值，產生總和為 1。例如，「產品包裝」及「品牌名稱」的預測值組合重要性為 0.654。最重要的組合會對應到「包裝設計」，其中「包裝設計」、「價格」及「優秀家用品」獎章佔這組預測值重要性的 95%。

### 多重共線性

預測值之間的高度相關將大幅降低迴歸模式’的穩定性。相關的預測值會導致不穩定的參數估計值。允差會反映自變數彼此之間的線性相關程度。這個量數是變數的變異未由方程式中其他自變數說明的比例。如果其他預測值可以說明某個預測值’的多數變異，則模式中不需要這個預測值。 接近 1 的允差值指出這個變數無法由其他預測值充分預測。相反的，允差很小的變數對模式的貢獻很少，而且會造成計算上的問題。此外，Pratt’ 的重要性量數的眾多負數值會指出多重共線性。

所有的允差量數都很高。沒有任何預測值可由其他預測值充分預測，且沒有出現多重共線性。

## 轉換圖

繪製原始類別數值，與它們的對應量化比對，可以顯示量化清單中可能沒注意到的趨勢。這種圖形一般稱為轉換圖。應該注意取得類似量化的類別。這些類別會影響相同類型的預測回應值。但是，轉換類別會指定圖形的基本外觀。

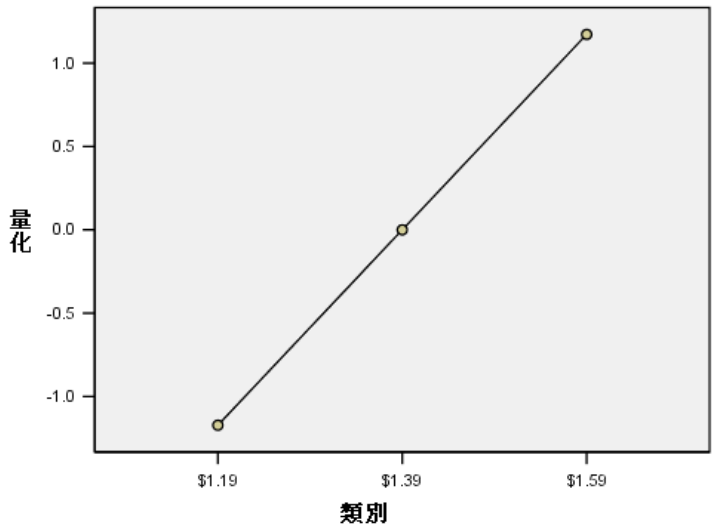
視為數值的變數會產生量化及原始類別之間的線性關係，對應到轉換圖中的直線。原始類別之間的順序及差異保留於量化內。

視為次序之變數的量化順序會對應到原始類別的順序。但是，類別之間的差異並不保留。如此一來，轉換圖呈現非遞減圖形但不須是直線。 如果連續類別對應類似的量化，類別區分可能是不必要的，且類別可以合併。這些類別會導致轉換圖中的高原圖形。但是，這個樣式也可能來自於在應該視為名義的變數上套用次序結構。如果之後將這個變數視為名義仍顯示相同樣式，則須合併類別。此外，如果視為次序之變數量化沿著直線產生，數值轉換可能會更適合。

對於視為名義的變數，沿著水平軸的類別次序會對應到用於代表類別的代碼順序。未建立類別順序或類別間距離的解讀。圖形可以假設任何非線性或線性形式。如果出現上升趨勢，應該嘗試次序處理。如果名義轉換圖顯示線性趨勢，數值轉換可能會更適合。

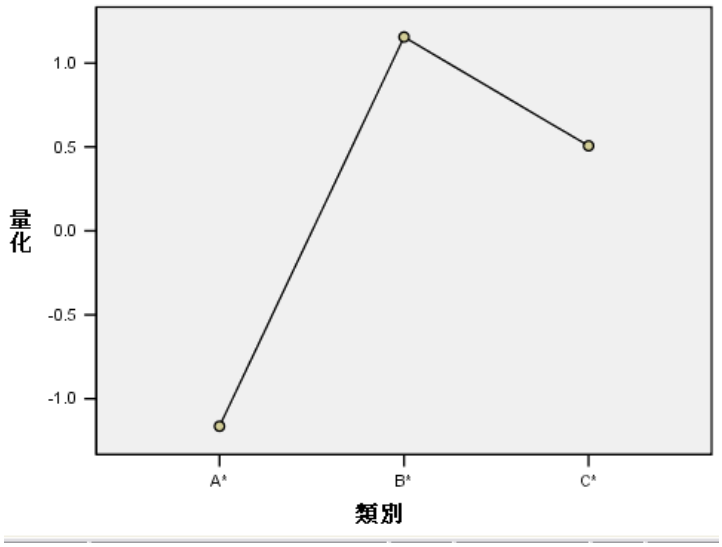
以下顯示「價格」的轉換圖，其中「價格」被視為數值。請注意，類別次序沿著直線對應到原始類別的次序。此外，$1.19 及 $1.39 (－1.173 及 0) 的量化之間的差異等於 $1.39 及 $1.59 (0 及 1.173) 的量化之間的差異。類別 1 和 3 與類別 2 距離相等的事實會保留於量化中。

圖表 9-21
價格轉換圖（數值）




「包裝設計」的名義轉換會產生以下轉換圖形。請注意第二類別擁有最大量化的不同非線性形狀。就迴歸而言，第二類別會使預測偏好排名往前進，而第一及第三類別的作用卻相反。

圖表 9-22
「包裝設計」（名義）的轉換圖

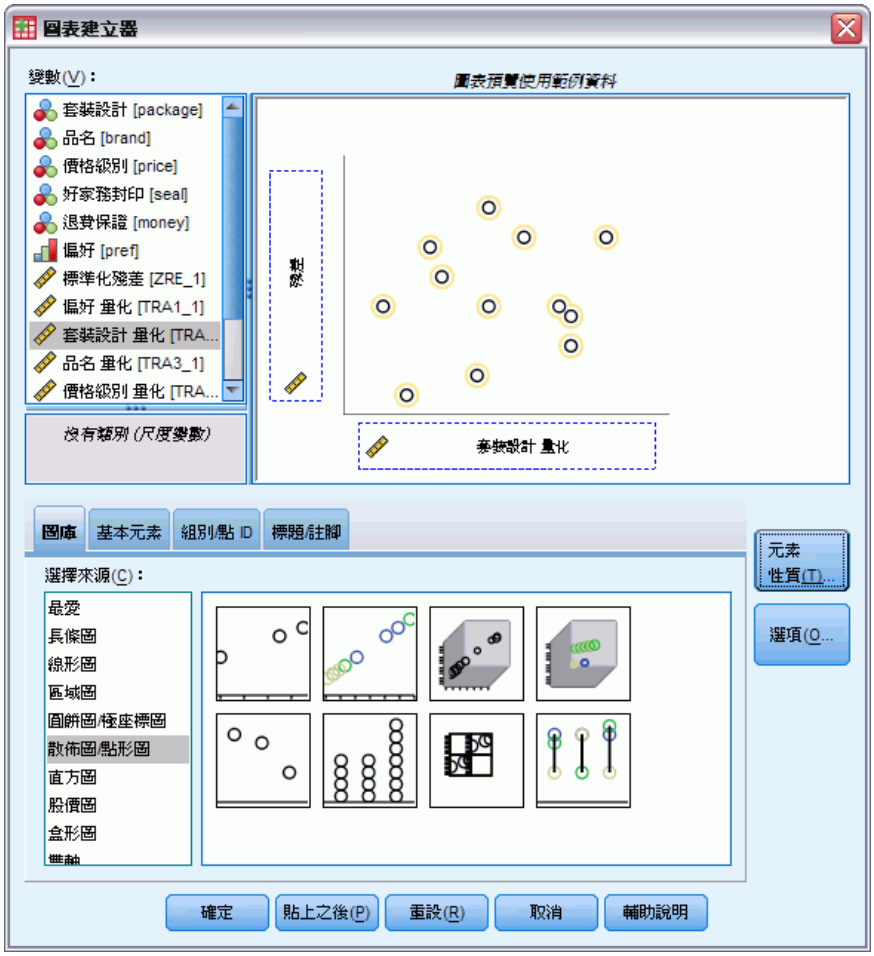



## 殘差分析

利用您儲存在作用中資料集的轉換資料及殘差，可讓您藉由「包裝設計」的轉換值來建立預測值的散佈圖。

若要取得這種散佈圖，叫回「圖表建立器」並按一下「**重設**」，清除您之前的選擇並還原預設選項。

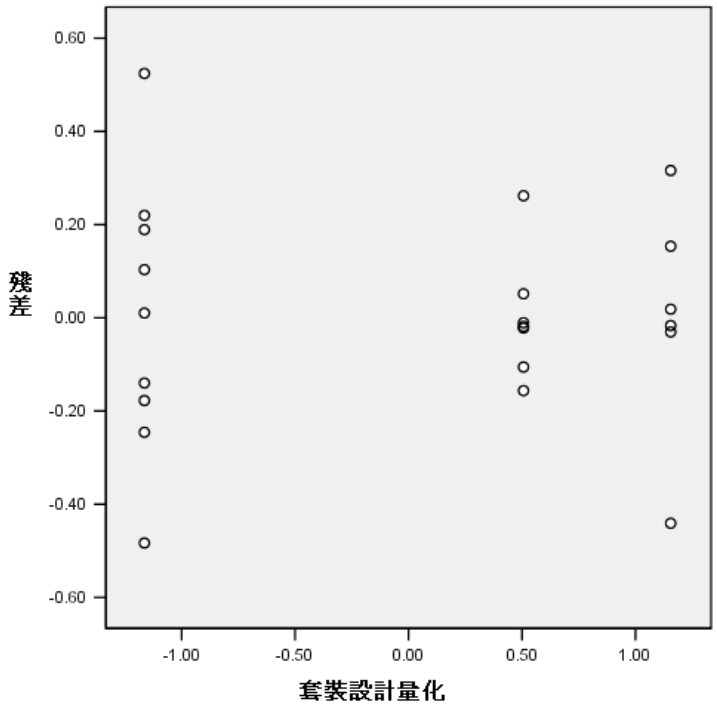
圖表 9-23
圖表建立器




- 選取「**散佈圖/點形圖**」圖庫並選擇「簡單散佈圖」。
- 選取「殘差」作為 y 軸的變數。
- 選取「包裝設計量化」作為 x 軸的變數。
- 按一下「**確定**」。

散佈圖顯示標準化殘差與「包裝設計」的最適分數比較。所有的殘差都介於 0 的兩個標準差之間。隨機散佈點置換了標準化線性迴歸散佈圖內的 U 型圖。可最佳化量化類別來改善預測功能。

圖表 9-24
類別迴歸的殘差




# 範例：臭氧資料

在這個範例中，您將使用較大的資料集來說明最適尺度轉換的選擇及作用。包括之前由 Breiman 及 Friedman (Breiman 和 Friedman 檢定(F), 1985)，和 Hastie 及 Tibshirani(Hastie 和 Tibshirani, 1990) 等人分析的六個氣象變數中的 330 個觀察值。下表描述原始變數。您的類別迴歸會嘗試自其餘的變數中預測臭氧濃度。先前研究人員在這些會阻礙標準迴歸方式的變數中發現非線性。

表格 9-2
原始變數

| 變數 | 說明 |
|---|---|
| 臭氧 | 每日臭氧濃度；列入 38 個類別其中之一 |
| ibh | 逆溫層底高度 |
| dpg | 氣壓梯度 (mm Hg) |
| vis | 能見度 (哩) |
| temp | 溫度 (華氏) |
| doy | 日期 |

這個資料集存放在 ozone.sav 中。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

## 離散化變數

如果變數擁有的類別多於實際可解讀的類別，您應該利用「離散化」對話方塊，修正類別，將類別範圍降到更便於管理的數字。

「日期」變數的最小值為 3，而最大值為 365。在類別迴歸中使用這個類別代表使用擁有 365 個類別的變數。同樣地，「能見度」(哩) 的範圍介於 0 與 350。若要簡化分析的解讀，請將這些變數離散化至長度為 10 的各個相等區間。

「逆溫層底高度」變數介於 111 與 5000。擁有這麼多類別的變數會導致非常複雜的關係。但是，將這個變數離散化至長度為 100 的各個相等區間會產生約 50 個類別。利用 50 個類別的變數比利用 5000 個類別的變數能更顯著地簡化解讀。

「氣壓梯度」(mm Hg) 介於 -69 與 107。這個程序會略過以分析中的負數編碼的任何類別，但是將這個變數離散化至長度為 10 的各個相等區間會產生約 19 個類別。

「溫度 (華氏)」介於華氏刻度 25 與 93 之間。為了將溫度資料以攝氏刻度分析，將這個變數離散化至長度為 1.8 的各個相等區間。

可能需要為變數進行不同的離散化。此處所用的選擇是完全主觀的選擇。如果您想要更少的類別，請選擇較大的區間。例如，「日期」可以分為月份或季節。

## 選擇轉換種類

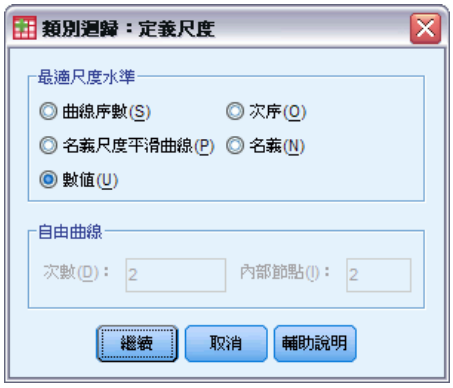
每個變數都可由多種不同水準的其中一個水準分析。但是，由於目標為預測回應值，所以您調整回應值的方法應該“就如同”使用數值最適尺度水準。因此，類別之間的順序及差異將保留於轉換變數內。

- 若要執行「類別迴歸」分析，請從功能表選擇：
  **分析(A) > 迴歸 > 最適尺度 (CATREG)...**

圖表 9-25
「類別迴歸」對話方塊

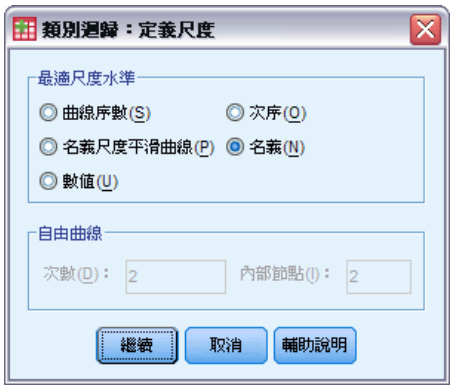
- 選取「每日臭氧濃度」作為依變數。
- 透過「日期」選取「逆溫層底高度」作為自變數。
- 選取「每日臭氧濃度」並按一下「**定義尺度**」。

圖表 9-26
「定義尺度」對話方塊



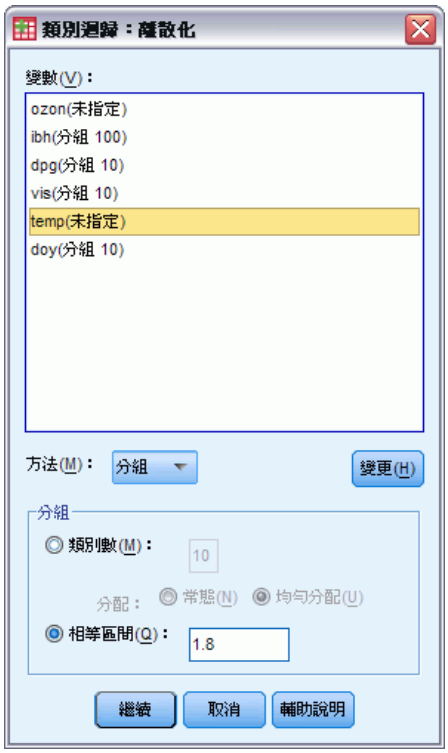

- 選取「**數值**」作為最適尺度水準。
- 按一下「**繼續**」。
- 透過「日期」選取「逆溫層底高度」，並按一下「類別迴歸」對話方塊中的「**定義尺度**」。

圖表 9-27
「定義尺度」對話方塊




- 選取「名義」作為最適尺度水準。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「離散化」。

圖表 9-28
「離散化」對話方塊




- 選取「ibh」。
- 選取「相等區間」並輸入 100 作為區間長度。
- 按一下「變更」。

- 選取「dpg」，「vis」，及「doy」。
- 輸入 10 作為區間長度。
- 按一下「變更」。
- 選取「temp」。
- 輸入 1.8 作為區間長度。
- 按一下「變更」。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「圖形」。

圖表 9-29
「圖形」對話方塊

- 透過「日期」選取「逆溫層底高度」的轉換圖。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「確定」。

圖表 9-30
模式摘要(M)

| | 複相關係數 R | R 平方 | 調過後的 R 平方 | 顯著預測錯誤 |
|---|---|---|---|---|
| 標準化資料 | .938 | .880 | .785 | .120 |

依變數:Daily ozone level
預測器: Inversion base height Pressure gradient (mm Hg) Visibility (miles) Temperature (degrees F) Day of the year

將所有預測值視為名義預測值會產生 0.880 的 $R^2$ 值。由於名義處理對量化沒有任何限制，因此大量的變異數說明於並不令人訝異。但是，解讀結果可能會很困難。

圖表 9-31
迴歸係數（所有預測值名義）

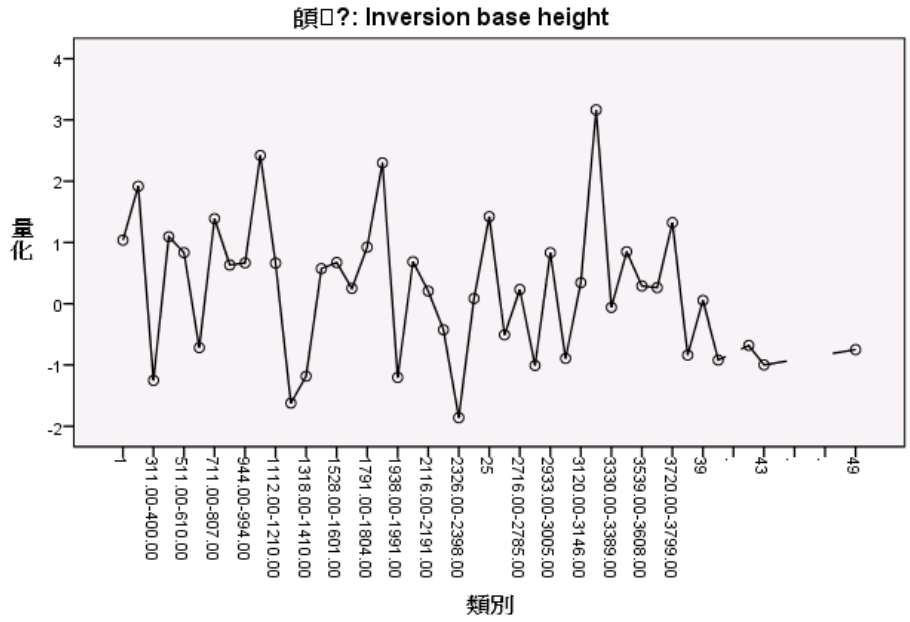
係數

| | 標準化係數 | | | | |
|---|---|---|---|---|---|
| | Beta 分配 | Bootstrap (1000) 標準誤 | df | F | 顯著性 |
| Inversion base height | .297 | .053 | 42 | 31.047 | .000 |
| Pressure gradient (mm Hg) | .326 | .055 | 16 | 34.793 | .000 |
| Visibility (miles) | .229 | .050 | 17 | 21.465 | .000 |
| Temperature (degrees F) | .577 | .091 | 35 | 40.562 | .000 |
| Day of the year | .420 | .070 | 36 | 36.171 | .000 |

依變數: Daily ozone level

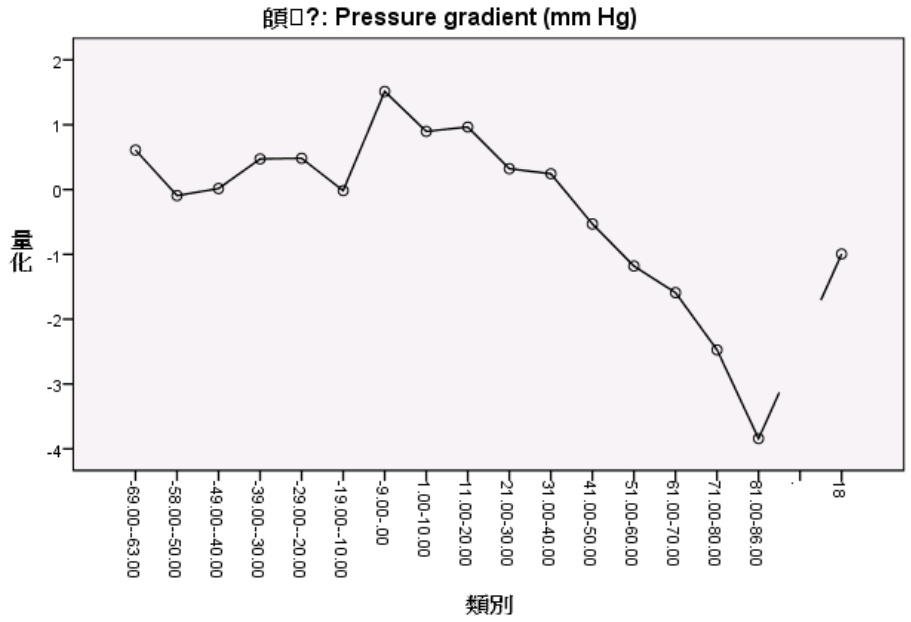
本表顯示預測值的標準化迴歸係數。解讀這些數值常見的錯誤是太專注於係數而忽略了量化。您不能就這麼斷言，舉例來說，逆溫層底高度的正值表示臭氧會隨著預測值的提高而增加。所有解讀都必須相對於已轉換變數，因此隨著逆溫層底高度量化的增加，預測的臭氧也會增加。若要檢查原始變數的作用，您必須建立類別與量化之間的關係。

圖表 9-32
「逆溫層底高度」（名義）的轉換圖




「逆溫層底高度」的轉換圖沒有顯示明顯樣式。經由鋸齒狀圖形證實，從低類別移到高類別會產生朝兩個方向發展的量化變動。因此，描述這個變數的作用需要專注於個別類別。在這個變數的量化上，套用次序或線性限制可能會顯著地降低配適度。

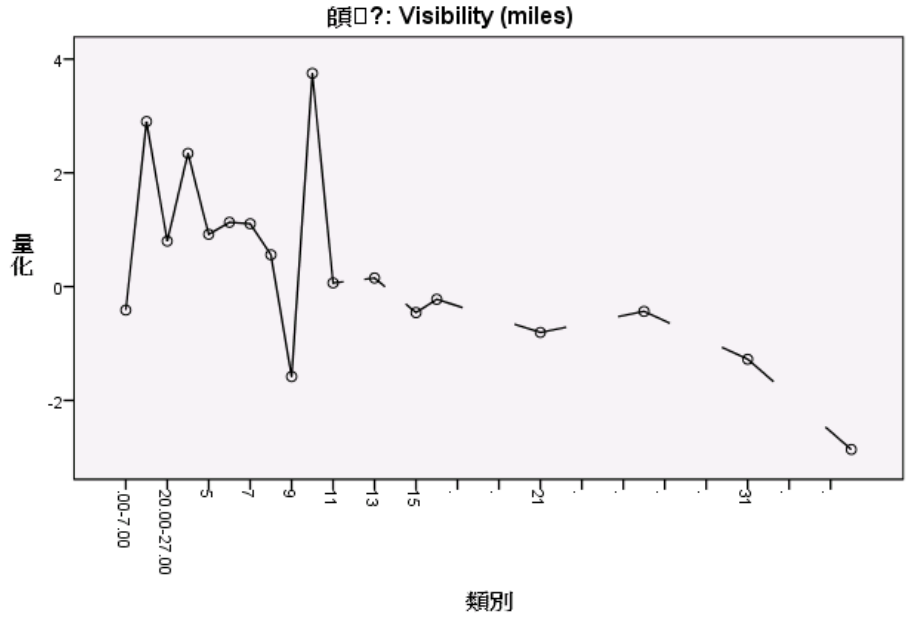
圖表 9-33
「氣壓梯度」（名義）的轉換圖




本圖顯示「氣壓梯度」的轉換圖。初始離散化的類別 (1 到 6) 會取得小型量化，且因此對預測回應值的貢獻較小。後續三個類別會取得較高的正值，導致預測的臭氧適度增加。

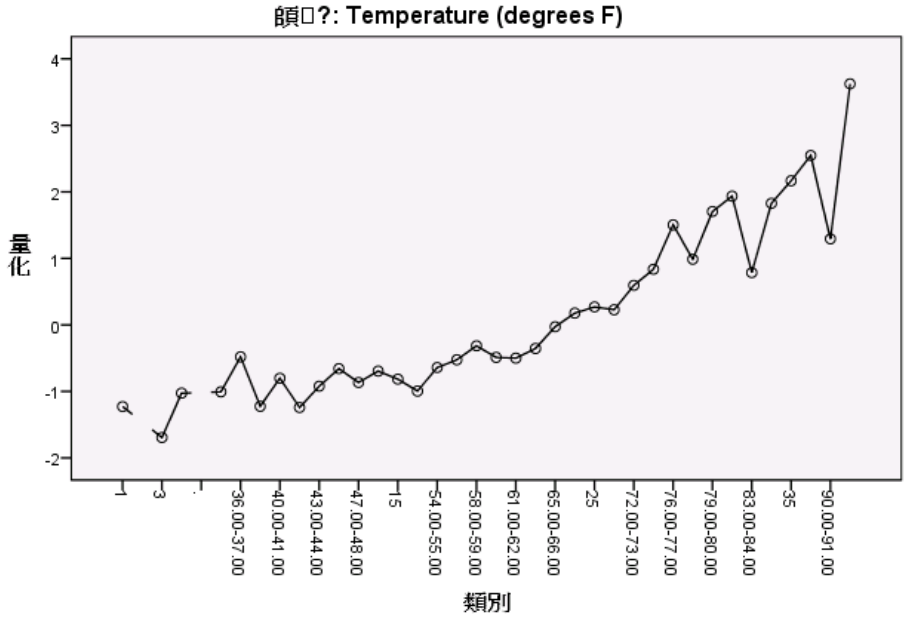
量化減少至類別 16，其中「氣壓梯度」最能降低預測的臭氧。雖然線條在這個類別之後上升，以次序尺度水準調整「氣壓梯度」來簡化效應解讀時，可能不會顯著地降低配適度。但是，0.04 的重要性量數及「氣壓梯度」的迴歸係數指出這個變數在迴歸中並不是非常有用。

圖表 9-34
「能見度」（名義）的轉換圖



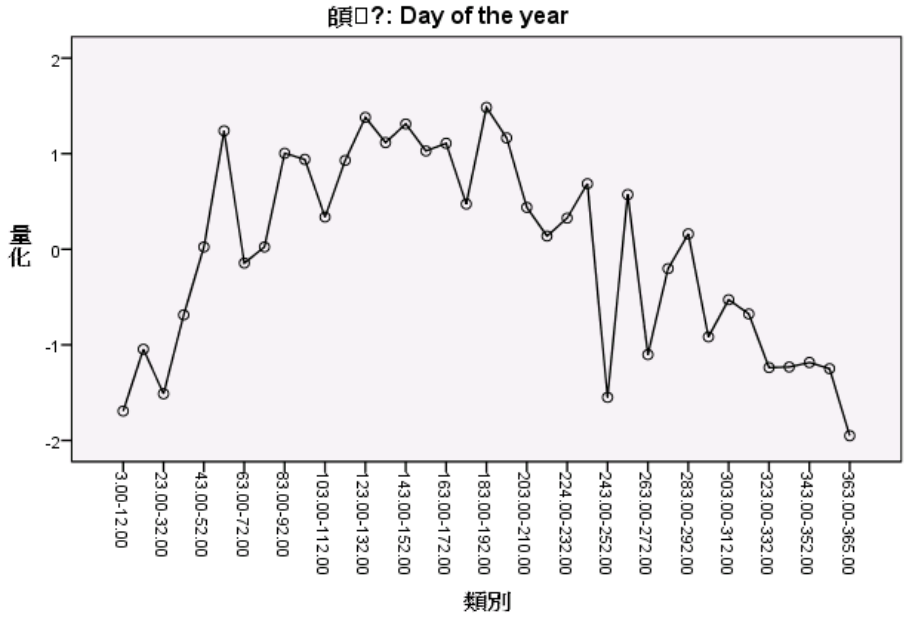

「能見度」的轉換圖，像「逆溫層低高度」的轉換圖，沒有顯示明顯樣式。在這個變數的量化上，套用次序或線性限制可能會顯著地降低配適度。

圖表 9-35
「溫度」（名義）的轉換圖




「溫度」的轉換圖顯示其他樣式。隨著類別增加，量化也傾向增加。結果，隨著「溫度」上升，預測的臭氧也傾向增加。這個樣式代表將「溫度」調整為次序水準。

圖表 9-36
「日期」（名義）的轉換圖

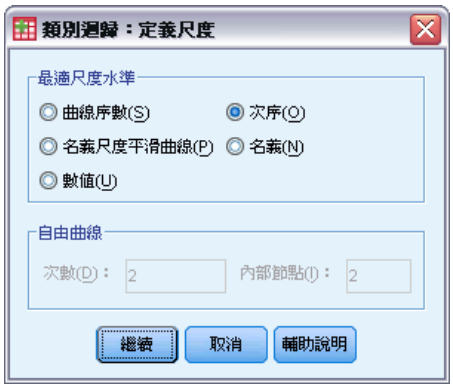



本圖顯示「日期」的轉換圖。量化傾向增加到統計圖中點，到了那一點後它們傾向減少，產生一個反向 U 型圖。考量「日期」的迴歸係數符號，原始組別會取得對預測臭氧有降低作用的量化。在中間類別方面，預測臭氧的量化作用會增加，在統計圖中間左右達到最大。

在那點之後，量化傾向降低預測的臭氧。雖然線條的鋸齒很多，整體形狀仍可辨識。因此，轉換圖代表「溫度」以次序水準測量，而其他所有預測值都以名義方式測量。

若要重新計算迴歸，請將「溫度」調整為次序水準，叫回「類別迴歸」對話方塊。

圖表 9-37
「定義尺度」對話方塊




- 選取「溫度」並按一下「**定義尺度**」。
- 選取「**次序**」作為最適尺度水準。

- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「儲存」。

圖表 9-38
「儲存」對話方塊

- 在「轉換變數」組別中，選取「將殘差儲存到作用中資料集」。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「確定」。

圖表 9-39
含有溫度（名義）之迴歸的模式摘要

**模式摘要**

| | 複相關係數 R | R 平方 | 調過後的 R 平方 | 顯著預測錯誤 |
|---|---|---|---|---|
| 標準化資料 | .934 | .872 | .787 | .128 |

依變數:Daily ozone level
預測器: Inversion base height Pressure gradient (mm Hg) Visibility (miles) Temperature (degrees F) Day of the year

這個模式會產生 0.872 的 $R^2$，因此當「溫度」量化被限制為必須排序時，變異數所說明之數值降低可加以忽略。

圖表 9-40
含有溫度（次序）的迴歸係數

係數

| | 標準化係數 | | | | |
|---|---|---|---|---|---|
| | Beta 分配 | Bootstrap (1000) 標準誤 | df | F | 顯著性 |
| Inversion base height | .298 | .042 | 42 | 51.376 | .000 |
| Pressure gradient (mm Hg) | .301 | .047 | 16 | 41.882 | .000 |
| Visibility (miles) | .224 | .044 | 17 | 25.659 | .000 |
| Temperature (degrees F) | .609 | .084 | 21 | 52.113 | .000 |
| Day of the year | .373 | .053 | 36 | 50.330 | .000 |

依變數: Daily ozone level

本表顯示「溫度」的尺度為次序的模式係數。 比較係數與「溫度」的尺度為名義的模式係數，並沒有太大的改變。

圖表 9-41
相關性、重要性及允差

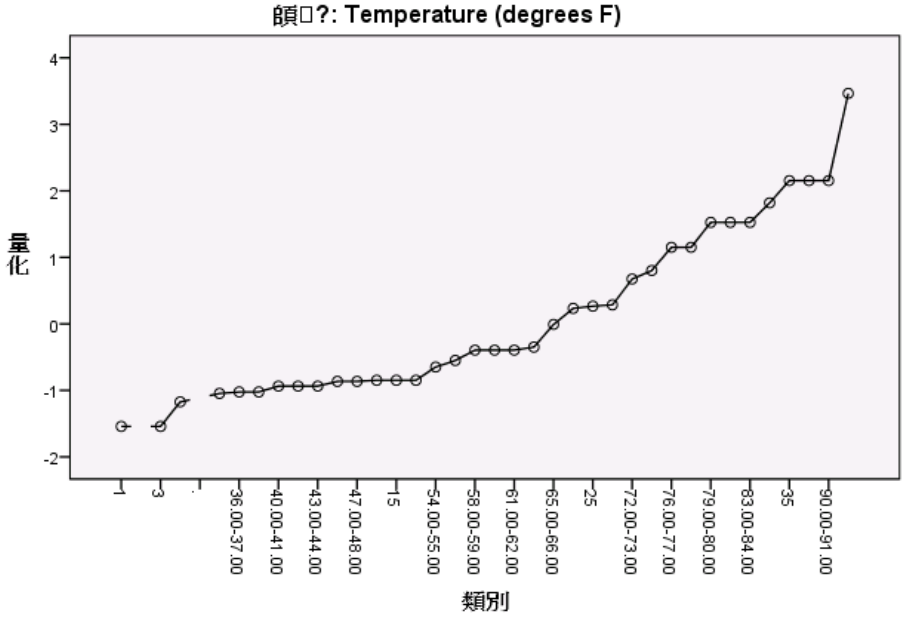
相關係數和容許度

| | 相關 | | | | 允差 | |
|---|---|---|---|---|---|---|
| | 淨相關 | 偏 | 部分 | 重要性 | 轉換後 | 轉換前 |
| Inversion base height | .438 | .627 | .288 | .150 | .930 | .596 |
| Pressure gradient (mm Hg) | .128 | .606 | .272 | .044 | .815 | .858 |
| Visibility (miles) | .365 | .518 | .216 | .094 | .933 | .752 |
| Temperature (degrees F) | .804 | .843 | .559 | .562 | .842 | .580 |
| Day of the year | .352 | .677 | .329 | .151 | .777 | .802 |

依變數: Daily ozone level

此外，重要性量數表示「溫度」對迴歸的重要性仍然比其他變數高出許多。然而現在，根據「溫度」的次序尺度水準及正迴歸係數，您可以斷言隨著「溫度」增加，預測的臭氧也會增加。

圖表 9-42
「溫度」（次序）的轉換圖




轉換圖將說明「溫度」量化的次序限制。名義轉換的鋸齒線在此處會由平滑的遞增線條取代。此外，沒有出現長期高原圖形，顯示不需要收合類別。

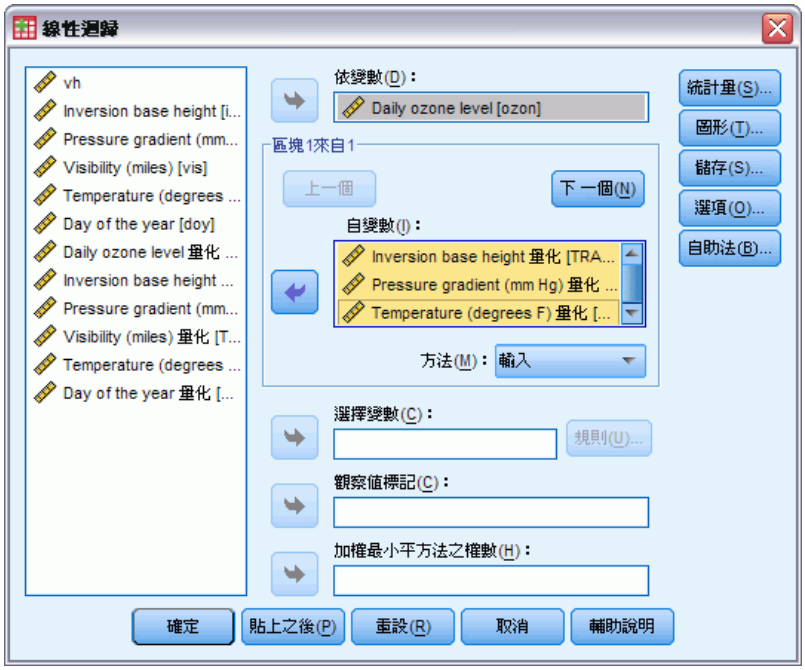
## 量化的最佳值

來自類別迴歸的轉換變數可用於標準線性迴歸中，產生相同的結果。但是，只有對產生量化的模式而言，這些量化才是最適合的。使用線性迴歸中的預測值子集不會對應到相同子集的最適尺度迴歸。

例如，您已經計算的類別迴歸有 0.875 的 $R^2$ 值。您已經儲存轉換後變數，所以為了符合只使用「溫度」、「氣壓梯度」，及「逆溫層底高度」作為預測值的線性迴歸，請從功能表中選擇：

**分析(A) > 迴歸 > 線性...**

圖表 9-43
「線性迴歸」對話方塊




- 選取「每日臭氧濃度量化」作為依變數。
- 選取「逆溫層底高度量化」、「氣壓梯度（mm Hg）量化」，及「溫度（華氏）量化」作為自變數。
- 按一下「確定」。

圖表 9-44
含最適尺度預測值子集之迴歸的模式摘要

**模式摘要**

| 模式 | R | R 平方 | 調過後的 R 平方 | 估計的標準誤 |
|---|---|---|---|---|
| 1 | .856[a] | .732 | .729 | 4.16711 |

a. 預測變數:(常數), Temperature (degrees F) 量化, Pressure gradient (mm Hg) 量化, Inversion base height 量化

利用回應值的量化，標準線性迴歸中的「溫度」、「氣壓梯度」，及「逆溫層底高度」將產生配適度為 0.732 的結果。若要將這個結果與只使用那三個預測值的類別迴歸配適度比較，請叫回「類別迴歸」對話方塊。

圖表 9-45
「類別迴歸」對話方塊

- 取消選取「能見度（哩）」及「日期」作為自變數。
- 按一下「確定」。

圖表 9-46
三個預測值之類別迴歸的模式摘要

| | 複相關係數 R | R 平方 | 調過後的 R 平方 | 顯著預測錯誤 |
|---|---|---|---|---|
| 標準化資料 | .892 | .796 | .735 | .204 |

依變數:Daily ozone level
預測器: Inversion base height Pressure gradient (mm Hg)
Temperature (degrees F)

類別迴歸分析的配適度為 0.796，比 0.732 的配適度更佳。這顯示只有在模式中包含全部五個變數時，量化在原始迴歸中取得的尺度性質才是最適尺度。

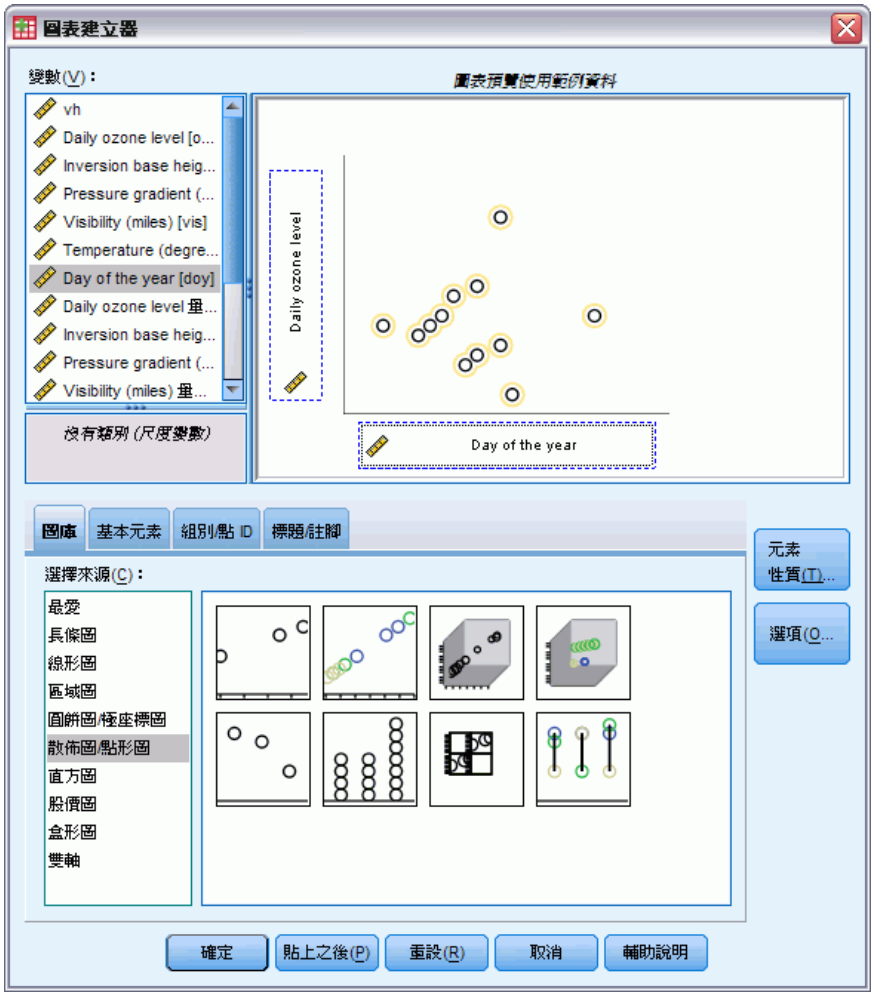
## 轉換的作用

轉換變數會造成原始回應值及轉換後變數的預測值線性原始組合之間的非線性關係。但是，有多個預測值時，成對關係會被模式中的其他變數混淆。

若要將您的分析焦點放在「每日臭氧濃度」及「日期」的關係，請從檢閱散佈圖開始。從功能表選擇：

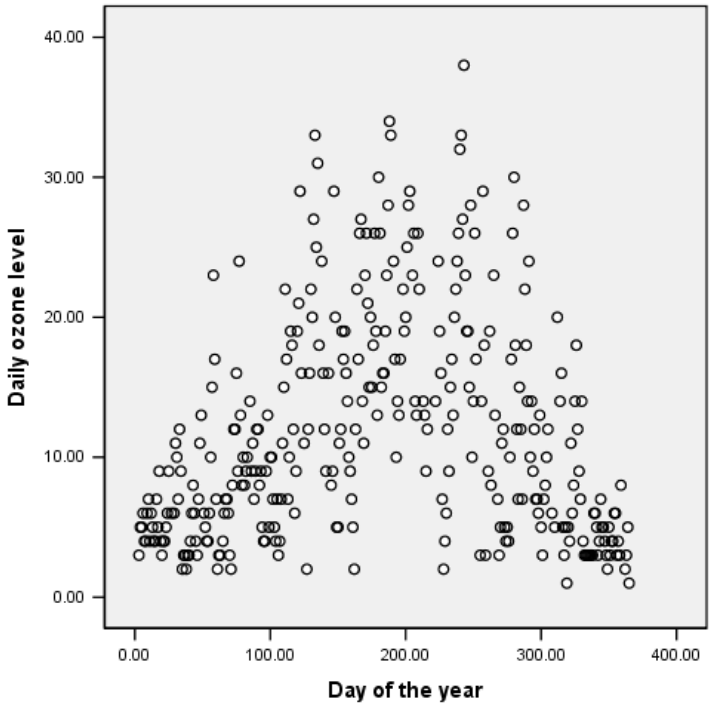
統計圖(G) > 圖表建立器(C)...

圖表 9-47
「圖表建立器」對話方塊




- 選取「**散佈圖/點形圖**」圖庫並選擇「簡單散佈圖」。
- 選取「每日臭氧濃度」作為 y 軸變數及「日期」作為 x 軸變數。
- 按一下「**確定**」。

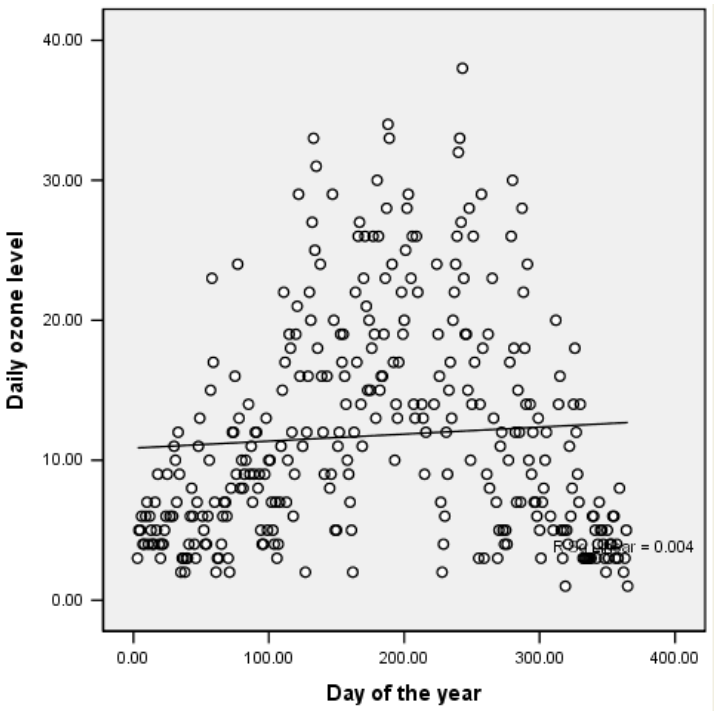
圖表 9-48
「每日臭氧濃度」及「日期」的散佈圖




本圖說明「每日臭氧濃度」及「日期」之間的關係。「日期」增加到接近 200 時，「每日臭氧濃度」也隨著增加。但是，「日期」增加到超過 200 時，「每日臭氧濃度」卻減少。這個反向的 U 型圖顯示這兩個變數之間是二次關係。線性迴歸無法擷取這個關係。

- 若要看到最適線條覆蓋在散佈圖的圖點上，請連按兩下來啟動圖形。
- 在「圖表編輯程式」中，選取一個點。
- 按一下「**在總數新增配適直線**」工具，並關閉「圖表編輯程式」。

圖表 9-49
散佈圖顯示最配適直線



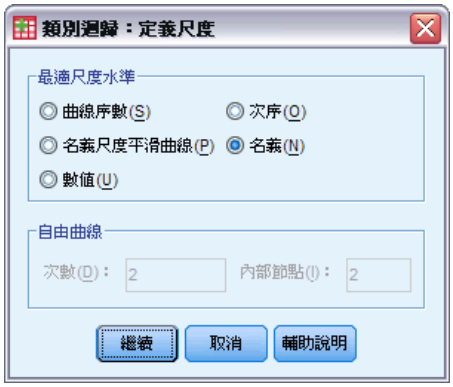

在「日期」上的「每日臭氧濃度」線性迴歸會產生 0.004 的 $R^2$ 值。這個配適度顯示「日期」沒有「每日臭氧濃度」的預測值。考慮到圖中所顯示的樣式，這點並不令人驚訝。但是，您可使用最適尺度，將二次關係線性化，並使用轉換後的「日期」預測回應值。

圖表 9-50
「類別迴歸」對話方塊

若要取得「日期」上的「每日臭氧濃度」類別迴歸，請叫回「類別迴歸」對話方塊。

- 透過「溫度（華氏刻度）」而取消選取「逆溫層底高度」作為自變數。
- 選取「日期」作為自變數。
- 按一下「**定義尺度**」。

圖表 9-51
「定義尺度」對話方塊

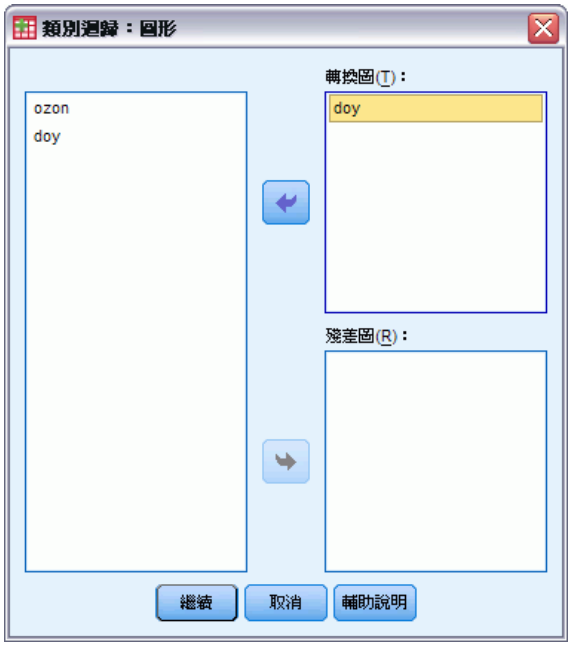



- 選取「**名義**」作為最適尺度水準。
- 按一下「**繼續**」。
- 在「類別迴歸」對話方塊中，按一下「**離散化**」。

圖表 9-52
「離散化」對話方塊

- 選取「doy」。
- 選取「**相等區間**」。
- 輸入 10 作為區間長度。
- 按一下「**變更**」。
- 按一下「**繼續**」。
- 在「類別迴歸」對話方塊中，按一下「**圖形**」。

圖表 9-53
「圖形」對話方塊




- 為轉換圖選取「doy」。
- 按一下「繼續」。
- 在「類別迴歸」對話方塊中，按一下「確定」。

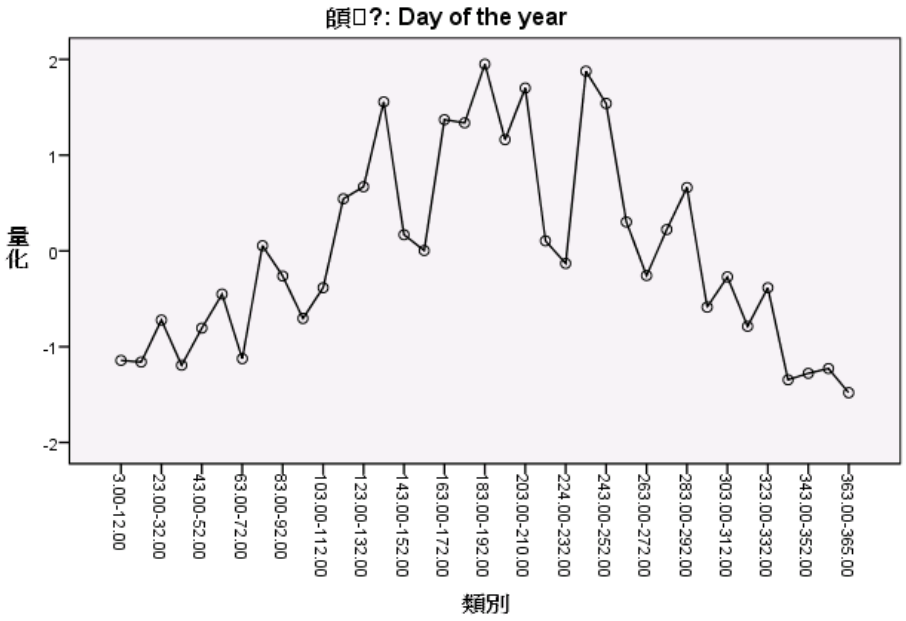
圖表 9-54
「每日臭氧濃度」及「日期」之類別迴歸的模式摘要

| | 複相關係數 R | R 平方 | 調過後的 R 平方 | 顯著預測錯誤 |
|---|---|---|---|---|
| 標準化資料 | .741 | .549 | .494 | .451 |

依變數:Daily ozone level
預測器: Day of the year

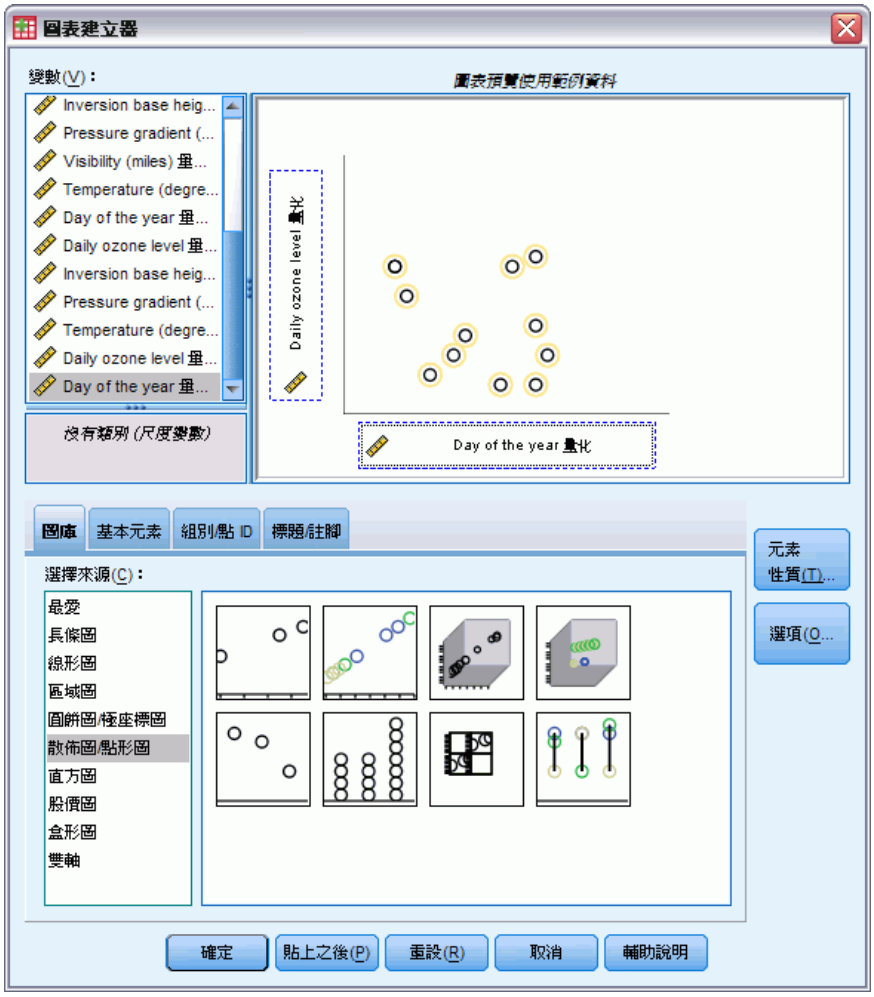
最適尺度迴歸將「每日臭氧濃度」視為數值及「日期」視為名義。產生 0.549 的 $R^2$ 值。雖然「臭氧濃度」內只有 55% 的變異可由類別迴歸說明，但相對於原始迴歸而言，這已經是一大進步。轉換「日期」可預測「臭氧濃度」。

圖表 9-55
「日期」（名義）的轉換圖




本圖顯示「日期」的轉換圖。「日期」的極端值都取得負量化，但集中值都取得正量化。透過套用這個轉換，「日期」數值高低對於預測的「每日臭氧濃度」有類似效果。

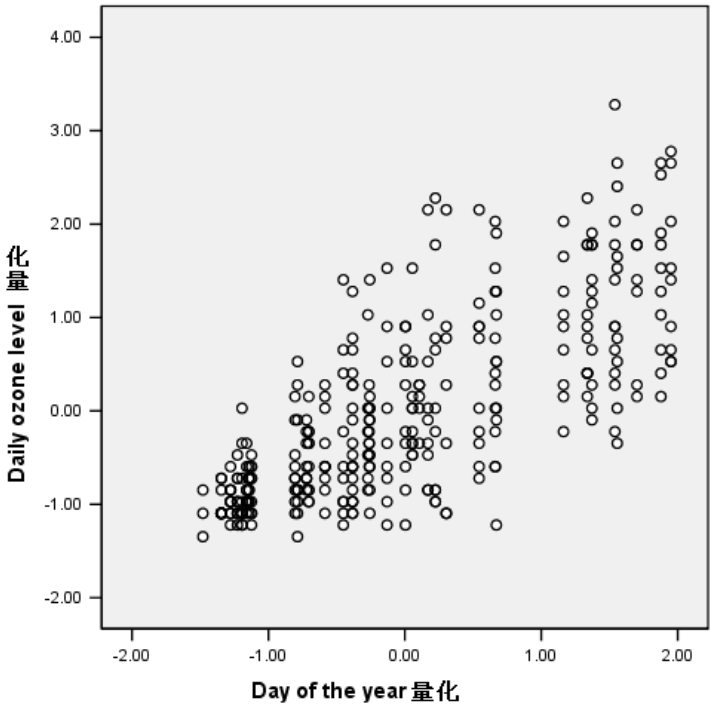
圖表 9-56
圖表建立器




若要看到轉換後變數的散佈圖，請叫回「圖表建立器」，並按一下「**重設**」來清除您之前的選擇。

- 選取「**散佈圖/點形圖**」圖庫並選擇「簡單散佈圖」。
- 選取「每日臭氧濃度量化 [TRA1_3]」作為 y 軸變數及「日期 [TRA2_3]」作為 x 軸變數。
- 按一下「**確定**」。

圖表 9-57
轉換後變數的散佈圖




本圖說明轉換後變數之間的關係。上升趨勢置換反向的 U 型圖。迴歸線有正斜率，顯示隨著「日期」增加，預測的「每日臭氧濃度」也增加。使用最適尺度可將關係線性化並可進行原本未被注意的解讀。

# 閱讀資料推薦

請參閱下列文字以取得類別迴歸的詳細資訊：

Buja, A. 1990. Remarks on functional canonical variates, alternating least squares methods and ACE (最小平方法與 ACE 交替使用下函數典型變異之淺見). Annals of Statistics (統計學年鑑), 18, .

Hastie, T., R. Tibshirani, 和 A. Buja. 1994. Flexible discriminant analysis (彈性判別分析). Journal of the American Statistical Association (美國統計協會彙報), 89, .

Hayashi, C. 1952. On the prediction of phenomena from qualitative data and the quantification of qualitative data from the mathematico-statistical point of view (從數學統計觀點論性質資料與性質資料量化之現象預測). Annals of the Institute of Statitical Mathematics (統計數學所年鑑), 2, .

Kruskal, J. B. 1965. Analysis of factorial experiments by estimating monotone transformations of the data (資料單調轉換評估之因子實驗分析). Journal of the Royal Statistical Society Series B (皇家統計協會彙報，系列 B), 27, .

Meulman, J. J. 2003. Prediction and classification in nonlinear data analysis: (非線性資料分析之預測與分類)Something old, something new, something borrowed, something blue (舊的、新的、借的、藍的). Psychometrika (心理學計量報導), 4, .

Ramsay, J. O. 1989. Monotone regression splines in action (單調迴歸曲線實作). Statistical Science (統計學), 4, .

Van der Kooij, A. J., 和 J. J. Meulman. 1997. MURALS:Multiple regression and optimal scaling using alternating least squares (使用替代最小平方之多重迴歸與最適尺度). 於： Softstat ’97, F. Faulbaum, 和 W. Bandilla，eds. Stuttgart: Gustav Fisher.

Winsberg, S., 和 J. O. Ramsay. 1980. Monotonic transformations to additivity using splines (利用曲線進行可加性單調轉換). Biometrika (生物統計學雜誌), 67, .

Winsberg, S., 和 J. O. Ramsay. 1983. Monotone spline transformations for dimension reduction (維度縮減之單調曲線轉換). Psychometrika (心理學計量報導), 48, .

Young, F. W., J. De Leeuw, 和 Y. Takane. 1976. Regression with qualitative and quantitative variables: (性質與數值變數迴歸)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 41, .

章節
10

# 類別主成份分析

類別主成份分析可以視為一種維度縮減的方式。可以藉由分析一組變數來顯示主要的維度變化。原始資料集可以被新的、較小的資料集所取代，如此損失的資訊最少。此方法會顯示變數之間、觀察值之間、以及變數和觀察值之間的關係。

類別主成份分析用來量化觀察資料的準則是，個體分數（成份分數）與每個量化變數有很高的相關性。有方法可以讓條件更確切。

在此以兩個範例說明類別主成份分析。第一個範例的資料集較小，適合說明與該程序相關的基本概念和解釋。第二個範例則適合檢驗實際的應用狀況。

## 範例：檢驗社會系統的相互關係

此範例檢驗 Guttman 的 (Guttman 值, 1968)，此表已由 Bell(Bell, 1961) 加以調整。Lingoes (Lingoes, 1968) 也曾討論過相關的資料。

Bell 以此表說明可能的社會團體。Guttman 過去曾使用此表的一部分，在這部分中有 5 個變數，分別說明 7 個理論社會團體的社會互動、團體歸屬感、成員實際接觸和關係正式性，而這 7 個群組包括：群眾（例如，足球場上的人）、觀眾（例如在戲院中和課堂上的人）、公眾（例如，報紙讀者和電視觀眾）、暴民（和群眾相似，但互動較為激烈）、原級團體（親密性）、次級團體（自願性）和現代社群（因親密的身體接近而導致鬆散的結盟和特殊服務的需求）。

下表中顯示資料集的變數，這些變數分類成 Guttman-Bell 所使用的 7 個社會團體，其中包括了其變數標記和與每個變數層級關聯的數值標記（類別）。這個資料集存放在 guttman.sav 中。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案. 除了選取主成份分析計算中要包括的變數外，您可以選取用於圖形中標記物件的變數。在此範例中，分析會包含資料中的前 5 個變數，而集群專門用來作為標記變數。當您指定類別主成份分析時，必須為每個分析變數指定最適尺度水準。在此例中，會指定所有分析變數的次序水準。

**表格 10-1**
**Guttman-Bell 資料集中的變數**

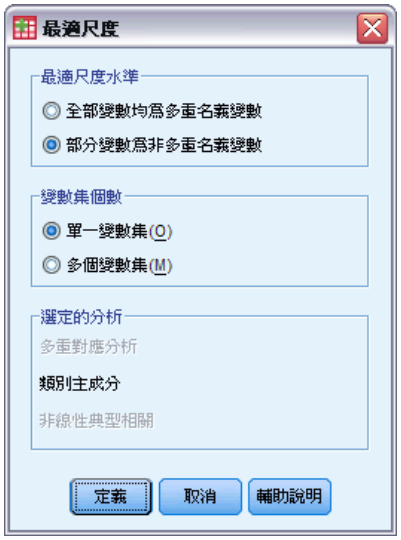
| 變數名稱 | 變數標記(L) | 數值標記 |
|---|---|---|
| intnsity | 交互作用強度 | 弱，低，中，高 |
| frquency | 交互作用頻率 | 極少，少，不常，常常 |
| blonging | 歸屬感 | 無，弱，不定，高 |
| proxmity | 實際接觸 | 疏遠，親近 |
| formlity | 關係正式性 | 無關係，正式，非正式 |
| 集群 | | 群眾，觀眾，公眾，暴民，原級團體，次級團體，現代社群 |

© Copyright SPSS Inc. 1989, 2010

# 執行分析

- 如需產生此資料集的類別主成份輸出，請從功能表中選擇：
  **分析(A) > 維度縮減(D) > 最適尺度...**

圖表 10-1
「最適尺度」對話方塊

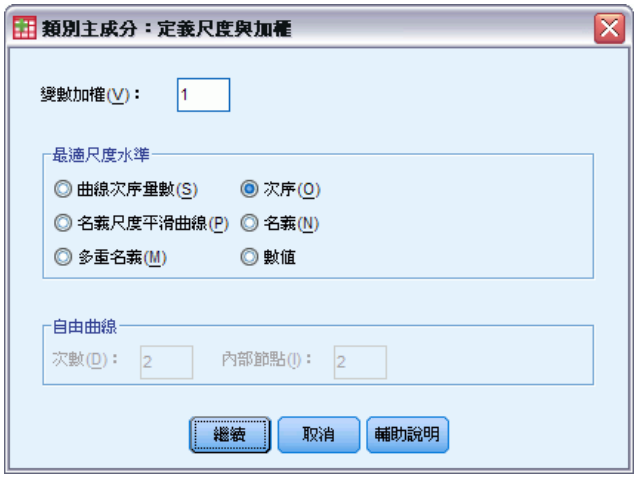



- 選擇「最適尺度水準」組別中的「**部份變數為非多重名義**」。
- 按一下「**定義**」。

圖表 10-2
「類別主成分」對話方塊

- 選取交互作用強度到關係正式性作為分析變數。
- 按一下**定義尺度與加權**。

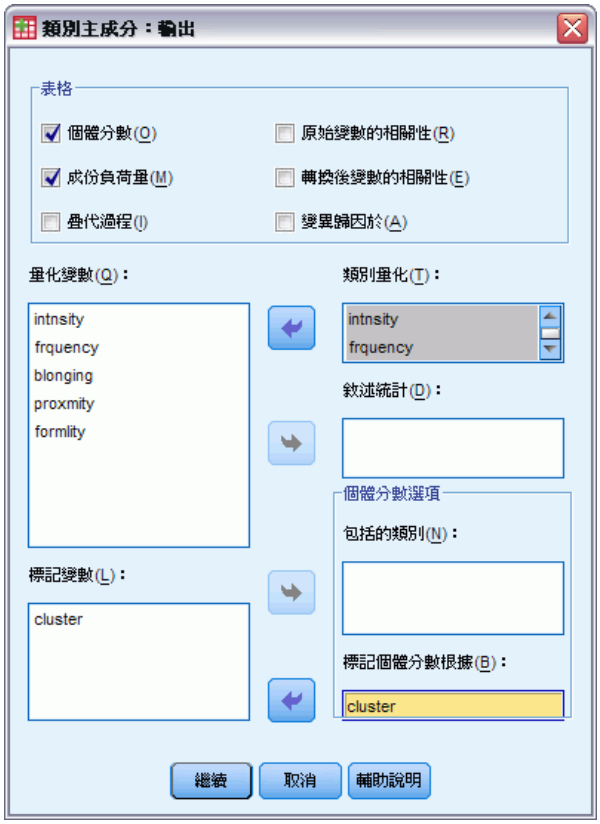
圖表 10-3
「定義尺度與加權」對話方塊




- 選取「最適尺度水準」組別中的**次序**。
- 按一下「繼續」。
- 選取「類別主成份」對話方塊中的集群作為標記變數。

▶ 按一下「輸出」。

圖表 10-4
「輸出」對話方塊




▶ 在「表格」組別中，選取個體分數，並取消選取轉換變數的相關性。。

▶ 選擇以產生 intnsity（交互作用強度）到 formlity（關係正式性）的類別量化。

▶ 選擇按照集群來標記個體分數。

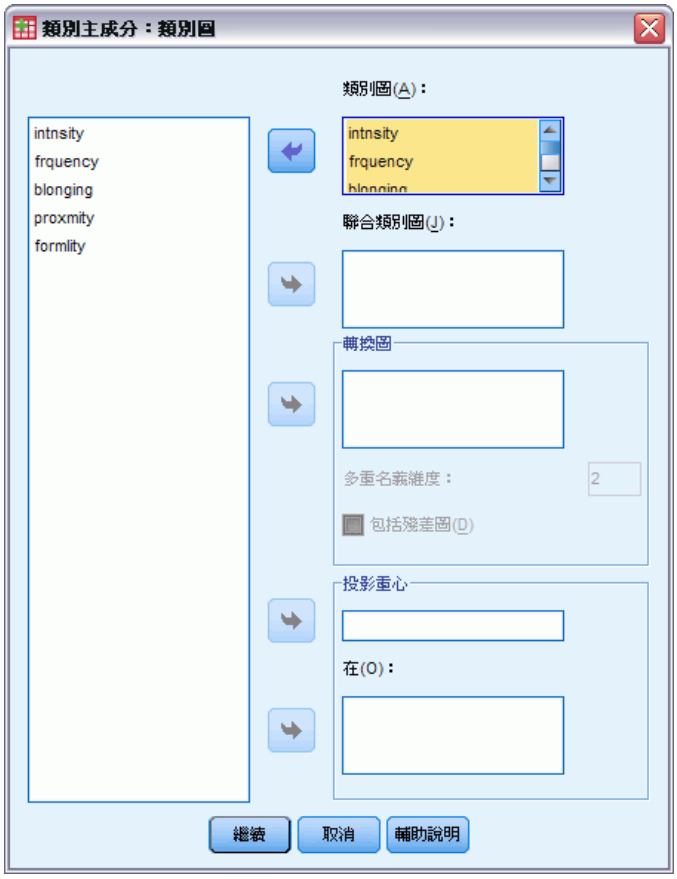
▶ 按一下「繼續」。

▶ 在「類別主成份」對話方塊的「圖形」群組中，按一下「個體」。

圖表 10-5
「個體與變數圖」對話方塊

- 在「圖形」組別中，選取「個體與變數（行列散佈圖）」。
- 在「標記個體」組別中，選擇依據變數來標記個體，然後選取「集群」作為標記個體所依據的變數。
- 按一下「繼續」。
- 在「類別主成份」對話方塊的「圖形」組別中，按一下「類別」。

圖表 10-6
「類別圖」對話方塊




- 選擇以產生 intnsity（交互作用強度）到 formlity（關係正式性）的聯合類別圖。
- 按一下「繼續」。
- 在「類別主成份」對話方塊中，按一下「確定」。

## 維度數

這些數字顯示類別主成份分析的某些初始輸出。在演算法的疊代歷程之後，會顯示模式摘要，其中包括每個維度的特徵值。這些特徵值等於典型主成份分析的特徵值。它們是表示每個維度有多少變異的量數。

圖表 10-7
疊代歷程(I)

| 疊代編號 | 變異數歸因於 | | 損失 | | |
|---|---|---|---|---|---|
| | 總數 | 增量 | 總數 | 重心座標 | 重心對向量座標的限制 |
| 0[a] | 4.515315 | .000000 | 5.484685 | 4.075583 | 1.409101 |
| 30[b] | 4.723893 | .000008 | 5.276107 | 4.276800 | 1.000196 |

a. 由於已達到收斂檢定值，因此停止疊代處理。

圖表 10-8
模式摘要(M)

| 維度 | Cronbach's Alpha 值 | 變異數歸因於 | |
|---|---|---|---|
| | | 總和 (特徵值) | 變異數的 % |
| 1 | .881 | 3.389 | 67.774 |
| 2 | .315 | 1.337 | 26.746 |
| 總數 | .986 [a] | 4.726 | 94.520 |

a. Cronbach's Alpha 平均值是以平均特徵值爲準。

特徵值可作為維度需求量的指標。在此例中，使用預設的維度數 2。這是所需的數目嗎？以一般的規則而言，當所有的變數是單一名義、次序或數值變數時，維度的特徵值應大於 1。因為二維解構成了 94.52% 的差異，所以第三維度可能無法再新增更多的資訊。

對於多重名義變數，沒有一個簡單的經驗法則可以決定適當的維度數。如果變數個數取代為類別總數減去變數個數，則上述規則仍適用。但單獨使用此規則時，則允許維度個數大於需求數。當選擇維度數時，最有益的原則便是選擇數目小者，以產生有意義的說明。模式摘要表中也會顯示 Cronbach 的 alpha 值 (信度的量數)，經由程序可讓此值最大化。

## 量化

系統會針對每個變數，呈現每個維度的量化、向量座標和重心座標。量化是指定給每個類別的數值。重心座標是在相同類別中之個體的平均分數。當向量座標落在代表個體空間之變數的線上時，向量座標即為類別座標。具有次序和數值尺度水準的變數需要這類座標。

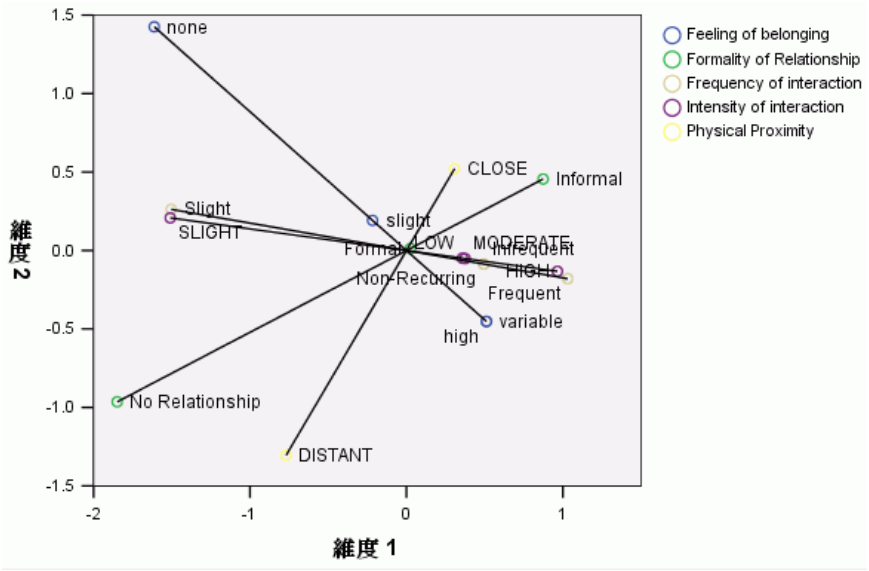
圖表 10-9
交互作用強度的量化

| 類別 | 比率 | 辨別變量 | 重心坐標 | | 矢量坐標 | |
|---|---|---|---|---|---|---|
| | | | 維度 | | 維度 | |
| | | | 1 | 2 | 1 | 2 |
| SLIGHT | 2 | -1.530 | -1.496 | .308 | -1.510 | .208 |
| LOW | 2 | .362 | .392 | .202 | .358 | -.049 |
| MODERATE | 1 | .379 | .188 | -1.408 | .374 | -.051 |
| HIGH | 2 | .978 | 1.010 | .194 | .965 | -.133 |

變數分量正常化。

略讀一下在類別點聯合圖形中的量化，您會發現如果水準為次序時，類別主成份分析無法將部份變數的類別明確的分隔開來。例如，交互作用強度和交互作用頻率這兩個變數等於或幾乎等於它們兩個中間類別的量化。這樣的結果可能意謂著要嘗試另一種類別主成份分析，或許某些類別要隱藏，或許要有不同的分析水準，例如（多重）名義。

圖表 10-10
聯合圖形類別點




類別點聯合圖形和成份負荷量圖很類似，但它也會顯示與最低量化（例如交互作用強度中的弱，或歸屬感中的無）對應的端點所在的位置。 兩個測量交互作用的變數，交互作用強度和交互作用頻率，顯示非常接近，並構成維度 1 大部分的差異，關係正式性也很接近實際接觸。

透過將目標鎖定在類別點上，您可以更清楚的了解其中的關係。交互作用強度和交互作用頻率不只是接近，其尺度的方向也很類似，也就是強度弱時其頻率也極少，強度高時其頻率為常常。您也可以看到親近的實際接觸，如握手，其關係類型為不正式，而實際接觸為疏遠者，其關係為無。

## 物件分數

您也可以要求個體分數的清單和圖形。個體分數的圖形在偵測偏離值、偵測一般個體組別或顯示一些特別圖樣等方面非常實用。

個體分數表會顯示由 Guttman-Bell 資料的社會組別所標記的個體分數清單。透過檢驗物件點的值，您可以識別圖形中特定的個體。

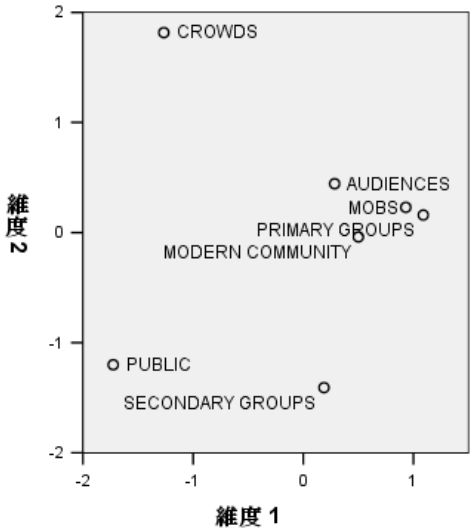
圖表 10-11
個體分數(B)

| 群集 | 維度 1 | 維度 2 |
|---|---|---|
| CROWDS | -1.266 | 1.816 |
| AUDIENCES | .284 | .444 |
| PUBLIC | -1.726 | -1.201 |
| MOBS | .931 | .229 |
| PRIMARY GROUPS | 1.089 | .159 |
| SECONDARY GROUPS | .188 | -1.408 |
| MODERN COMMUNITY | .500 | -.039 |

變數分量正常化。

第一個維度似乎將 CROWDS 和 PUBLIC（這兩者有相對大的負分）與 MOBS 和 PRIMARY GROUPS（這兩者有相對大的正分）區分。第二個維度有三個區塊：PUBLIC 和 SECONDARY GROUPS（這兩者為大的負值），CROWDS（大的正值），和介於它們之間的其他社會團體。透過檢閱個體分數的圖形，便可以很容易了解其關係。

圖表 10-12
個體分數圖

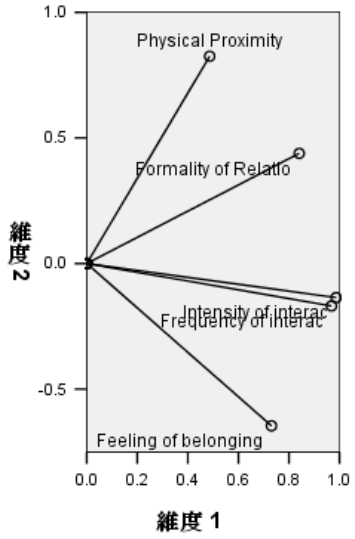



在圖形中，您可以看到 PUBLIC 和 SECONDARY GROUPS 在底端，CROWDS 在頂端，而其他社會團體在中間。根據分析單位的其他資訊，來檢驗各個體間的樣式。在此狀況下，您會知道個體的分類。而在其他狀況下，您可以使用補充變數來標記個體。 您也可以看到類別主成份分析不會區分 MOBS 和 PRIMARY GROUPS。雖然多數人通常不認為他們的家人是暴民，但在所使用的變數上，這兩個團體的 5 個變數中，有 4 個變數有相同的分數。很明顯的，您可能想要探索出所使用變數和類別的可能缺點。例如交互作用強度高和關係不正式者，可能對這兩個團體代表不同的意義。或者，您可能會考慮較高的維度解。

# 成份負荷量

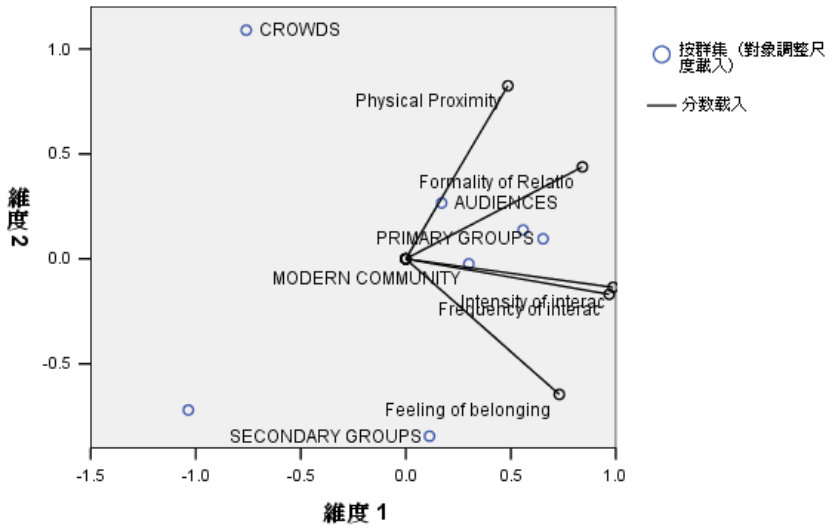
此圖顯示成份負荷量圖。向量（直線）在長度上相對較長，再次表示前兩個維度構成所有量化變數的大部分差異。在第一個維度上，所有變數都具有高（正）的成份負荷量。第二個維度主要與方向相反的量化變數歸屬感和 實際接觸相關。這表示在維度 2 中負分數高的個體，其歸屬感的分數高，實際接觸的分數低。因此，第二個維度一方面表示這兩個變數的對比，一方面表示與量化變數交互作用強度和交互作用頻率間的關係小。

圖表 10-13
成份負荷量(M)



如需檢驗個體和變數間的關係，請檢閱個體和成份負荷量的行列散佈圖。變數向量會指向變數最高類別的方向。例如，對於實際接觸和歸屬感，其最高類別分別是親近和高。因此，CROWDS會具有實際接觸親近而無歸屬感的特性，而SECONDARY GROUPS的實際接觸不親近但歸屬感高。

圖表 10-14
行列散佈圖(B)




## 額外維度

增加維度數會增加所造成的變異量，並會在較低的維度解中顯示隱含的差異性。如同先前所述，無法區分 MOBS 和 PRIMARY GROUPS 這兩個維度。但是，增加維度的範圍，可讓這兩個團體之間有差異性。

### 執行分析

- ▶ 如需取得三維的解，請開啟「類別主成份」對話方塊。
- ▶ 輸入 3 作為解中的維度數。
- ▶ 在「類別主成份」對話方塊中，按一下「**確定**」。

## 模式摘要

圖表 10-15
模式摘要(M)

| 維度 | Cronbach's Alpha 值 | 變異數歸因於 | |
|---|---|---|---|
| | | 總和 (特徵值) | 變異數的 % |
| 1 | .885 | 3.424 | 68.480 |
| 2 | -.232 | .844 | 16.871 |
| 3 | -.459 | .732 | 14.649 |
| 總數 | 1.000 | 5.000 | 99.999 |

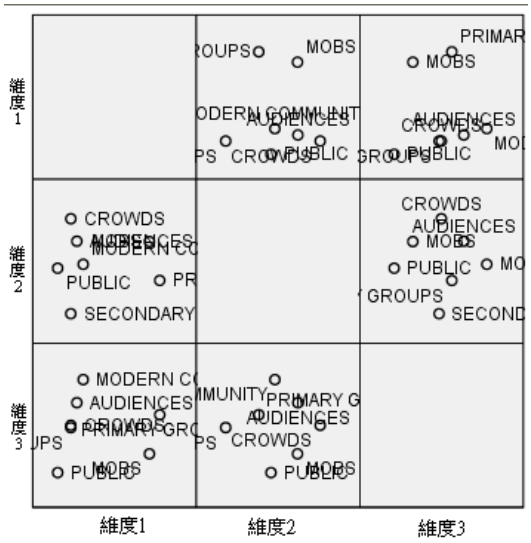
a. Cronbach's Alpha 平均值是以平均特徵值爲準。

三維解的特徵值為 3.424、0.844 和 0.732，幾乎構成所有的變異。

## 物件分數

在散佈圖矩陣中會繪製出三維解中的個體分數。在散佈圖矩陣中，每個維度會相對於每個其他維度，繪製出一系列的二維散佈圖。請注意，三維中的前兩個特徵值不等於二維解中的特徵值，換句話說，解並非巢狀。因為維度 2 和維度 3 中的特徵值現在小於 1（假設 Cronbach 的 alpha 值為負數），所以二維解對您較適用。在此提供三維解以供說明之用。

圖表 10-16
三維個體分數散佈圖矩陣




在圖形的頂端列中，會顯示第一個維度將 PRIMARY GROUPS 和 MOBS 與其他團體分開。請注意，個體沿著垂直軸的順序在任何圖形的頂端列中都不會變更；這些圖形每一個都使用維度 1 作為 y 軸。

圖形的中間列會顯示對維度 2 的解釋。第二個維度會有些改變，與二維解略有不同。第二個維度先前原本有三個區塊，但現在這些個體沿著軸展開。

第三個維度可幫助區分 MOBS 與 PRIMARY GROUPS，這在二維解中是不會發生的。

再仔細看一下維度 2 對維度 3，以及維度 1 對維度 2 圖。在由維度 2 和維度 3 所定義的平面上，這些個體約略構成一個矩形，其中的 CROWDS、MODERN COMMUNITY、SECONDARY GROUPS 和 PUBLIC 是矩形的四個頂點。在此平面上，MOBS 和 PRIMARY GROUPS 會分別顯示為 PUBLIC-CROWDS 和 SECONDARY GROUPS-MODERN COMMUNITY 的凸組合。但是，如同先前所提，他們沿著維度 1 與其他的組別分離。AUDIENCES 未沿著維度 1 與其他組別分離，並顯示為 CROWDS 和 MODERN COMMUNITY 的組合。

### 成份負荷量

圖表 10-17
三維成份負荷量

| | 維度 | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Intensity of interaction | .980 | -.005 | -.201 |
| Frequency of interaction | .521 | -.643 | .561 |
| Feeling of belonging | .980 | -.002 | -.197 |
| Physical Proximity | .519 | .656 | .549 |
| Formality of Relationship | .981 | .004 | -.193 |

知道個體如何分離並不能表示哪個變數對應到哪個維度。但是使用成份負荷量就可以做到這點。第一個維度主要對應到歸屬感、交互作用強度和關係正式性，第二個維度會將交互作用頻率和實際接觸分開，第三個維度會將它們彼此分開。

## 範例：飲食失調症狀學

飲食失調是一種消耗性疾病，與飲食行為失序有關，它是一種嚴重的身體意象扭曲，也是對體重的一種迷思，而同時影響到心理和體重。每年都有數百萬的人受到影響，其中又特別以青少年的危險性較高。如果能及早發現症狀，便能及早診治，成效也最好。

健康專業人員會嘗試透過心理評估和健康檢查來診斷出飲食失調，但卻很難將病患分類到數個不同的飲食失調等級中，因為對於厭食/暴食的行為，目前並沒有一個標準化的症狀學。是不是有一些症狀可以清楚的將病患區分成 4 個組別呢？而這些病患又有哪些共用的症狀？

為了試著回答這些問題，研究人員(Van der Ham, Meulman, Van Strien, 和 Van Engeland, 1997)對 55 個已知有飲食失調的少年進行研究，結果如下表所示。

表格 10-2
病患診斷

| 診斷 | 病患號碼 |
|---|---|
| 厭食症 | 25 |
| 厭食症合併暴食症 | 9 |
| 厭食後暴食症 | 14 |
| 非典型飲食失調 | 7 |
| 總和(T) | 55 |

每個病患在四年之中被訪問四個回合，所以得到總數為 220 的觀察值。在每次的訪察中，會依照下表所列的 16 個症狀，將病患評分。目前遺漏了第二次訪察的病患 71，第二次訪察的病患 76，以及第三次訪察的病患 47 的症狀分數，因此只剩下 217 個有效觀察值。資料存放在 anorectic.sav 中。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

**表格 10-3**
**修改後的 Morgan-Russell 子量尺，可用來測量健康**

| 變數名稱 | 變數標記(L) | 低分區（分數 1） | 高分區（分數 3 或 4） |
|---|---|---|---|
| weight | 體重 | 正常範圍之外 | 標準 |
| mens | 經期 | 經期不順 | 經期規律 |
| fast | 限制攝食（禁食中） | 低於 1200 卡路里 | 正常/規律進餐 |
| binge | 暴飲暴食 | 一週多於一次 | 未暴飲暴食 |
| vomit | 催吐 | 一週多於一次 | 未催吐 |
| purge | 通便 | 一週多於一次 | 未通便 |
| hyper | 過動 | 無法安靜 | 未過動 |
| fami | 家庭關係 | 貧窮 | 小康 |
| eman | 家庭束縛力 | 非常依賴 | 適當 |
| frie | 朋友 | 沒有好朋友 | 兩個或多個好朋友 |
| school | 學校/僱用記錄 | 未就學/就業 | 記錄良好 |
| satt | 性態度 | 不足 | 適當 |
| sbeh | 性行為 | 不足 | 愉悅 |
| mood | 心理狀態（情緒） | 非常沮喪 | 標準 |
| preo | 注重食物和體重 | 完整(C) | 漠不關心 |
| body | 身體知覺 | 失常 | 標準 |

在此狀況下適合使用主成份分析，因為研究的目的是為了弄清症狀和不同飲食失調等級之間的關係。此外，類別主成份分析可能比典型主成份分析更適用，因為症狀是以次序尺度所評分的。

## 執行分析

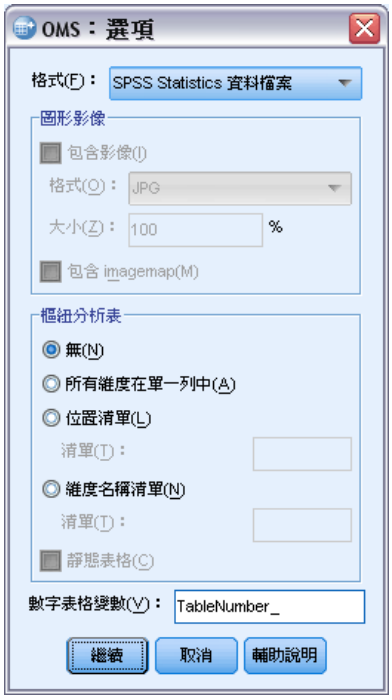
為了能適當地為每個診斷檢驗出病程的結構，您要產生出可用的投影重心圖表，作為散佈圖的資料。使用「輸出管理系統」便可達成此目的。

▶ 若要開啟 OMS 要求，請從功能表中選擇：
**公用程式(U) > OMS 控制台(O)...**

圖表 10-18
輸出管理系統控制台

- 選取「表格」作為輸出類型。
- 選取 CATPCA 作為指令。
- 選取「投影重心」作為表格類型。
- 選取「輸出目的地」群組中的檔案，並輸入 projected_centroids.sav 作為檔名。
- 按一下「選項」。

圖表 10-19
「選項」對話方塊



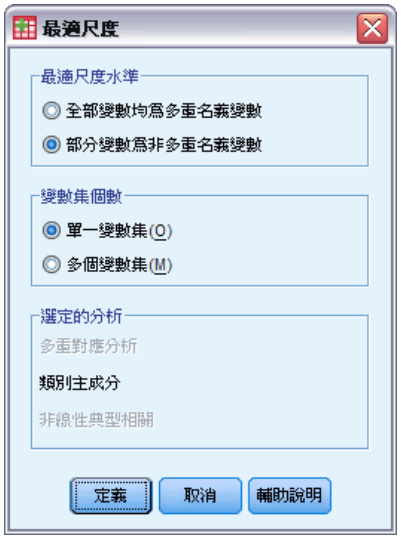

- 選取「IBM SPSS Statistics資料檔」作為輸出格式。
- 輸入 TableNumber_1 作為表格數變數。
- 按一下「繼續」。

圖表 10-20
輸出管理系統控制台

- 按一下「新增」。
- 按一下「確定」，再按一下「確定」來確認 OMS 階段作業。

  「輸出管理系統」現在已設定完成，並將投影重心表的結果寫入 projected_centroids.sav 檔中。
- 如需產生此資料集的類別主成份輸出，請從功能表中選擇：
  分析(A) > 維度縮減(D) > 最適尺度...

圖表 10-21
「最適尺度」對話方塊




- 選擇「最適尺度水準」組別中的「部份變數為非多重名義」。
- 按一下「定義」。

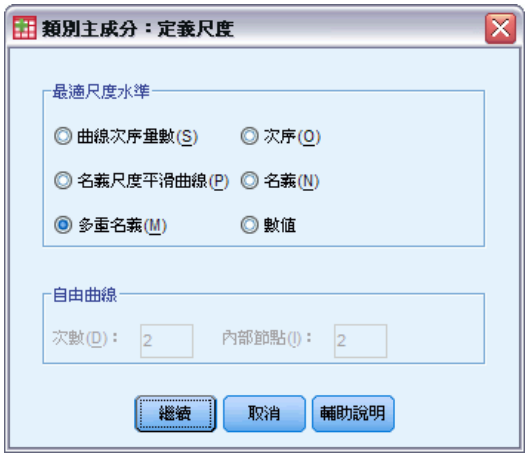
圖表 10-22
「類別主成分」對話方塊

- 選取「體重」到「身體知覺」作為分析變數。
- 按一下定義尺度與加權。

圖表 10-23
「定義尺度與加權」對話方塊

- 選取「次序」作為最適尺度水準。
- 按一下「繼續」。
- 選取「次數/診斷交互作用」作為補充變數，並按一下「類別主成份」對話方塊中的「定義尺度」。

圖表 10-24
「定義尺度」對話方塊




- 選取「多重名義量數」作為最適尺度水準。
- 按一下「繼續」。

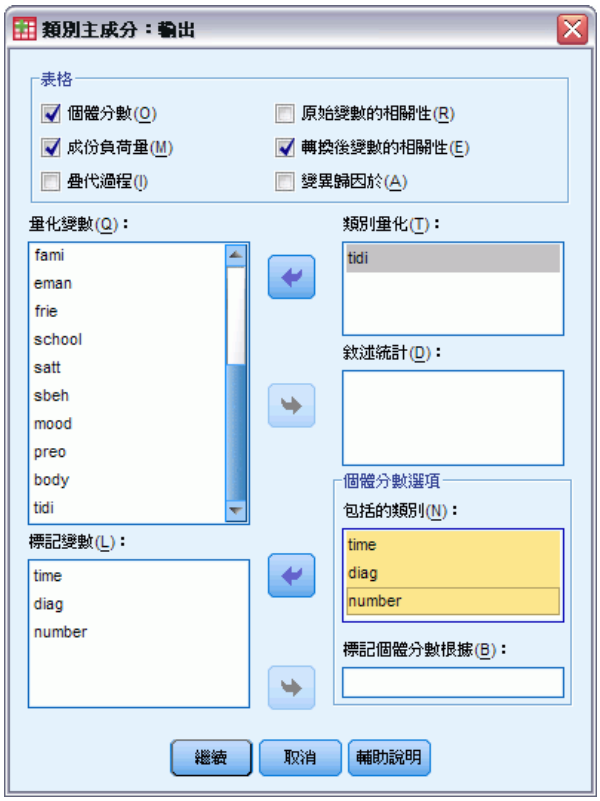
圖表 10-25
「類別主成分」對話方塊

- 選取「訪察次數」到「病患號碼」作為標記變數。
- 按一下「選項」。

圖表 10-26
「選項」對話方塊

- 選擇按照「變數名稱或數值」來標記圖形。
- 按一下「繼續」。
- 在「類別主成份」對話方塊中，按一下「輸出」。

圖表 10-27
「輸出」對話方塊




- 在「表格」組別中選取「個體分數」。
- 要求 tidi 的類別量化。
- 選擇以包括 time、diag 和 number 的類別。
- 按一下「繼續」。
- 在「類別主成份」對話方塊中，按一下「儲存」。

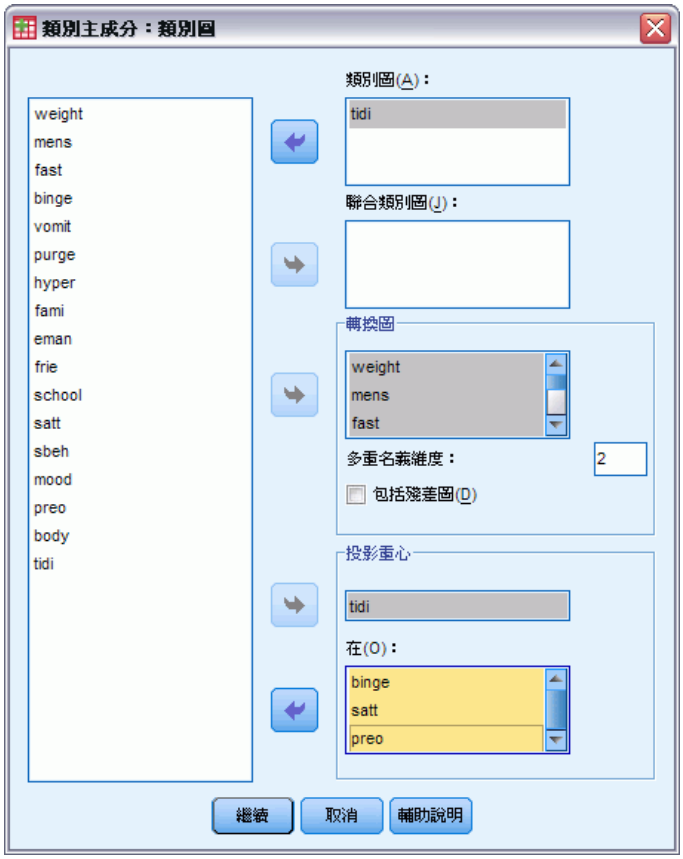
圖表 10-28
「儲存」對話方塊

- 在「轉換變數」組別中，選取「儲存至作用中資料集」。
- 按一下「繼續」。
- 在「類別主成份」對話方塊中，按一下「個體」。

圖表 10-29
「個體與變數圖」對話方塊

- 選擇按照變數來標記個數。
- 選取 time 和 diag 作為標記個體所使用的變數。
- 按一下「繼續」。
- 在「類別主成份」對話方塊中，按一下「類別」。

圖表 10-30
「類別圖」對話方塊




- 要求 tidi 的類別圖。
- 要求 weight 到 body 的轉換圖。
- 選擇在 binge、satt 和preo 上投影 tidi 的重心。
- 按一下「繼續」。
- 在「類別主成份」對話方塊中，按一下「確定」。

此程序會產生受試者的分數（具平均數 0 和單位變異數）和類別量化，可使受試著分數和轉換變數的相關平均數平方最大化。在目前的分析中，會限制類別量化以反映次序資訊。

最後，為了將投影重心表資訊寫入 projected_centroids.sav 中，您必須結束 OMS 要求。叫回「OMS 控制台」。

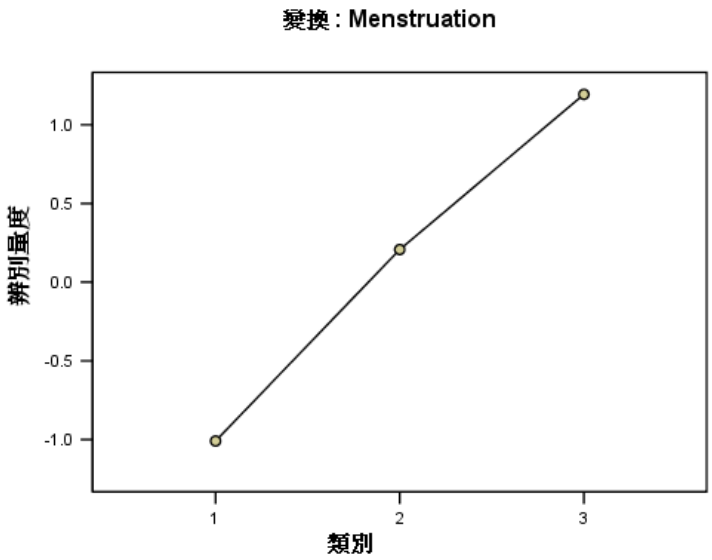
圖表 10-31
輸出管理系統控制台

- 按一下「結束」。
- 按一下「確定」，再按一下「確定」來確認。

## 轉換圖

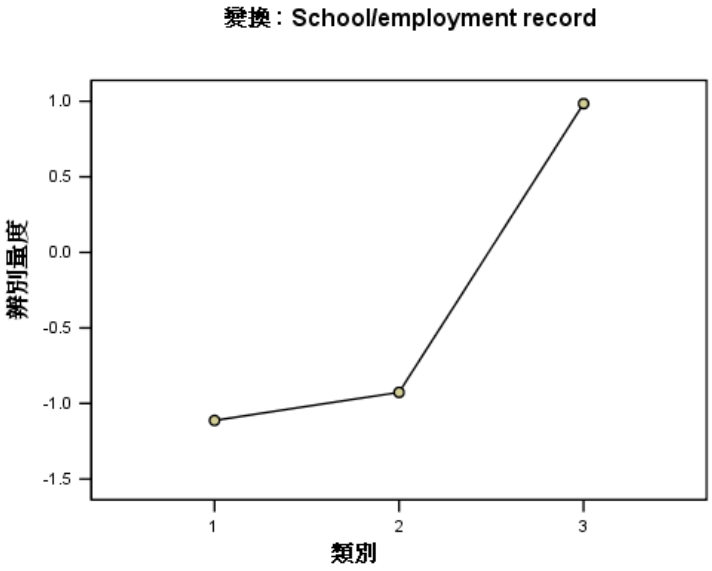
轉換圖會在水平軸上顯示原始類別數，垂直軸則提供最適量化。

圖表 10-32
「經期」轉換圖



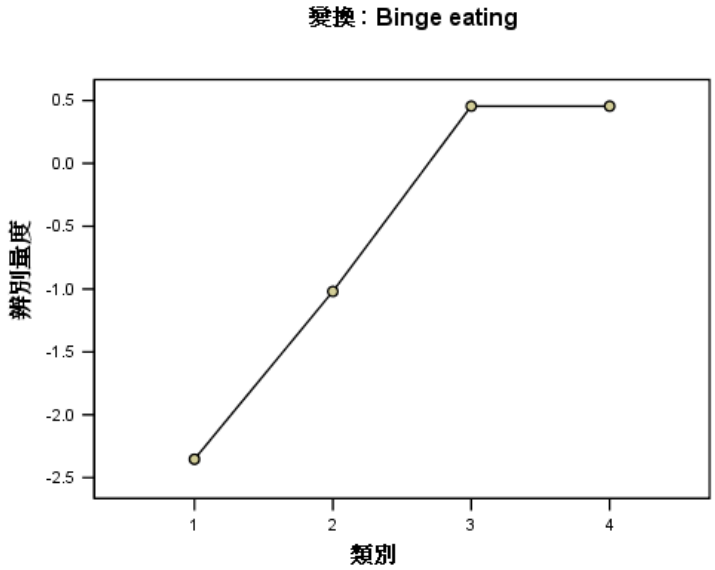

有些變數，例如經期，會得到幾乎線性的轉換，因此在此分析中，您可以將它們解譯為數值。

圖表 10-33
「學校/僱用記錄」轉換圖




其他變數，如學校/僱用記錄的量化不會得到線性轉換，故應該在次序尺度水準上解譯。第二個和第三個類別間的差異，比第一個和第二個類別之間的差異還重要的多。

圖表 10-34
「暴飲暴食」轉換圖




在暴飲暴食的量化中產生一個有趣的觀察值。類別 1 到類別 3 取得線性轉換，但類別 3 和 4 的量化值一樣。此結果顯示病患間類別 3 和類別 4 的分數是沒有差異的，故建議您將類別 4 的分數記錄為類別 3 的分數，以雙成份解來使用數值尺度水準。

## 模式摘要

圖表 10-35
模式摘要(M)

| 維度 | Cronbach's Alpha 值 | 變異數歸因於 | |
|---|---|---|---|
| | | 總和 (特徵值) | 變異數的 % |
| 1 | .874 | 5.550 | 34.690 |
| 2 | .522 | 1.957 | 12.234 |
| 總數 | .925[a] | 7.508 | 46.924 |

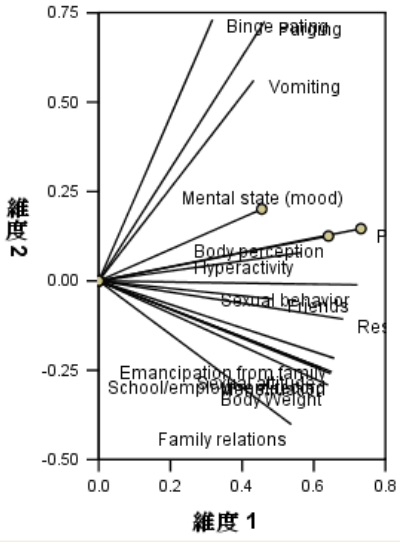
a. Cronbach's Alpha 平均值是以平均特徵值爲準。

若想了解模式對資料的配適度，請檢閱模式摘要。約有 47% 的總差異可由雙成份模式解釋，35% 由第一個維度解釋，12% 由第二個維度解釋。因此在各個體水準上，幾乎有一半的變化性可由雙成份模式來解釋。

## 成份負荷量

為了開始解釋解中的兩個維度，請先檢閱成份負荷量。所有變數在第一個維度中都有正成份負荷量，這表示有共同的因子與所有的變數正相關。

圖表 10-36
成份負荷量圖




第二個維度可區分變數。暴飲暴食、催吐和通便這三個變數集束在第二維度中，具有高的正負荷量。這些特徵一般便視為是暴食行為的代表。

無家庭束縛、學校/僱用記錄、性態度、體重和經期等變數構成另一個配套，您可以在此配套中包含限制攝食（禁食中）和家庭關係，因為他們的向量很接近主集群，且這些變數會視為厭食特徵（禁食、體重、經期），或是社會心理特性（束縛性、學校/工作記錄、性態度、家庭關係）。此配套的向量與 binge、vomit 和 purge 的向量是正交的（相互垂直），這表示這個變數集與暴食變數集無相關。

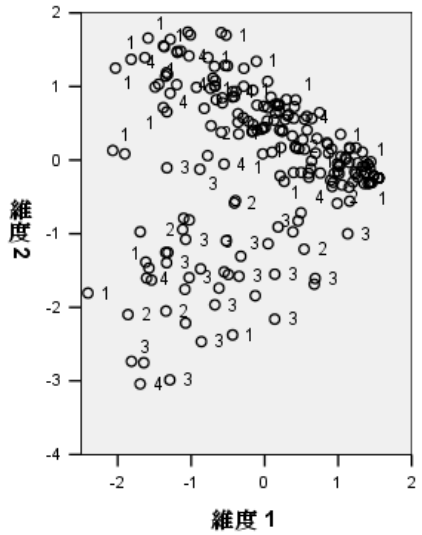
朋友、心理狀態（情緒）和 過動 似乎不是非常適合該解。您可以在圖中的每個向量的長度觀察到這個現象。所指定變數其向量的長度會符合其配適度，且這些變數的向量最短。根據雙成份解，您或許可以將這些變數從飲食失調的建議症狀學中剔除。但是，這些變數在較高的維度解中可能更配適。

性行為、注重食物和體重和身體知覺變數構成另一組理論症狀組別，這一組與病患對其身體的體驗有關。當與變數的兩個正交配套相關時，這些變數有相當長的向量，而且與第一個維度緊密相關，因此可提供有關“共同”因子的有用資訊。

## 物件分數

以下數字顯示個體分數的圖形，其中會以診斷類別來標記受訪者。

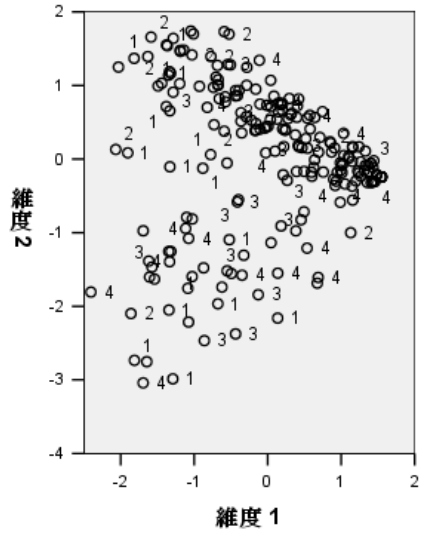
圖表 10-37
按照診斷標記個體分數圖




此圖並無助於解釋第一個維度，因為病患並非沿著第一個維度的診斷來區分。但是，其中仍有一些與第二個維度有關的資訊。厭食的受訪者 (1) 和具有典型飲食失調的病患 (4) 組成一組，位在構成暴食的受訪者和病患 (2 和 3) 之上。因此，第二個維度會將暴食病患與其他病患分隔，如同您在上一節所見 (在第二個維度中暴食配套的變數具有高正成份負荷量)。這是有道理的，因為傳統上與暴食相關症狀，在第二個維度中會有高成份負荷量值。

此數字顯示個體分數的圖形，其中會以診斷次數來標記受訪者。

圖表 10-38
以訪察次數來標記個體分數



按照次數來標記個體，顯示出第一個維度與次數有關，因為其中似乎與診斷次數的進度有關，從第一次訪察時開始大部分集中在左側，其餘的向右進展。請注意，您可以連接此圖形中的次數點，方法是先儲存個體分數，然後使用 x軸上的維度 1 分數和 y 軸上的維度 2 分數來建立散佈圖，再使用病患號碼設定標記。

將以次數標記的個體分數圖與以診斷標記的圖加以比較，可察覺出異常的個體。例如在以次數標記的圖中，有一個病患他第 4 回合的診斷是落在圖形中所有其他點的左邊。這就是異常，因為這些點一般的傾向是：愈後面次數的落點會往右。有趣的是，落點似乎不在其應有位置上的這個次數，其診斷也異常，因為該病患有厭食行為，但其分數落在暴食集群的病患中。再看一下個體分數表，您會發現這是病患 43，他已經被診斷出有厭食症，他的個體分數如下表所示。

**表格 10-4**
**病患 43 的個體分數**

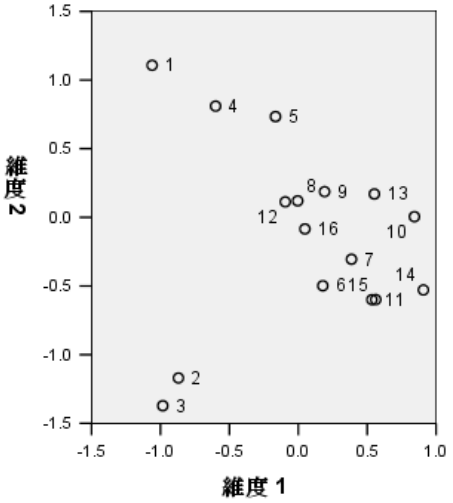
| 時間 | 維度 1 | 維度 2 |
|---|---|---|
| 1 | -2.031 | 1.250 |
| 2 | -2.067 | 0.131 |
| 3 | -1.575 | -1.467 |
| 4 | -2.405 | -1.807 |

該病患在第 1 次的分數是典型的厭食，在維度 1 中的負分數高，對應到身體不好的形象，在維度 2 的分數為正，對應到厭食症狀或是不良的社會心理行為。但是，與大部分病患不同的是，在維度 1 中只有一點或是沒有進展。在維度 2 中，有一些明顯朝向「正常」的進展（約為 0，介於厭食與暴食行為之間），但該病患變成展現暴食症狀。

## 檢驗病程結構

若要更了解兩個維度與四個診斷類別和四個次數點的關係，您可以將次數/診斷交互作用建立為補充變數，建立的方式是將病患診斷的四個類別和訪察次數的四個類別交叉分類。因此，次數/診斷交互作用有 16 個類別，其中的第一個類別是指第一次訪察時的厭食症病患。 第五個類別是指第二次訪察時的厭食症病患，以此類推，第六個類別是指第四次訪察時的典型飲食失調病患。使用次數/診斷交互作用補充變數可研究在一段時間內不同組別的病程。對於變數提供了多重名義尺度水準，而各類別點也顯示以下的數字。

圖表 10-39
次數/診斷交互作用的類別點




從此圖形中，有些結構是很明顯的：在第二個維度，第一次訪察的診斷類別清楚地區分厭食症和典型飲食失調以及厭食症合併暴食症和在厭食症後發生暴食症。之後，還要看出其他的圖樣就有點困難了。

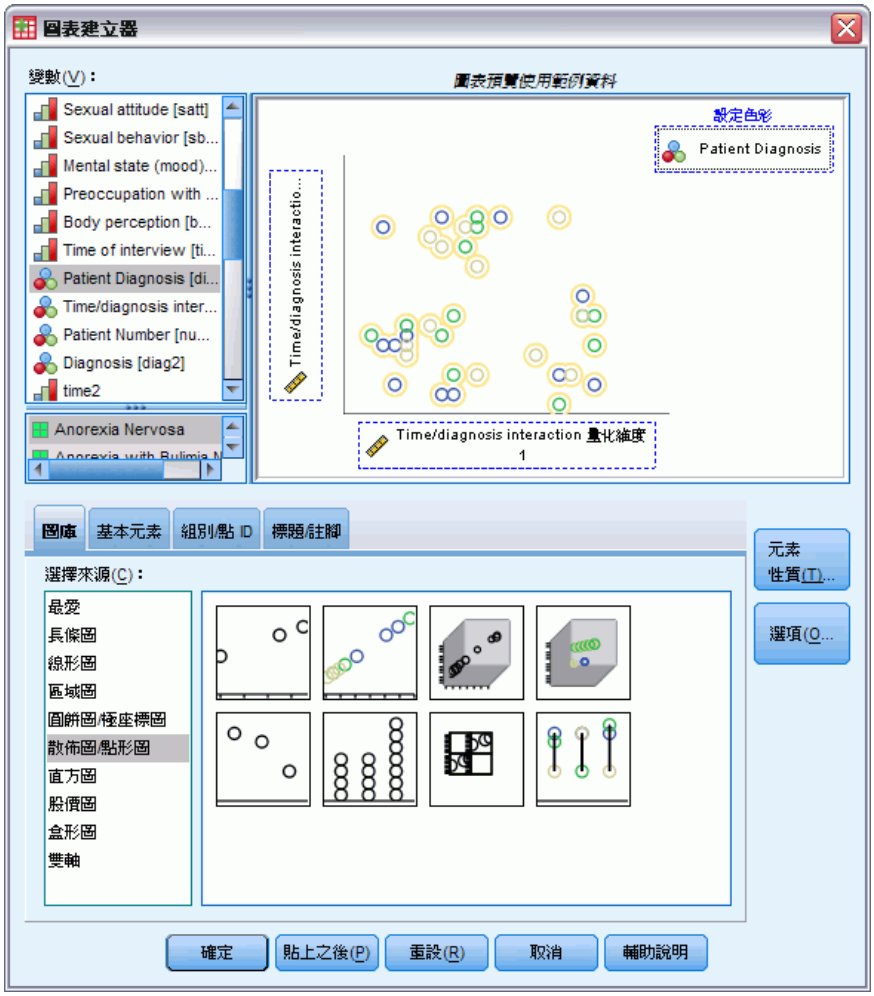
但是，您可以根據量化建立一個散佈圖，讓圖樣更明顯可見。若要進行此功能，在功能表中選擇：

**統計圖(G) > 圖表建立器(C)...**

圖表 10-40
散佈圖/點形圖圖庫

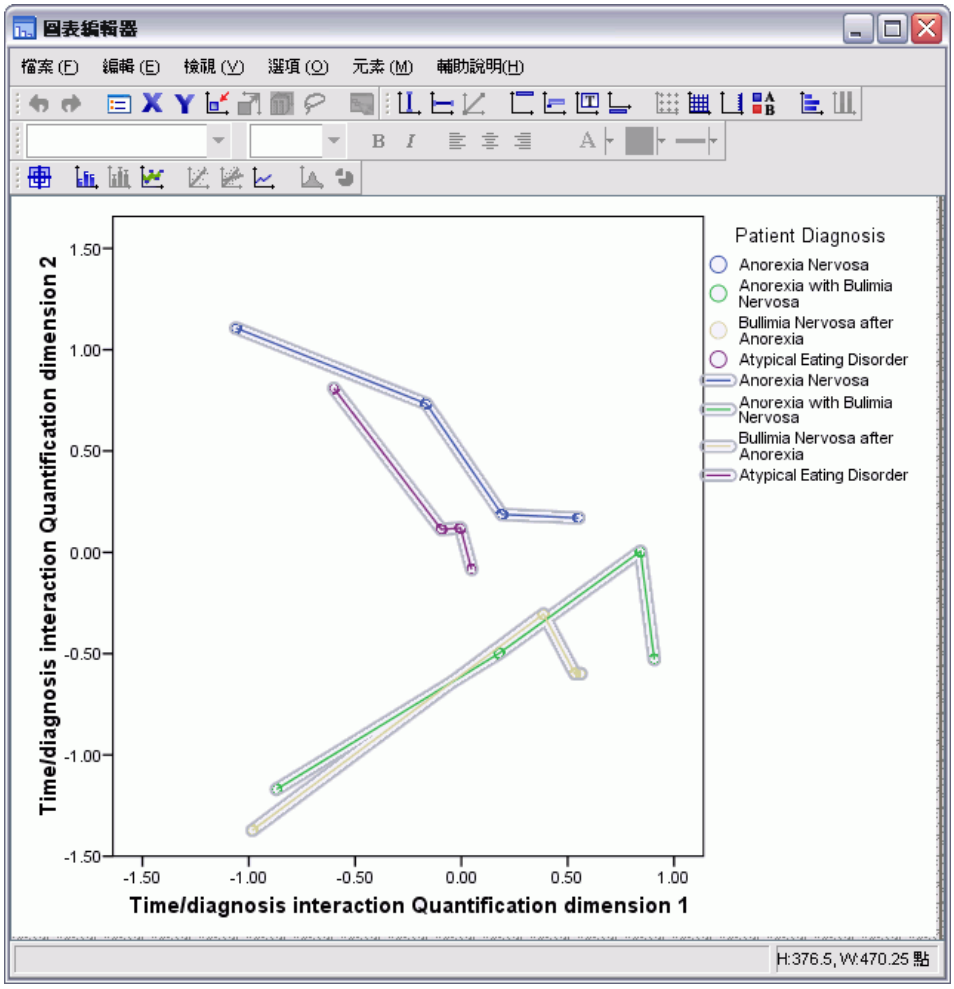
► 選取「**散佈圖/點形圖**」圖庫並選擇「分組散佈圖」。

圖表 10-41
圖表建立器



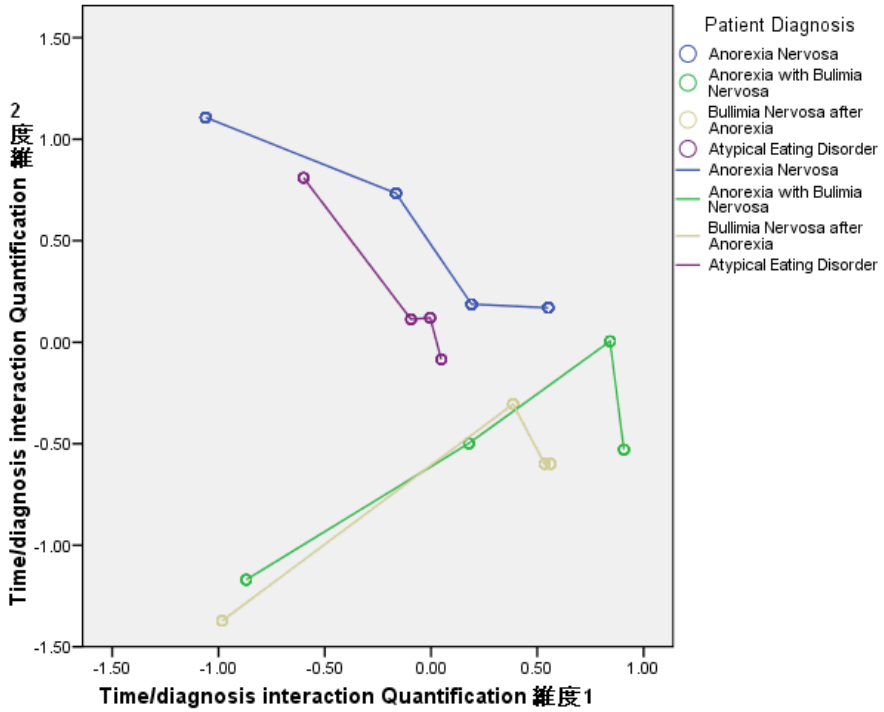

- 選取「次數/診斷交互作用量化維度 2」作為 y 軸變數，和「次數/診斷交互作用量化維度 1」作為 x 軸變數。
- 選擇按照「病患診斷」來設定顏色。
- 按一下「**確定**」。

圖表 10-42
病程結構




- 然後，若要將這些點連接，請在圖表上連按兩下，再按一下「圖表編輯程式」中的**新增內插法線條**工具。
- 關閉「圖表編輯程式」。

圖表 10-43
病程結構




透過將一段時間內每個診斷類別的類別點連接，則所產生的圖樣立即顯示出第一個維度與時間、第二個維度、診斷有關，如同您先前從個體分數圖中的判定相同。

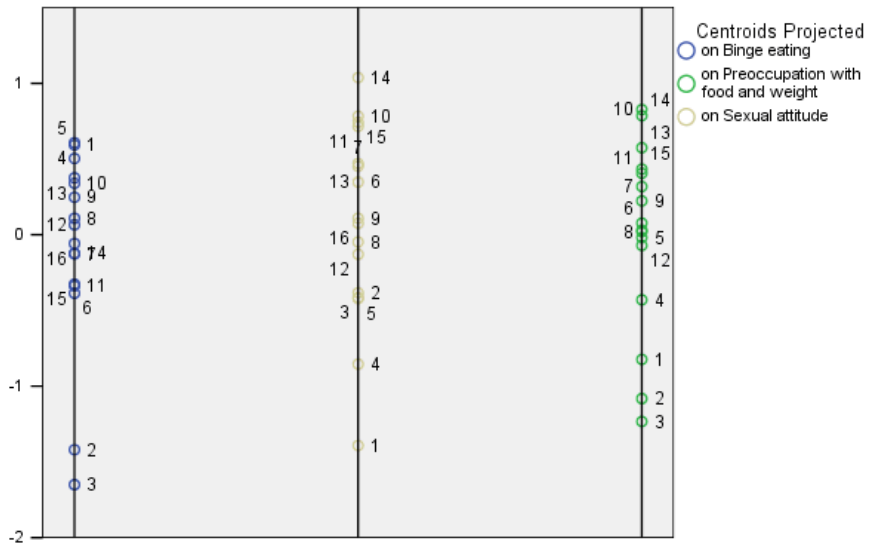
但是，此圖更進一步的顯示經過一段時間，該病的傾向會愈趨相同。此外，對於所有的組別，第一次和第二次訪察之間的進度最大；厭食症病患在第二次到第三次訪察之間顯示有些進度，但其他組別只顯示一點點的進度。

## 所選變數的微分發展

在每個症狀配套中，由成份負荷量所識別的一個變數會被選為該配套的“代表”。在暴食配套中會選取「暴飲暴食」；在厭食/社會心理配套中會選取「性態度」；在第三個配套中，則是「身體知覺」。

為了檢驗出病程的可能差異，所以會計算出暴飲暴食、性態度和注重食物和體重上的次數/診斷交互作用投影，並使用以下數字繪製圖形。

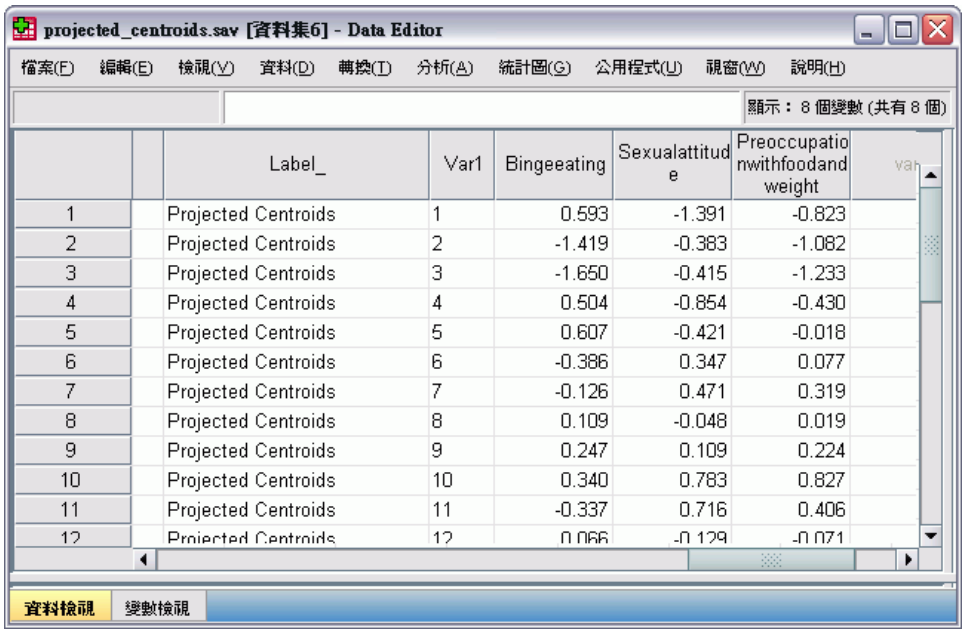
**圖表 10-44**
**「暴飲暴食」、「性態度」和「注重食物和體重」上的「次數/診斷交互作用」投影重心。**




此圖顯示在第一次訪察時，暴飲暴食的症狀會將暴食病患 (2 和 3) 與其他 (1 和 4) 區分開來，性態度會將厭食病患與非典型病患 (1 和 4) 與其他 (2 和 3) 區分開來，而注重食物和體重則沒有很確實的區分病患。在許多的應用上，此圖便足以說明症狀和診斷間的關係，但由於次數多且複雜，圖案因此變得不清楚了。

為了能檢視經過一段時間後的這些投影狀況，您必須要能夠繪製出投影重心表的內容。透過 OMS 要求，將此資訊儲存至 projected_centroids.sav，可能可以繪製出。

**圖表 10-45**
Projected_centroids.sav




| | Label_ | Var1 | Bingeeating | Sexualattitude | Preoccupationwithfoodandweight | var |
|---|---|---|---|---|---|---|
| 1 | Projected Centroids | 1 | 0.593 | -1.391 | -0.823 | |
| 2 | Projected Centroids | 2 | -1.419 | -0.383 | -1.082 | |
| 3 | Projected Centroids | 3 | -1.650 | -0.415 | -1.233 | |
| 4 | Projected Centroids | 4 | 0.504 | -0.854 | -0.430 | |
| 5 | Projected Centroids | 5 | 0.607 | -0.421 | -0.018 | |
| 6 | Projected Centroids | 6 | -0.386 | 0.347 | 0.077 | |
| 7 | Projected Centroids | 7 | -0.126 | 0.471 | 0.319 | |
| 8 | Projected Centroids | 8 | 0.109 | -0.048 | 0.019 | |
| 9 | Projected Centroids | 9 | 0.247 | 0.109 | 0.224 | |
| 10 | Projected Centroids | 10 | 0.340 | 0.783 | 0.827 | |
| 11 | Projected Centroids | 11 | -0.337 | 0.716 | 0.406 | |
| 12 | Projected Centroids | 12 | 0.066 | -0.129 | -0.071 | |

Bingeeating、Sexualattitude 和 Preoccupationwithfoodandweight 這三個變數包含在每個感興趣症狀上的投影重心值。觀察值號碼 (1 到 16) 會對應到次數/診斷交互作用。您必須計算出新的變數，這些變數可將「次數」和「診斷」值分離出。

▶ 從功能表選擇：
**轉換(T) > 計算變數(C)...**

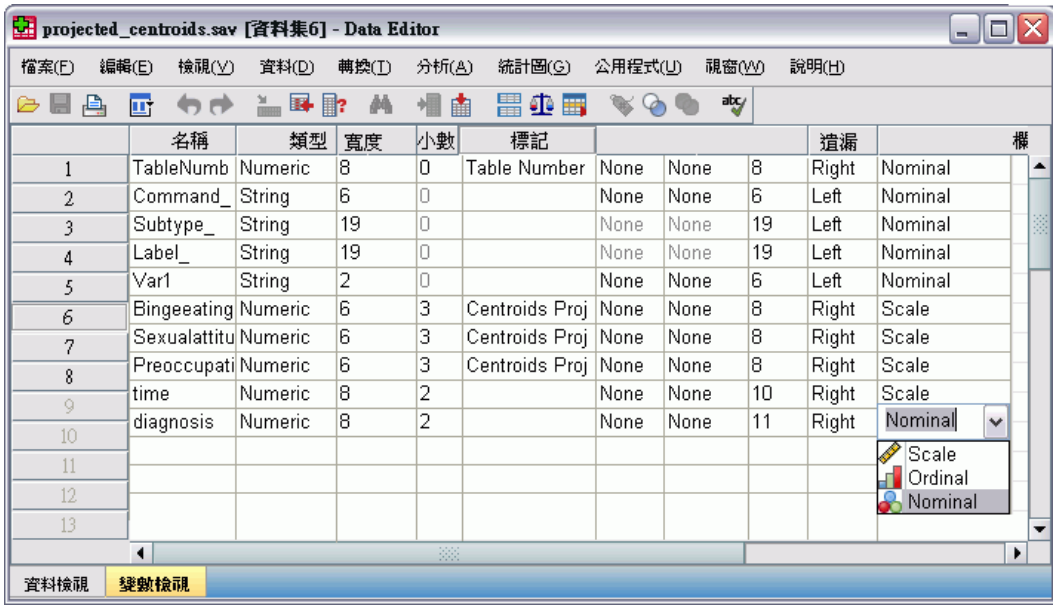
圖表 10-46
計算變數對話方塊

- ▶ 輸入 time 作為目標變數。
- ▶ 輸入 trunc(($casenum-1)/4) + 1 作為數值運算式。
- ▶ 按一下「確定」。

圖表 10-47
計算變數對話方塊

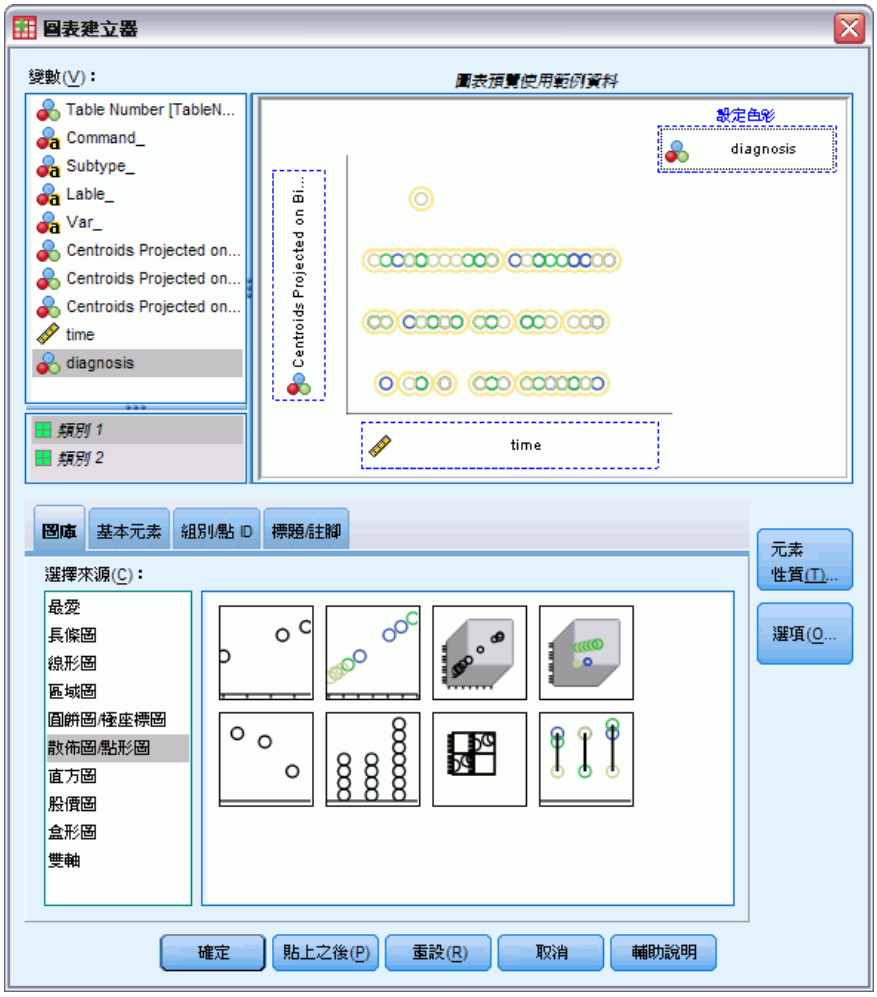
- ▶ 叫回「計算變數」對話方塊。
- ▶ 輸入 diagnosis 作為目標變數。
- ▶ 輸入 mod($casenum-1, 4) + 1 作為數值運算式。
- ▶ 按一下「確定」。

圖表 10-48
Projected_centroids.sav



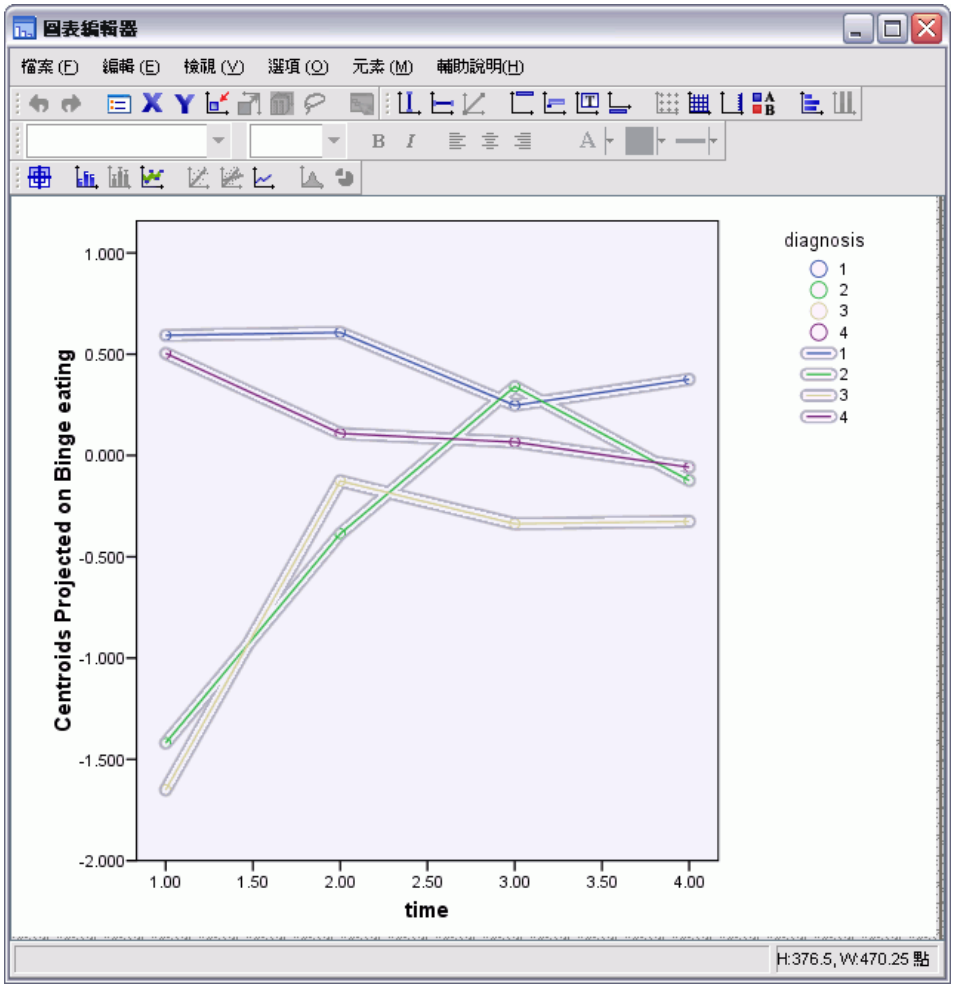

在「資料檢視」中，將diagnosis的測量由「尺度」變更為「名義」。

圖表 10-49
圖表建立器




- 最後，若要檢視一段時間內暴飲暴食的診斷次數投影重心，請叫出「圖表建立器」，並按一下「**重設**」，將您先前的選擇清除。
- 選取「**散佈圖/點形圖**」圖庫並選擇「分組散佈圖」。
- 選取「在暴飲暴食上投影的重心」作為 y 軸變數，選取「次數」作為 x 軸變數。
- 選擇按照「診斷」來設定顏色。
- 按一下「**確定**」。

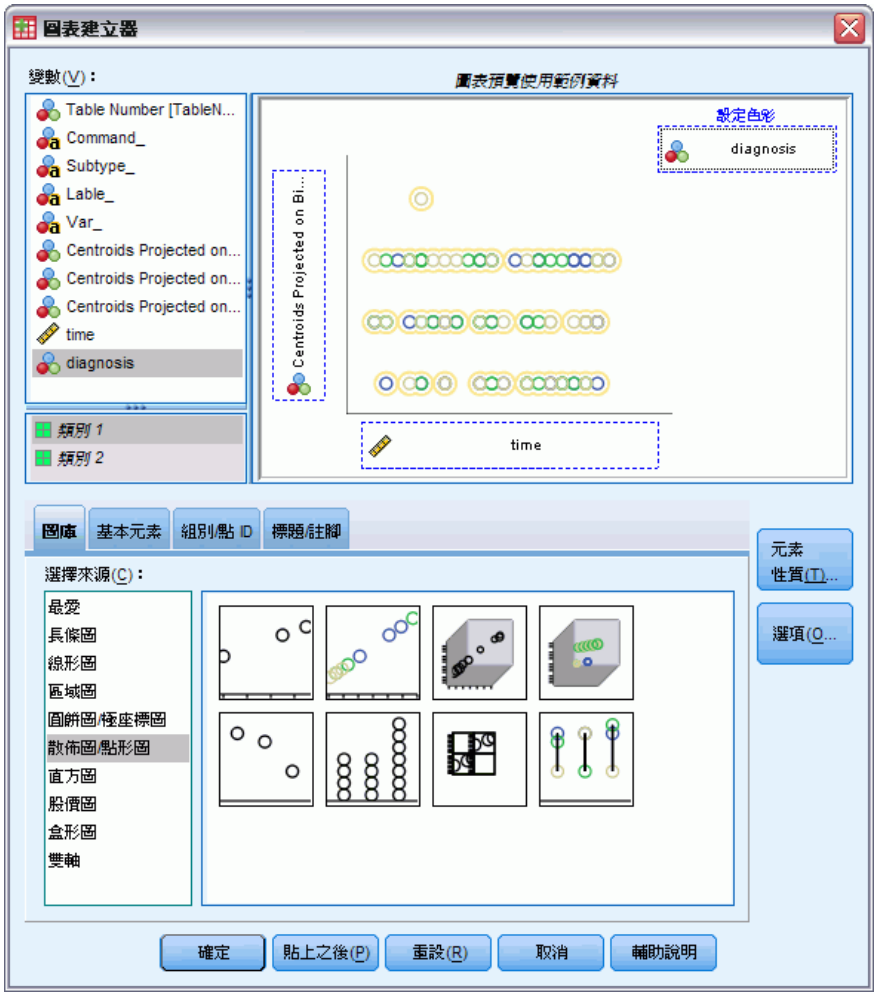
圖表 10-50
在一段時間內，在「暴飲暴食」上的「診斷次數」投影重心




- 然後，若要將這些點連接，請在圖表上連按兩下，再按一下「圖表編輯程式」中的**新增內插法線條**工具。
- 關閉「圖表編輯程式」。

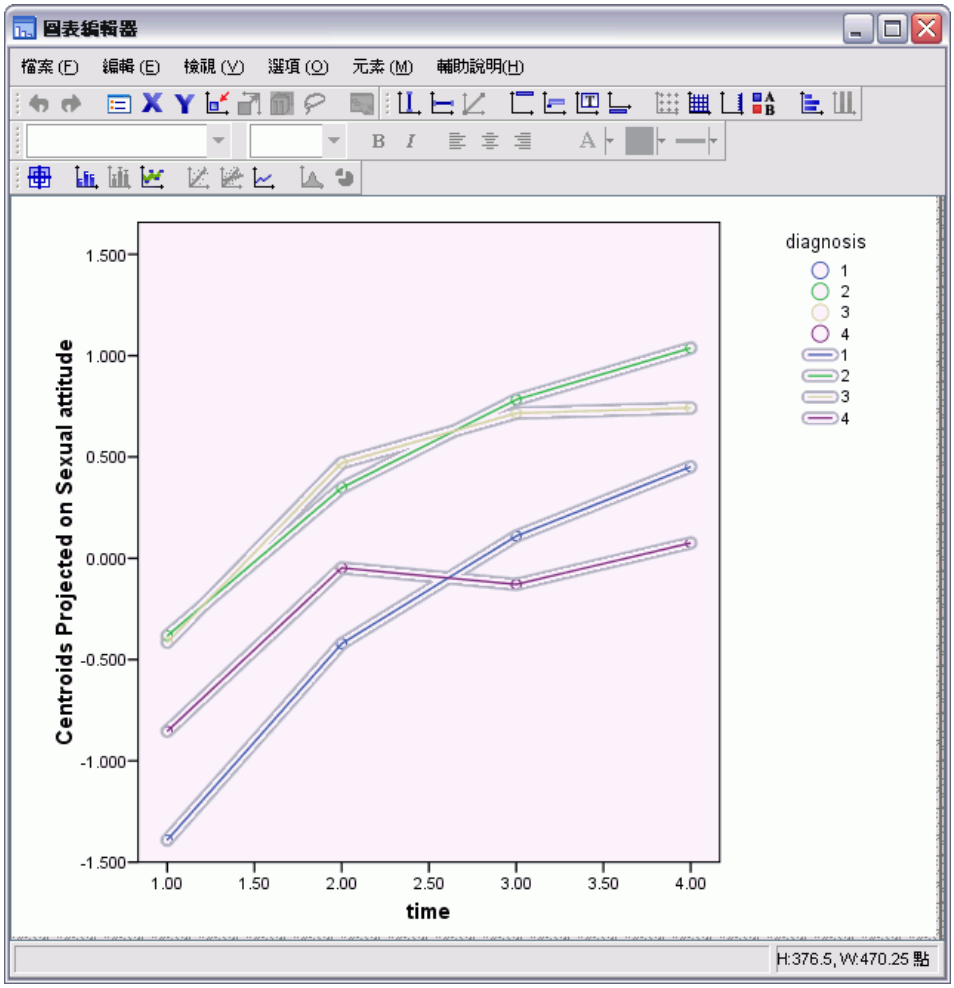
關於暴飲暴食，很明顯的厭食組別開始值與暴食組別的不同。此差異隨著時間而縮小，厭食組別幾乎無變更，而暴食群組卻顯示出變化進度。

圖表 10-51
圖表建立器




- 叫出「圖表建立器」。
- 取消選取「在暴飲暴食上投影的重心」作為 y 軸變數，並選取「在性態度上投影的重心」作為 x 軸變數。
- 按一下「確定」。

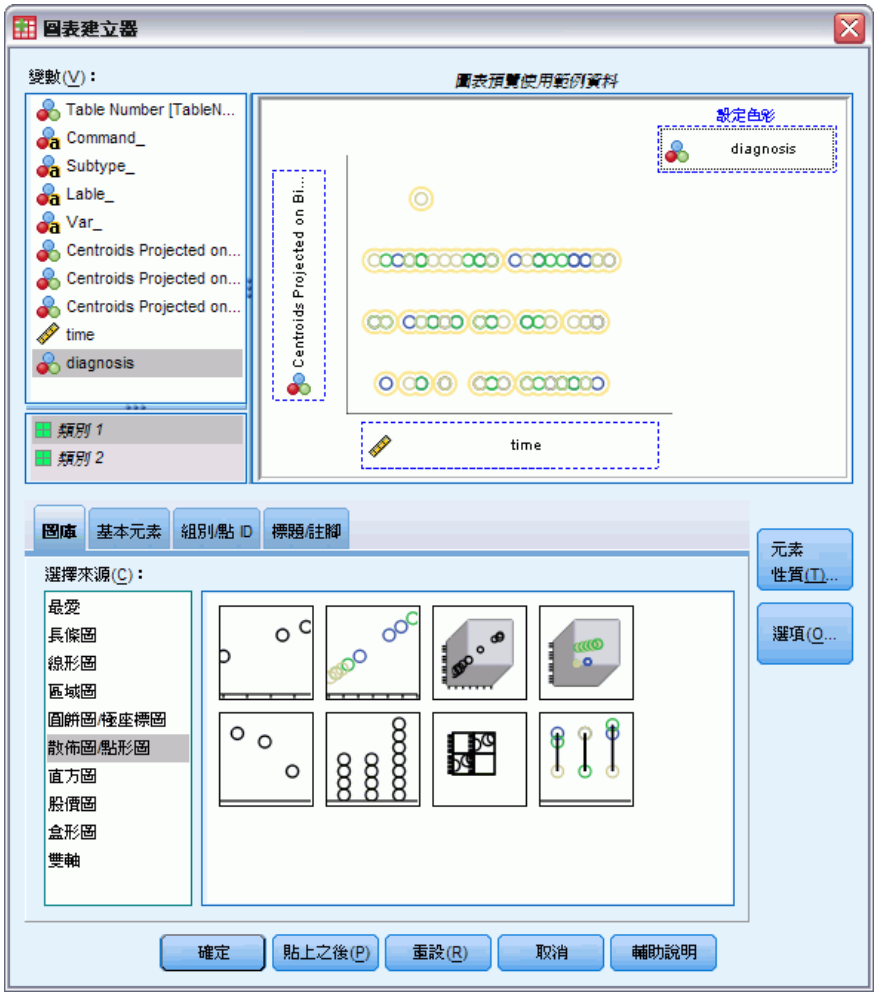
圖表 10-52
在一段時間內，在「性態度」上的「診斷次數」投影重心




- 然後，若要將這些點連接，請在圖表上連按兩下，再按一下「圖表編輯程式」中的**新增內插法線條**工具。
- 關閉「圖表編輯程式」。

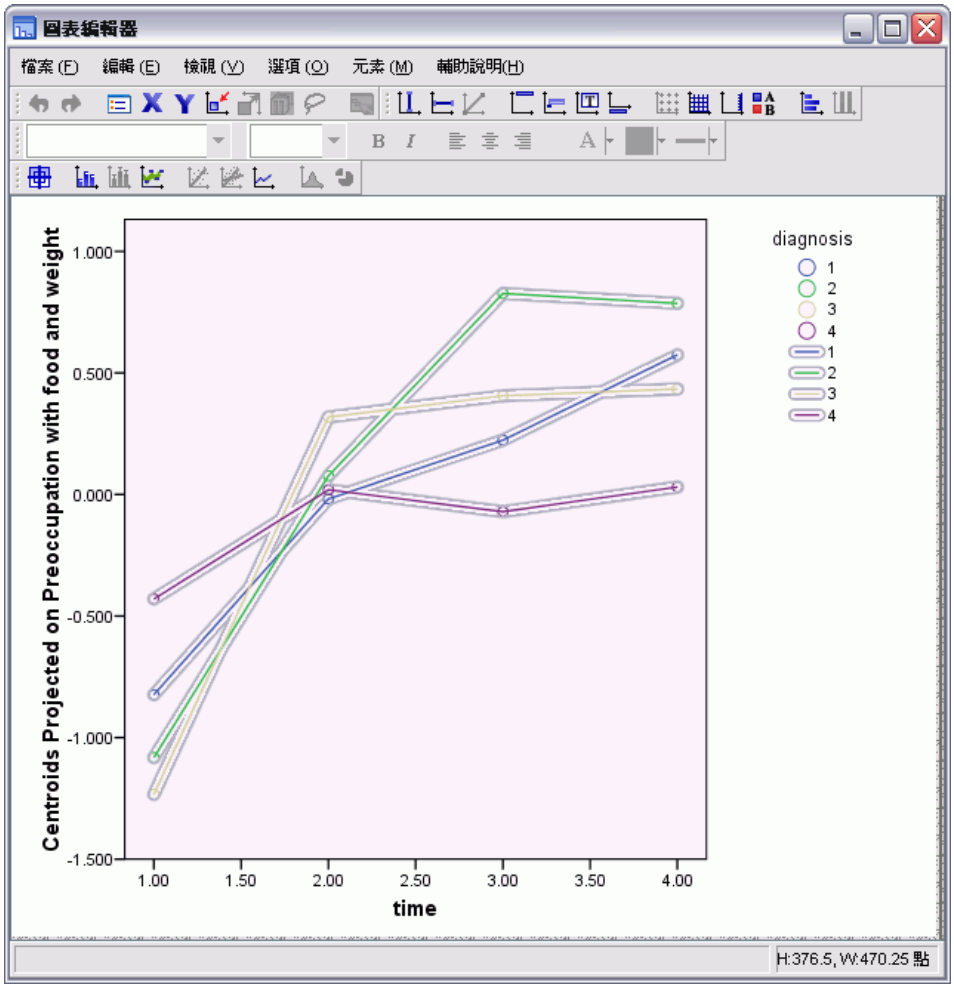
有關於性態度，在經過一段時間後，四個軌線多多少少有些平行，而且所有組別都顯示了進度。但是，暴食組別比厭食組別的分數高（較佳）。

圖表 10-53
圖表建立器




- 叫出「圖表建立器」。
- 取消選取「在性態度上投影的重心」作為 y 軸變數，並選取「在注重食物和體重上投影的重心」作為 x 軸變數。
- 按一下「確定」。

圖表 10-54
在一段時間內，在「身體知覺」的「診斷次數」投影重心




- 然後，若要將這些點連接，請在圖表上連按兩下，再按一下「圖表編輯程式」中的**新增內插法線條**工具。
- 關閉「圖表編輯程式」。

身體知覺是代表核心症狀的變數；核心症狀由四個不同的組別分攤。除了非典型飲食失調病患之外，該厭食組別和兩個暴食組別在開始端和結尾端有非常類似的水準。

# 閱讀資料推薦

請參閱下列文字以取得類別主成份分析的詳細資訊：

De Haas, M., J. A. Algera, H. F. J. M. Van Tuijl, 和 J. J. Meulman. 2000. Macro and micro goal setting: (微觀及宏觀目標設定)In search of coherence (尋求一貫性). Applied Psychology (應用心理學), 49, .

De Leeuw, J. 1982. Nonlinear principal components analysis (非線性主成份分析). 於： COMPSTAT Proceedings in Computational Statistics (計算統計學 COMPSTAT 會議記錄), Vienna (維也納): Physica Verlag.

Eckart, C., 和 G. Young. 1936. The approximation of one matrix by another one of lower rank (一個矩陣及另一個低等級矩陣的近似法). Psychometrika (心理學計量報導), 1, .

Gabriel 檢定(G), K. R. 1971. The biplot graphic display of matrices with application to principal components analysis (矩陣行列散佈圖形顯示與其在主成份分析之應用). Biometrika (生物統計學雜誌), 58, .

Gifi, A. 1985. PRINCALS。Research Report UG-85-02 (研究報告 UG-85-02). Leiden (萊頓): Department of Data Theory, University of Leiden (萊頓大學資料理論系).

Gower, J. C., 和 J. J. Meulman. 1993. The treatment of categorical information in physical anthropology (體質人類學之類別資訊處理). International Journal of Anthropology (國際人類學刊), 8, .

Heiser, W. J., 和 J. J. Meulman. 1994. Homogeneity analysis: (均齊性分析)Exploring the distribution of variables and their nonlinear relationships (探索變數分配及其非線性關係). 於： Correspondence Analysis in the Social Sciences: (社會學對應分析)Recent Developments and Applications (近期發展與應用), M. Greenacre, 和 J. Blasius，eds. 紐約: Academic Press (學術出版社).

Kruskal, J. B. 1978. Factor analysis and principal components analysis: (因子分析與主成份分析)Bilinear methods (雙線性方法). 於： International Encyclopedia of Statistics (國際統計學百科), W. H. Kruskal, 和 J. M. Tanur，eds. 紐約: The Free Press (自由出版社).

Kruskal, J. B., 和 R. N. Shepard. 1974. A nonmetric variety of linear factor analysis (線性因子分析之非計量變化). Psychometrika (心理學計量報導), 39, .

Meulman, J. J. 1993. Principal coordinates analysis with optimal transformations of the variables: (最適變數轉換之主座標分析)Minimizing the sum of squares of the smallest eigenvalues (將最小特徵值平方和最小化). British Journal of Mathematical and Statistical Psychology (英國數學暨統計心理學雜誌), 46, .

Meulman, J. J., 和 P. Verboon. 1993. Points of view analysis revisited: (再論觀點分析)Fitting multidimensional structures to optimal distance components with cluster restrictions on the variables (以集群限制的變數將多元結構配適至最適距離成份). Psychometrika (心理學計量報導), 58, .

Meulman, J. J., A. J. Van der Kooij, 和 A. Babinec. 2000. New features of categorical principal components analysis for complicated data sets, including data mining (複雜資料集類別主成份分析之新功能 - 含資料採礦). 於： Classification, Automation and New Media (分類、自動化與新媒體), W. Gaul, 和 G. Ritter，eds. 柏林: Springer-Verlag.

Meulman, J. J., A. J. Van der Kooij, 和 W. J. Heiser. 2004. Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data (次序及非次序資料非線性最適尺度轉換之主成分分析). 於： Handbook of Quantitative Methodology for the Social Sciences (社會學量化方法論手冊), D. Kaplan，ed. Thousand Oaks, Calif. (加州千橡): Sage Publications, Inc..

Theunissen, N. C. M., J. J. Meulman, A. L. Den Ouden, H. M. Koopman, G. H. Verrips, S. P. Verloove-Vanhorick, 和 J. M. Wit. 2003. Changes can be studied when the measurement instrument is different at different time points (在不同時間點依量測工具改變來研究變更). Health Services and Outcomes Research Methodology (健康服務與結果研究方法論), 4, .

Tucker, L. R. 1960. Intra-individual and inter-individual multidimensionality (個體內與個體間多元性). 於： Psychological Scaling: (心理學尺度法)Theory & Applications (理論與應用), H. Gulliksen, 和 S. Messick，eds. 紐約: John Wiley and Sons.

Vlek, C., 和 P. J. Stallen. 1981. Judging risks and benefits in the small and in the large (判斷小規模及大規模之風險與惠益). Organizational Behavior and Human Performance (組織行為與人員績效), 28, .

Wagenaar, W. A. 1988. Paradoxes of gambling behaviour (賭博行為之弔詭). 倫敦: Lawrence Erlbaum Associates, Inc.

Young, F. W., Y. Takane, 和 J. De Leeuw. 1978. The principal components of mixed measurement level multivariate data: (混合測量水準多變量資料之主成份)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 43, .

Zeijl, E., Y. te Poel, M. du Bois-Reymond, J. Ravesloot, 和 J. J. Meulman. 2000. The role of parents and peers in the leisure activities of young adolescents (青少年休閒活動中父母與友伴的角色). Journal of Leisure Research (休閒研究雜誌), 32, .

章節
11

# 非線性典型相關分析

非線性典型相關分析的目的是為了判斷兩個或多個變數集與另一變數集之間的相似程度。其目標與線性典型相關分析相同，是為了儘量說明低維度空間中各集合之間的差異性。但和線性典型相關不同的是，非線性典型相關不會假定測量水準間隔，或假設關係是線性的。還有另外一個重要的不同，就是非線性典型相關分析會透過將每個集合中變數的線性組合，同時與未知集（個體分數）進行比較，來建立集合間的相似性。

## 範例：調查結果分析

本章的範例取自 (Verdegaal, 1985) 調查。在調查中記錄了來自 15 個受訪者對 8 個變數的回應。資料集中的變數、變數標記和數值標記（類別）如下表所示。

**表格 11-1**
**調查資料**

| 變數名稱 | 變數標記 | 數值標記 |
|---|---|---|
| age | 年齡（年為單位） | 20 - 25, 26 - 30, 31 - 35, 36 - 40, 41 - 45, 46 - 50, 51 - 55, 56 - 60, 61 - 65, 66 - 70 |
| marital | 婚姻狀況 | 單身、已婚、其他 |
| pet | 有養寵物 | 無、貓、狗、其他（非貓或狗）、各種家畜 |
| news | 最常閱讀的報紙 | 無、Telegraaf、Volkskrant、NRC、其他 |
| music | 喜愛的音樂 | 古典、新浪潮、流行、各種音樂、不喜愛音樂 |
| live | 喜愛的居住環境 | 城市、郊區、鄉下 |
| math | 數學測驗分數 | 0 - 5, 6 - 10, 11 - 15 |
| language | 語言測試分數 | 0 - 5, 6 - 10, 11 - 15, 16 - 20 |

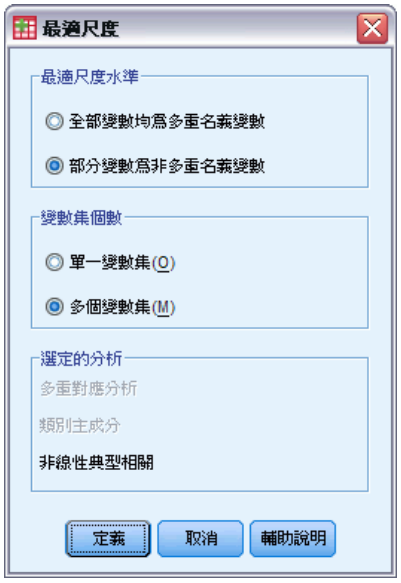
可以在 verd1985.sav 中找到此資料集。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案. 前 6 個變數是要注意的變數，它們分成 3 個集合。集 1 包括 age 和 marital，集 2 包括 pet 和 news，集 3 包括 music 和 live。Pet 調整為多重名義量數，age 調整為次序量數，所有其他的變數調整為單一名義量數。此分析需要隨機起始組態。根據預設，該起始組態是數值。但是，當有些變數被視為單一名義量數，且沒有排序的可能性時，最好選擇隨機起始組態。這是此研究中大部分變數的狀況。

© Copyright SPSS Inc. 1989, 2010

## 檢驗資料

- 若要得到此資料集的非線性典型相關分析，從功能表中選擇：
  **分析 > 維度縮減 > 最適尺度...**

圖表 11-1
「最適尺度」對話方塊

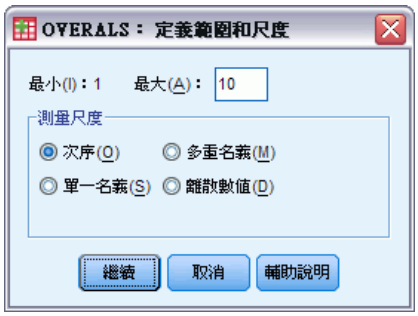



- 選擇「最適尺度水準」組別中的「**部份變數為非多重名義**」。
- 選擇「變數集個數」組別中的「**多重集合**」。
- 按一下「**定義**」。

圖表 11-2
「非線性典型相關分析」對話方塊

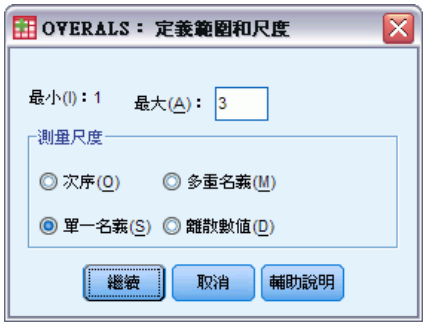
- 選擇「年齡（年為單位）」和「婚姻狀況」作為第一個集合的變數。
- 選擇「age」並按一下「定義範圍與尺度」。

圖表 11-3
「定義範圍與尺度」對話方塊




- 輸入 10 作為此變數的最大值。
- 按一下「繼續」。
- 在「非線性典型相關分析」對話方塊中，選擇「marital」，並按一下「定義範圍與尺度」。

圖表 11-4
「定義範圍與尺度」對話方塊




- 輸入 3 作為此變數的最大值。
- 選擇「單一名義量數」作為測量尺度。
- 按一下「繼續」。
- 在「非線性典型相關分析」對話方塊中按一下「下一步」來定義下一個變數集。

圖表 11-5
「非線性典型相關分析」對話方塊

- 選擇「有養寵物」和「最常閱讀的報紙」作為第二個集合的變數。
- 選擇「pet」並按一下「定義範圍與尺度」。

圖表 11-6
「定義範圍與尺度」對話方塊

- 輸入 5 作為此變數的最大值。
- 選擇「多重名義量數」作為測量尺度。
- 按一下「繼續」。
- 在「非線性典型相關分析」對話方塊中，選擇「news」，並按一下「定義範圍與尺度」。

圖表 11-7
「定義範圍與尺度」對話方塊

- 輸入 5 作為此變數的最大值。
- 選擇「單一名義量數」作為測量尺度。
- 按一下「繼續」。
- 在「非線性典型相關分析」對話方塊中按一下「下一步」來定義最後一個變數集。

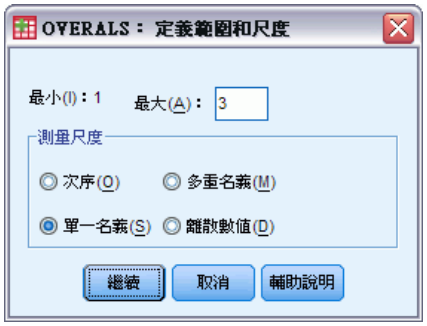
圖表 11-8
「非線性典型相關分析」對話方塊

- 選擇「喜愛的音樂」和「喜愛的居住環境」作為第三個集合的變數。
- 選擇「music」並按一下「**定義範圍與尺度**」。

圖表 11-9
「定義範圍與尺度」對話方塊

- 輸入 5 作為此變數的最大值。
- 選擇「**單一名義量數**」作為測量尺度。
- 按一下「繼續」。
- 在「非線性典型相關分析」對話方塊中，選擇「live」，並按一下「**定義範圍與尺度**」。

圖表 11-10
「定義範圍與尺度」對話方塊




- 輸入 3 作為此變數的最大值。
- 選擇「**單一名義量數**」作為測量尺度。
- 按一下「**繼續**」。
- 在「非線性典型相關分析」對話方塊中，按一下「**選項**」。

圖表 11-11
「選項」對話方塊

- 在「顯示」組別中，取消選擇「**重心**」，並選擇「**加權值及成份負荷量**」。
- 在「圖形」組別中，選擇「**類別重心圖**」和「**轉換**」。
- 選擇「**使用隨機起始組態**」。
- 按一下「**繼續**」。
- 在「非線性典型相關分析」對話方塊中，按一下「**確定**」。

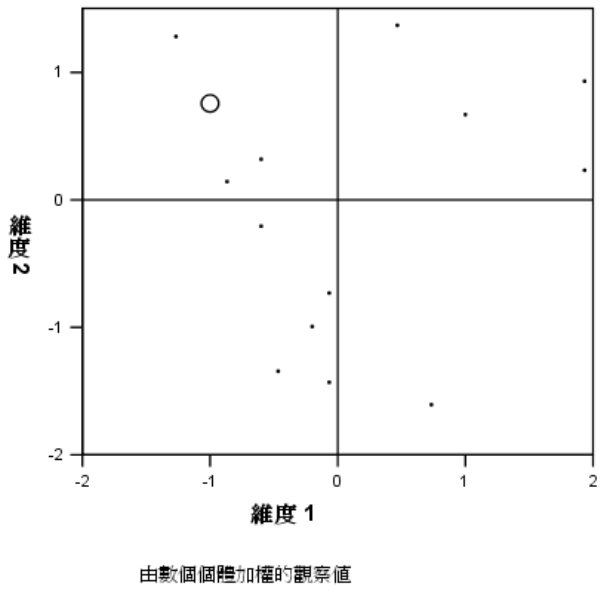
具有最適尺度的類別典型相關分析會在列出變數及其最適尺度水準的清單之後，產生一個表格，其中顯示類別中個體的頻率。如果有遺漏值的話，此表格特別的重要，因為幾近空白的種類較可能主控解答。在此例中，沒有遺漏值。

第二個初步檢查是檢驗偏離值的個體分數圖。偏離值與其他個體在量化上有如此的差異，使偏離值位在圖形的邊界，也因此主控一或多個維度。

如果您發現偏離值，您可以使用以下兩個方法來處理。只要將偏離值從資料中除去，並再執行一次非線性典型相關分析。或者是，您可以嘗試將這些類別合併，來記錄偏離個體的極端回應。

如個體分數圖所示，調查資料中沒有偏離值。

**圖表 11-12**
**個體分數**




## 說明集合之間的相似度

以非線性典型相關分析來測量集合之間相關性的方法有數種（每一個方法會在不同的表格或表格集中說明）。

### 分析摘要

適合度值和不適值可以告訴您，在集合間的關聯性方面，非線性典型相關分析解適合最適量化資料的程度。分析摘要表會顯示調查範例的適合度值、不適值和特徵值。

圖表 11-13
分析摘要

| | | 維度 | | |
|---|---|---|---|---|
| | | 1 | 2 | 總和 |
| 損失 | 集合 1 | .240 | .183 | .423 |
| | 集合 2 | .184 | .408 | .593 |
| | 集合 3 | .171 | .205 | .376 |
| | 平均數 | .199 | .265 | .464 |
| | 特徵值 | .801 | .735 | |
| | 配適 | | | 1.536 |

不適度是由整個維度和集合所分攤的。對於每個維度和集合，不適度代表個體分數中，集合內的變數加權組合無法說明的變異部分。不適度的平均數會標記為 Mean。在此範例中，整個集合的不適值平均數為 0.464。請注意，第二個維度所發生的不適值比第一維度的多。

每個維度的特徵值等於 1 減去維度的不適度平均數，並表示每個維度顯示多少個關係。特徵值的總和可達總適合度。對於 Verdegaal 的資料，第一個維度所說明的實際適合度為 0.801 / 1.536 = 52%。

最大適合度值等於維度數目，若有取得則表示關係很完美。整個集合與維度上的平均不適值可告訴您最大適合度和實際適合度的差異。適合度加上不適度平均數等於維度數目。完美的相似度幾乎不會發生，而且通常是應用在資料不重要的層面上。

另一個具兩個變數集合的常用統計方法是典型相關。由於典型相關與特徵值有關，因此未提供額外的資訊，但它不包含在非線性典型相關分析輸出。對於兩個變數集合而言，以下公式可取得每個維度的典型相關：

$$\rho_d = 2 \times E_d - 1$$

其中 d 是維度數目，E 是特徵值。

您可以使用以下公式來歸納兩個以上集合的典型相關：

$$\rho_d = ((K \times E_d) - 1)/(K - 1)$$

其中 d 是維度數目，K 是集合數目，E 是特徵值。就我們的範例而言，

$$\rho_1 = ((3 \times 0.801) - 1)/2 = 0.702$$

以及

$$\rho_2 = ((3 \times 0.735) - 1)/2 = 0.603$$

## 加權值及成份負荷量

另一個相關的量數是每個集合的線性組合和個體分數之間的多重相關。如果在集合中沒有變數是多重名義量數，您可以計算此量數，方法是將加權值與集合中每個變數的成份負荷量相乘，將乘積相加，然後取和的平方根。

圖表 11-14
加權值

| 集合 | | 維度 1 | 維度 2 |
|---|---|---|---|
| 1 | Age in years | .680 | .789 |
| | Marital status | .296 | -1.016 |
| 2 | Newspaper read most often | -.845 | -.361 |
| 3 | Music preferred | .631 | -.749 |
| | Neighborhood preference | -.484 | -.780 |

圖表 11-15
成份負荷量

| 集合 | | 維度 1 | 維度 2 |
|---|---|---|---|
| 1 | Age in years[a.,b] | .834 | .259 |
| | Marital status[b.,c] | .651 | -.604 |
| 2 | Pets owned[d.,e] 維度 1[b.,c] | .397 | -.431 |
| | 2[b.,c] | -.277 | .680 |
| | Newspaper read most often[b.,c] | -.667 | -.391 |
| 3 | Music preferred[b.,c] | .786 | -.500 |
| | Neighborhood preference[b.,c] | -.687 | -.540 |

a. 最適尺度水準：次序

b. 分析標的空間中單一量化變數的投影圖

c. 最適尺度水準：單一名義量數

d. 最適尺度水準：複名義量數

e. 分析標的空間中多重量化變數的投影圖

這些值是此範例中變數的加權值和成份負荷量。最適尺度變數（年齡（年為單位）和婚姻狀況）的第一個加權總和，與第一維度個體分數的多重相關 (R) 如下：

$$\begin{aligned} R &= \sqrt{(0.701 \times 0.841 + (-0.273 \times -0.631))} \\ &= \sqrt{(0.5895 + 0.1723)} \\ &= 0.873 \end{aligned}$$

對於每個維度，1 - 不適度 = $R^2$。例如，從分析摘要表中，1 - 0.238 = 0.762，其平方根為 0.873（加上一些捨入誤差）。結果，小的遺失度值表示最適尺度變數的加權總和與維度之間有大的多重相關。加權值並非多重名義量數所專有。對於多重名義量數變數，使用 1 - 每個集合的不適度。

### 分割適合度和不適度

每個集合的不適度會由非線性典型相關分析以數種方式分割。透過對調查範例進行非線性典型相關分析，適合度表可呈現分析後產生的多重適合度、單一適合度和單一不適度表。請注意，多重適合度減去單一適合度等於單一不適度。

圖表 11-16
分割適合度和不適度

| | | 多重適合度 | | | 單一適合度 | | | 單一損失 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 維度 | | 維度 | 維度 | | 維度 | 維度 | | 維度 |
| 集合 | | 1 | 2 | 總和 | 1 | 2 | 總和 | 1 | 2 | 總和 |
| 1 | Age in years[a] | .494 | .676 | 1.170 | .462 | .622 | 1.085 | .032 | .054 | .085 |
| | Marital status[b] | .089 | 1.033 | 1.122 | .088 | 1.033 | 1.120 | .001 | .000 | .001 |
| 2 | Pets owned[c] | .402 | .439 | .841 | | | | | | |
| | Newspaper read most often[b] | .724 | .187 | .911 | .714 | .130 | .844 | .010 | .057 | .067 |
| 3 | Music preferred[b] | .421 | .577 | .998 | .398 | .561 | .960 | .022 | .016 | .039 |
| | Neighborhood preference[b] | .234 | .609 | .843 | .234 | .608 | .843 | .000 | .000 | .000 |

a. 最適尺度水準：次序

b. 最適尺度水準：單一名義量數

c. 最適尺度水準：複名義量數

單一不適度表示不適是由將變數限制在一量化集合（也就是單一名義量數、次序量數或名義量數）上所產生的。如果單一不適度很大，最好能將變數視為多重名義量數。但在此範例中，單一適合度和多重適合度幾乎相同，這表示多個座標幾乎是在由加權值所指定方向的一直線上。

多重適合度等於每個變數的多重類別座標變異數。這些量數就好比是同質性分析中的區別量數。您可以檢驗多重適合度表，來看看哪一個變數最具有區別度。例如，請先看一下婚姻狀況和最常閱讀的報紙的多重適合度表。適合度值是兩個維度交叉後的摘要，也就是婚姻狀況的 1.122，和最常閱讀的報紙 的 0.911。此資訊告訴我們一個人的婚姻狀況比他最常閱讀的報紙更具區別度。

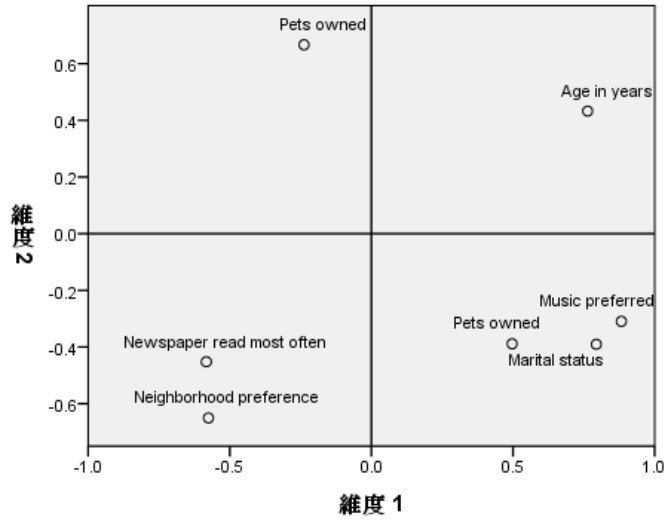
單一適合度對應到每個變數的平方加權值，而且等於單一類別座標的變數異。因此，加權值等於單一類別座標的標準差。透過檢驗單一適合度在各維度上的分散狀況，我們可以看到最常閱讀的報紙變數主要在第一維度上進行區別，而我們也可以看到婚姻狀況幾乎完全在第二個維度上區別。換句話說，最常閱讀的報紙在第一個維度上會比第二個維度更加分開，而婚姻狀況則呈現相反的圖樣。相反的，年齡（單位年）在第一個維度和第二個維度上都具有區別度，因此各類別沿著這兩個維度的分散狀況是一樣的。

## 成份負荷量

以下數字說明調查資料的成份負荷量圖。當沒有遺漏資料時，成份負荷量圖等於量化變數和個體分數之間的 Pearson 相關。

原點到每個變數點之間的距離為約該變數的重要性。在此圖形上並未繪製出典型變數，但會藉由穿過原點繪出的水平線和垂直線來表示。

圖表 11-17
成份負荷量



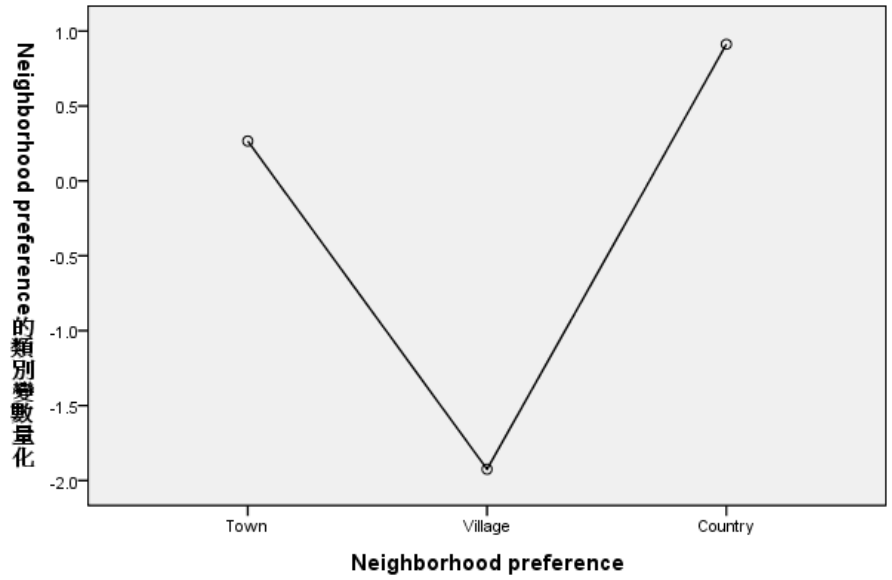

變數之間的關係是很明顯的。此外，圖中有兩個方向與水平軸和垂直軸方向不一致。一個方向是由年齡（年為單位）、最常閱讀的報紙和喜愛的居住環境所決定。另一個方向由婚姻狀況、喜愛的音樂和有養寵物所定義。有養寵物變數是多重名義量數，因此在圖形上繪製出兩個點。每個量化會解釋為單一變數。

## 轉換圖

可調整之每個變數的不同水準會對量化上套用限制。轉換圖說明了量化和選定的最適尺度水準所產生的原始類別兩者之間的關係。

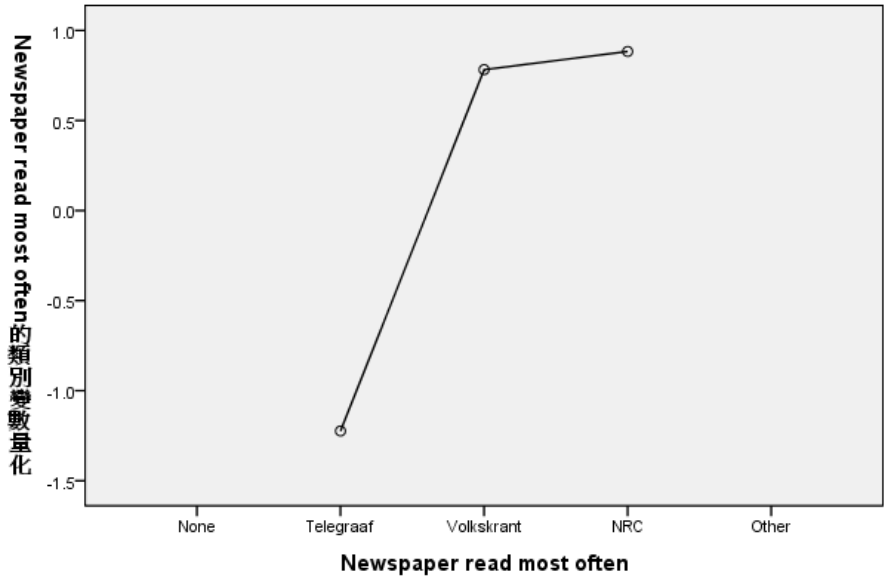
在喜愛的居住環境作為名義量數的轉換圖中，會顯示出一個 U 型的樣式，其中間類別的量化最低，兩端類別的值彼此類似。此樣式可表示原始變數與其轉換後變數之間的二次關係。在此並不建議使用喜愛的居住環境的另一個最適尺度水準。

圖表 11-18
「喜愛的居住環境（名義）」轉換圖

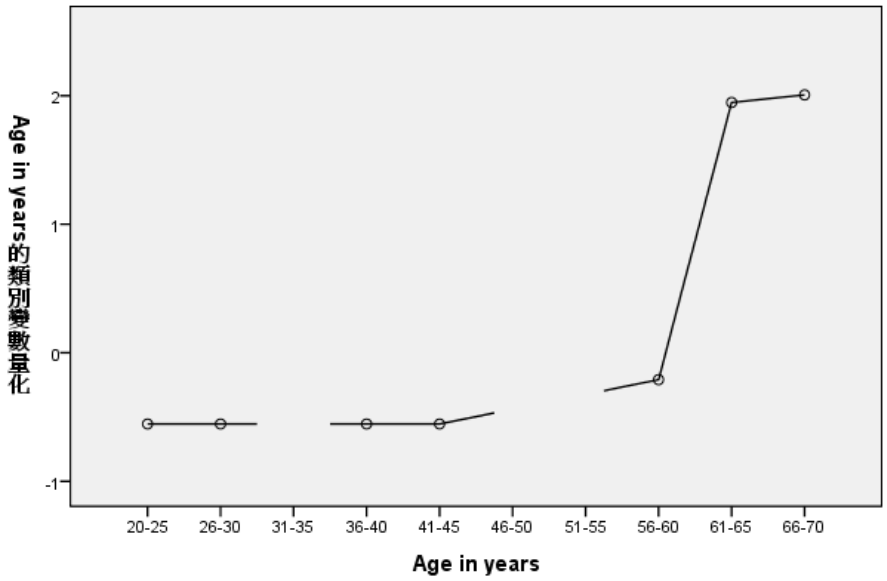



相反的，最常閱讀的報紙的量化，對應到三個包含所觀察的觀察值之類別的上升趨勢。第一個類別的量化值最低，第二個類別的量化值較高，第三個類別的量化值最高。雖然變數已調整為名義量數，但類別順序是從量化值中取得的。

圖表 11-19
「最常閱讀的報紙（名義）」轉換圖



圖表 11-20
「年齡（年為單位）」的轉換圖（次序）




年齡 (年為單位)的轉換圖顯示 S 狀曲線。前四個年齡最小的觀察類別都有相同的負量化值，而兩個年齡最大的類別有相似的正值。因此，可嘗試將所有年齡較小的類別合併成一個共同類別 (也就是低於 50)，並將兩個年齡最大的類別合併成一個類別。但是，幾個年青組別的量化值完全相等，這表示將量化值的順序限制為原始類別的順序可能不符合需求。 因為 26 – 30，36 – 40，41 – 45 組別的量化值不能小於 20 – 25 組別的量化值，所以這些值的設定會等於邊界值。讓這些值均小於最年青組別的量化值 (也就是將年齡視為名義量數)，可提高適合度。所以雖然年齡可視為次序變數，但在本例中將它視為次序變數是不適當的。此外，將年齡視為數值量數，並因此在類別之間維持相當的距離，會大大地降低適合度。

## 單一類別對多重類別座標

對於視為單一名義、次序或數值量數的每個變數，已決定了量化、單一類別座標和多重類別座標。系統會呈現年齡 (年為單位)的這些統計值。

圖表 11-21
「年齡 (年為單位)」的座標

| | | | 單一類別變數 | | 多重類別座標 | |
|---|---|---|---|---|---|---|
| | | | 維度 | | 維度 | |
| | 邊際次數 | 量化 | 1 | 2 | 1 | 2 |
| 20-25 | 3 | -.554 | -.377 | -.437 | -.192 | -.139 |
| 26-30 | 5 | -.554 | -.377 | -.437 | -.404 | -.623 |
| 31-35 | 0 | .000 | | | | |
| 36-40 | 1 | -.554 | -.377 | -.437 | -.318 | -.733 |
| 41-45 | 1 | -.554 | -.377 | -.437 | -.356 | -.534 |
| 46-50 | 0 | .000 | | | | |
| 51-55 | 0 | .000 | | | | |
| 56-60 | 2 | -.209 | -.142 | -.165 | -.435 | .087 |
| 61-65 | 1 | 1.947 | 1.324 | 1.536 | 1.710 | 1.204 |
| 66-70 | 2 | 2.006 | 1.364 | 1.583 | 1.215 | 1.711 |
| 遺漏 | 0 | | | | | |

每個沒有記錄觀察值的類別其量化值為 0。對於年齡 (年為單位)，這包括了 31 - 35、46 - 50 和 51 - 55 類別。這些類別未限制為與其他類別一起排序，且不會影響任何計算。

對於多重名義變數，每個類別在每個維度上有不同的量化值。對於所有其他轉換類型，無論解的維度狀況為何，一個類別只有一個量化值。每個單一類別座標集合代表類別位置在個體空間內一直線上的位置。所指定類別的座標等於乘上變數維度加權值的量化值。例如，在年齡 (年為單位)的表格，類別 56-60 的單一類別座標 (-0.142, -0.165) 是乘上維度加權值 (0.680, 0.789) 的量化值 (-0.209)。

對於視為單喔名義、次序或數值量數的變數，其多重類別座標代表在套用次序或線性限制之前，個體分數中的類別座標。這些數值是未限制的不適度最小化。對於多重名義變數，這些座標代表類別的量化值。

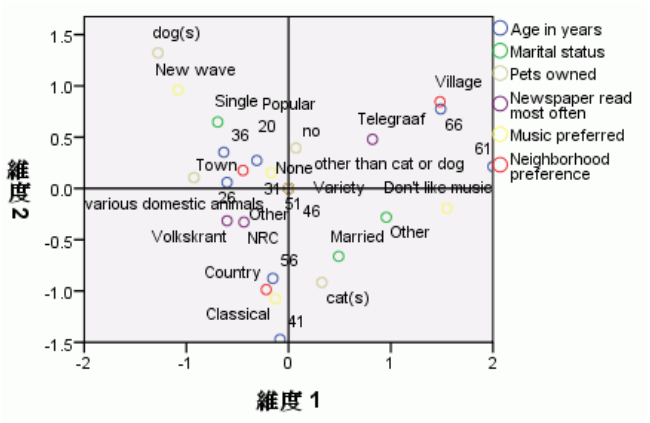
若在類別與其量化值之間的關係設定限制，則透過將單一類別座標與多重類別座標比較，則可顯示限制的效果。在第一個維度上，年齡 (年為單位)的多重類別座標會減少至類別 2，並相對地維持在同一個水準上，一直到類別 9 為止；在此類別中有大幅的增加。第二個維度證明了有類似的樣式。在套用尺度限制的單一類別維度中，會將這些關係移除。現在在這兩個維度上，座標是不會減少的。兩個維度集合的差異結構意謂著，將變數視為名義量數這種處理方式是適當的。

## 重心和投影重心

解釋以變數標記之重心圖的方式，與同質性分析中類別量化，或是非線性主成份分析中的多重類別座標的方式相同。這類圖形本身即表現出變數如何適當的分隔個體組別 (重心位在許多個體重力的中央)。

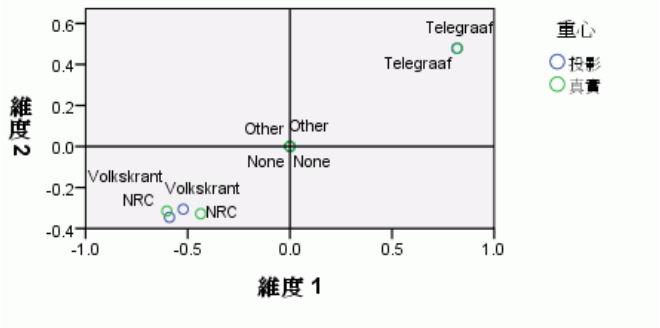
請注意，年齡 (年為單位)沒有非常明確的分隔。年齡較小的類別在圖形的左側分為一組。如先前所示，次序量數是非常嚴格的尺度水準，所以無法在年齡 (年為單位)上實行。

圖表 11-22
由變數標記的的重心



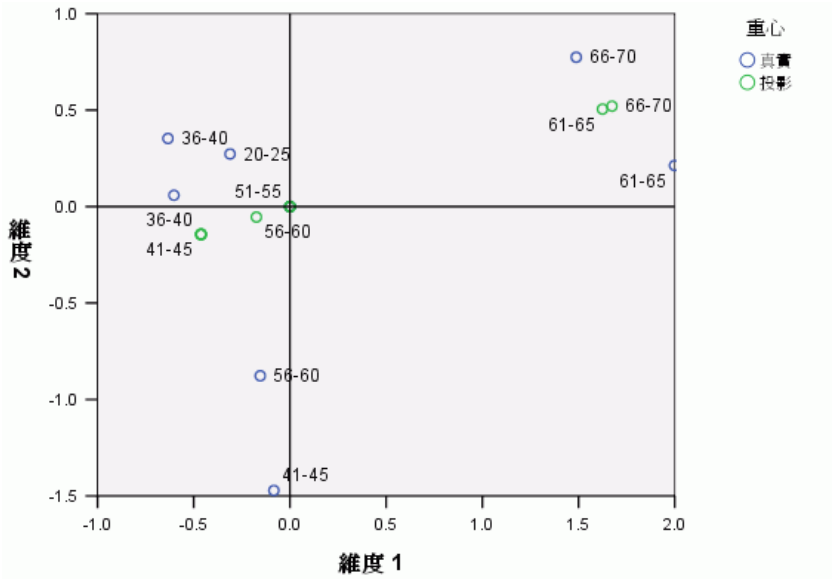

當您要求重心圖時，也會產生由數值標記所標記每個變數的重心和投影重心圖。投影重心在個體空間中並非排列在一直線上。

圖表 11-23
「最常閱讀的報紙」的重心和投影重心



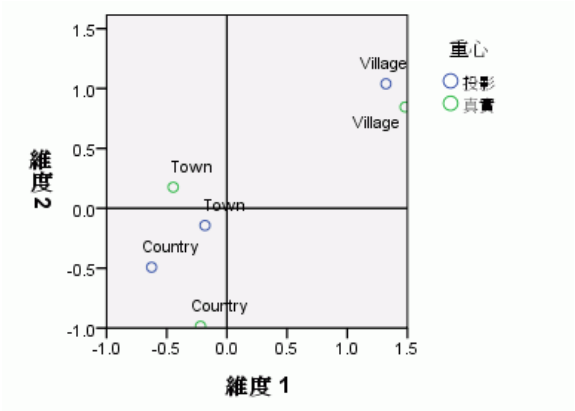

真正的重心會投影在由成份負荷量所定義的向量上。這些向量已經新增到重心圖中，以幫助區分投影重心和實際重心。兩個互相垂直的參考線延伸穿過原點，投影重心會落在這兩條參考線所劃出的其中一個象限中。單一名義量數或是數值變數的方向解釋，可從投影重心的位置得知。例如，最常閱讀的報紙指定為單一名義量數。投影重心會顯示 Volkskrant 和 NRC 與 Telegraaf 成對比。

圖表 11-24
「年齡（年為單位）」的重心和投影重心




從重心投影中，可以很明顯看到年齡（年為單位）的問題。將年齡（年為單位）視為次數量數，意謂著年齡組別的順序必須保留。為了滿足此限制，所有小於 45 歲的年齡組別會投影到同一個點上。沿著由年齡（年為單位）、最常閱讀的報紙和喜愛的居住環境所定義的方向，無法區隔年齡較小的組別。這樣的發現意謂著要將變數視為名義量數。

圖表 11-25
「喜愛的居住環境」的重心和投影重心




若要了解變數之間的關係，請找出特定類別（變數）對於重心圖中的類別集群有什麼樣的意義。看一下圖形的右上角和左下角，可描述出年齡（年為單位）、最常閱讀的報紙和喜愛的居住環境三者間的關係。在右上角的年齡組別是年長的應答者，他們閱覽的報紙是 Telegraaf，喜愛住在郊區。看一下每個圖形的左下角，您會看到年青到中年的應答者讀的是 Volkskrant 或 NRC，想要住在鄉村或城市。但是，很難將年青組別分隔。

現在將注意力放在重心圖的左上角和右下角，可對於另一個方向（喜愛的音樂、婚姻狀況和有養寵物）進行同樣類型的解釋。在左上角，我們發現有單身的人傾向養狗，喜愛新浪潮音樂。已婚的人和婚姻為其他類別的人有養貓，前者喜愛古典音樂，後者不喜愛音樂。

## 其他分析

分析的結果表示將年齡（年為單位）視為次序量數似乎不適當。雖然年齡（年為單位）是在次序水準上測量的，但它與其他變數的關係並不單調。為了調查將最適尺度水準變更為單一名義量數的效果，您可以重新執行該分析。

### 執行分析

- ▶ 叫回「非線性典型相關分析」對話方塊，並瀏覽到第一個資料集。
- ▶ 選擇「age」並按一下「**定義範圍與尺度**」。
- ▶ 在「定義範圍與尺度」對話方塊中，選擇「**單一名義量數**」作為尺度範圍。
- ▶ 按一下「**繼續**」。
- ▶ 在「非線性典型相關分析」對話方塊中，按一下「**確定**」。

雙維度解的特徵值分別是 0.806 和 0.757，總適合度為 1.564。

**圖表 11-26**
**雙維度解的特徵值**

| | | 維度 | | |
|---|---|---|---|---|
| | | 1 | 2 | 總和 |
| 損失 | 集合 1 | .249 | .115 | .363 |
| | 集合 2 | .176 | .408 | .584 |
| | 集合 3 | .157 | .205 | .363 |
| | 平均數 | .194 | .243 | .436 |
| | 特徵值 | .806 | .757 | |
| | 配適 | | | 1.564 |

多重適合度和單一適合度表顯示，根據多重適合度值總和的證明，年齡（年為單位）仍為具高區別度的變數。但先前的結果的解釋卻相反，單一適合值顯示區別度幾乎完全沿著第二個維度。

**圖表 11-27**
**分割適合度和不適度**

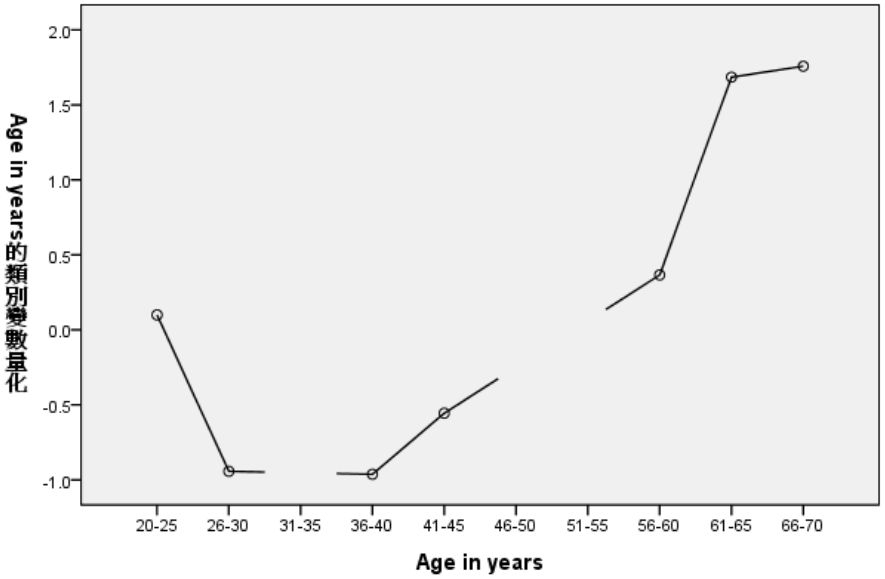
| | | 多重適合度 | | | 單一適合度 | | | 單一損失 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 維度 | | 維度 | 維度 | | 維度 | 維度 | | 維度 |
| 集合 | | 1 | 2 | 總和 | 1 | 2 | 總和 | 1 | 2 | 總和 |
| 1 | Age in years[a] | .246 | 1.197 | 1.443 | .195 | 1.188 | 1.384 | .051 | .008 | .059 |
| | Marital status[a] | .273 | 1.136 | 1.409 | .272 | 1.135 | 1.407 | .001 | .000 | .002 |
| 2 | Pets owned[b] | .530 | .392 | .921 | | | | | | |
| | Newspaper read most often[a] | .639 | .185 | .824 | .631 | .149 | .780 | .008 | .036 | .044 |
| 3 | Music preferred[a] | .604 | .438 | 1.041 | .603 | .437 | 1.040 | .000 | .001 | .001 |
| | Neighborhood preference[a] | .075 | .822 | .897 | .075 | .822 | .897 | .000 | .000 | .000 |

a. 最適尺度水準：單一名義量數

b. 最適尺度水準：複名義量數

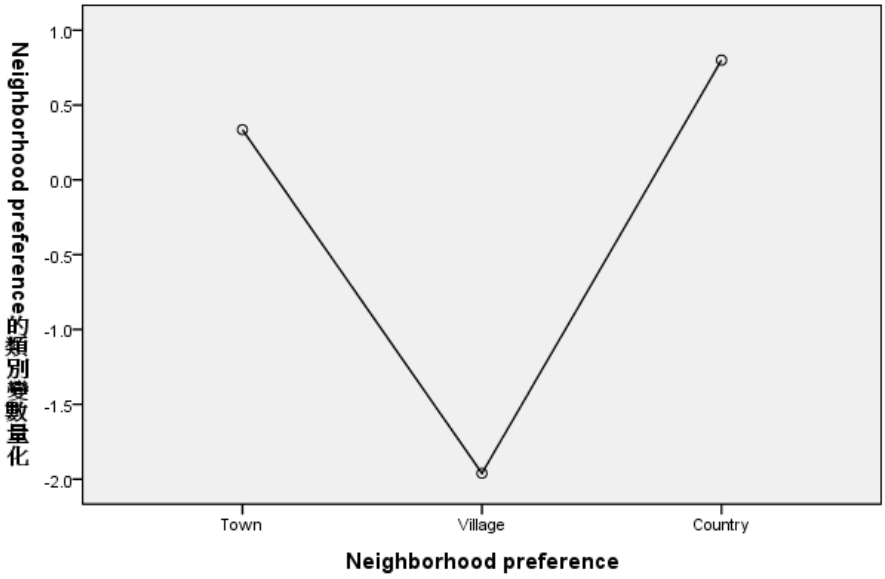
回到年齡（年為單位）的轉換圖。名義變數的量化是沒有限制的，所以將年齡（年為單位）視為次數量數時，便不再發生無減少的趨勢了。減少的趨勢一直到 40 歲為止，之後為增加的趨勢，與 U 形（二次）關係對應。兩個年齡較大的類別仍得到相似的分數，後續的分析可能要將這些類別組合在一起。

**圖表 11-28**
**「年齡（年為單位）」的轉換圖（名義）**



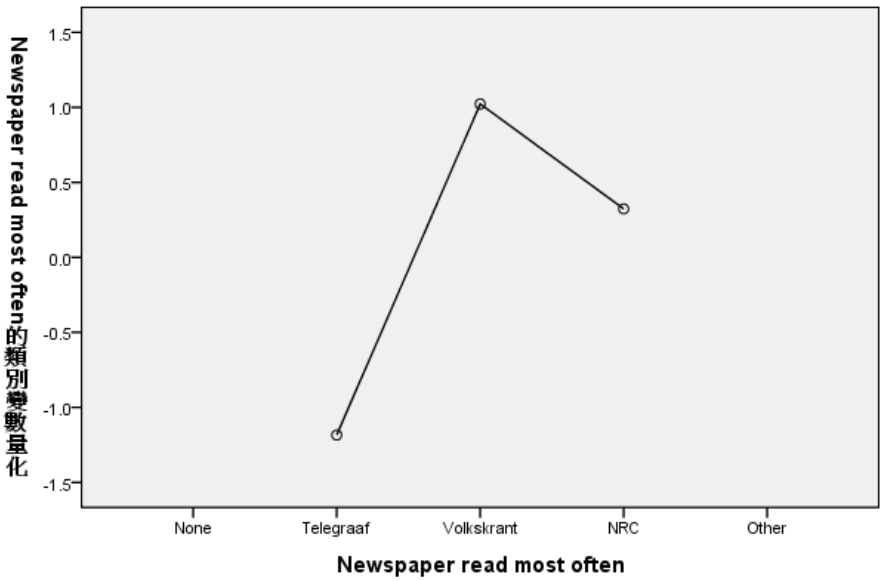
在此會顯示喜愛的居住環境轉換圖。將年齡（年為單位）視為名義量數所影響喜愛的居住環境的量化並不顯著。中間類別的量化值最小，兩端類別則為很大的正值。

圖表 11-29
「喜愛的居住環境（年齡名義）」轉換圖



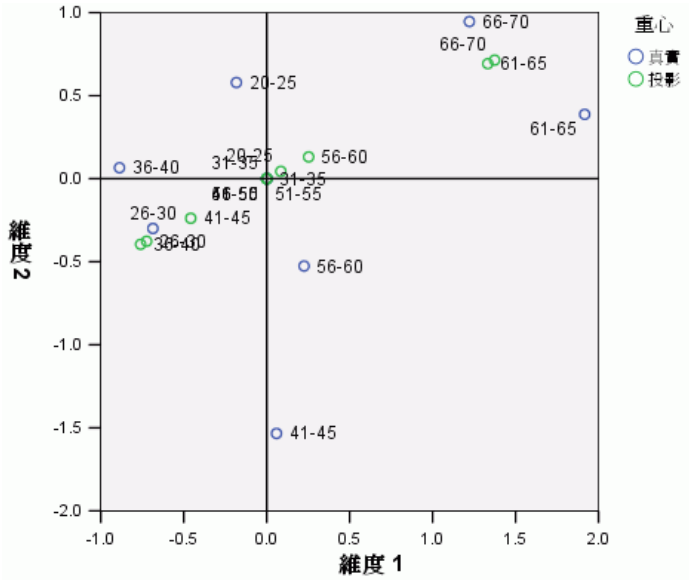

在最愛閱覽的報紙轉換圖中發現一項變更。先前在量化值中呈現增加的趨勢，可能意謂著要將此變數視為次序變數。但是，將年齡（年為單位）視為名義量數，會將此趨勢從新的量化值中移除。

圖表 11-30
「最常閱讀的報紙（年齡名義）」轉換圖



此圖是年齡（單位年）的重心圖。請注意，類別沒有沿著時間順序，落入與投影重心交會的線。20 – 25 組別位在中間，而非兩端。類別量數的散佈程度，比先前在次序量數時的散佈程度還大。

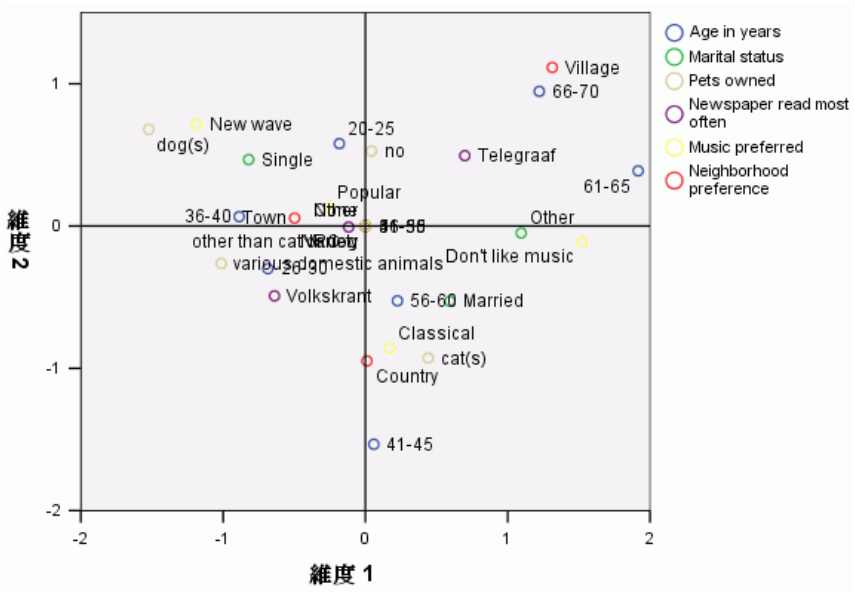
**圖表 11-31**
**「年齡（年為單位）」的重心和投影重心（名義）**




現在可以透過重心圖來解釋年輕組別。Volkskrant 和 NRC 類別與先前的分析相較起來分的更開，因此可分開解釋每個類別。26 和 45 之間的組別閱覽 Volkskrant 報，並喜愛住在鄉下。20 – 25 和 56 – 60 年齡組別閱讀 NRC 報，前者喜愛住在城市，後者喜愛住在鄉下。最年長組別閱讀 Telegraaf，喜愛住在鄉下。

其他方向（喜愛的音樂、婚姻狀況 和 有養寵物）的解釋，基本上與先前的分析一樣。唯一的明顯差異是婚姻狀況為其他的人有養貓，或沒養寵物。

圖表 11-32
由變數標記的重心（年齡名義）




## 一般建議

在您檢驗初始結果後，您可能要變更非線性典型相關分析的一些規格，讓分析更精準。在此提供建構分析的一些祕訣：

- 儘可能建立最多的集合。您要預測的重要變數獨自放在一個集合中。
- 將您認為是預測值的變數一起放在一個集合中。如果有許多預測值，嘗試將它們分隔成數個集合。
- 將每個名義變數獨自放在不同的集合中。
- 如果變數與其他變數高度相關，且您不要此關係主控解，那麼將這些變數一起放在同一個集合中。

# 閱讀資料推薦

請參閱下列文字以取得非線性典型相關分析的詳細資訊：

Carroll, J. D. 1968. Generalization of canonical correlation analysis to three or more sets of variables (三個或以上變數集合之典型相關分析泛論). 於： Proceedings of the 76th Annual Convention of the American Psychological Association, 3 (第 76 屆美國心理學協會大會記錄，第 3 冊), 華盛頓特區: American Psychological Association (美國心理學協會).

De Leeuw, J. 1984. Canonical analysis of categorical data (類別資料之典型分析), 第二版 ed. Leiden (萊頓): DSWO Press (DSWO 出版社).

Horst, P. 1961. Generalized canonical correlations and their applications to experimental data (概化典型相關及其實驗性資料應用). Journal of Clinical Psychology (臨床心理學雜誌), 17, .

Horst, P. 1961. Relations among m sets of measures (m 集量測之間關係 ). Psychometrika (心理學計量報導), 26, .

Kettenring, J. R. 1971. Canonical analysis of several sets of variables (多個變數集合的典型分析). Biometrika (生物統計學雜誌), 58, .

Van der Burg, E. 1988. Nonlinear canonical correlation and some related techniques (非線性典型相關及相關技巧舉例). Leiden (萊頓): DSWO Press (DSWO 出版社).

Van der Burg, E., 和 J. De Leeuw. 1983. Nonlinear canonical correlation (非線性典型相關). British Journal of Mathematical and Statistical Psychology (英國數學暨統計心理學雜誌), 36, .

Van der Burg, E., J. De Leeuw, 和 R. Verdegaal. 1988. Homogeneity analysis with k sets of variables: (K 變數集之均齊性分析)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 53, .

Verboon, P., 和 I. A. Van der Lans. 1994. Robust canonical discriminant analysis (穩健典型判別分析). Psychometrika (心理學計量報導), 59, .

章節
12

# 對應分析

**對應表格**是二因子表格，其中的儲存格包含有部分行和列間的對應測量。對應測量可以是任何列和行變數間的相似性、相關性、混淆性、關聯性或交互作用的標示。最常見的對應表格類型是交叉表列，其中的儲存格含有次數個數。

這樣的表格可在「交叉表」程序中輕易取得。但是，交叉表列不一定能清楚說明兩個變數間的關係特性。如果變數是名義（沒有繼承順序或等級）、且包含各種類別時，此情形尤其明顯。您可以從列出 職業 和 早餐穀物 的10$x$9交叉表列，得知所觀察的儲存格次數與期望值有顯著的差異，但是很難從交叉表列看出哪些職業別的人有相近的口味，以及這些口味是什麼。

「對應分析」可讓您以多維度空間的圖形方法，檢查兩個名義變數之間的關係。他會計算列和行的分數，並依據分數產生圖形。在圖形中，相似的類別會聚集在一起。如此一來，您可以很容易就分辨出相似的變數類別，或是哪兩個變數類別是相關的。「對應分析」程序也可以讓您將補充點嵌入作用點定義的空間。

如果依據分數排列的類別是不需要的或是反直覺的，強制讓部分類別的分數相等會使順序限制更為嚴重。例如，假設您期望變數抽煙行為- 含有類別不抽煙、輕度、中度和重度，擁有對應此順序的分數。但是，如果分析將類別的順序排列成不抽煙、輕度、重度和中度，強制讓重度和中度的分數相等會保留類別依其分數所形成的排列順序。

從距離的觀點來看，對應分析的解讀是依據使用的常態化方法。「對應分析」程序可以用來分析變數各類別之間，或是各變數間的差異。搭配預設的常態化，程序可以分析列和行變數間的差異。

對應分析演算法可以進行多種分析。集中列和行、使用卡方距離相當於標準對應分析。但是，使用結合歐基里得直線距離的另一個集中選項，可允許另一個低維度空間矩陣的表示法。

以下將說明三個範例。第一個範例使用相當小的對應表格，並列舉在對應分析中繼承的概念。第二個範例說明實用的行銷方法。最後一個範例使用距離表格，以多維度尺度法表示。

## 常態化

常態化是用於將慣量分配至列分數和行分數。在各種常態化之下，對應分析方法的某些方面是不會變更的，如奇異值、每個維度的慣量和貢獻。列和行分數、以及其變異會受到影響。對應分析散佈慣量的方式有很多種。三種最常見的方式有：只將慣量散佈至列分數中、只將慣性散佈至行分數中，以及將慣性對稱地散佈至列分數和行分數兩者。

**列主成分分析。** 以列主成份分析常態化而言，列點間的歐基里得直線距離接近對應表格中列間的卡方距離。而且列分數，就是行分數的加權平均數。將行分數標準化，得到至重心 1 的加權距離平方和。因為此方法會最大化列類別間的距離，如果您主要是對列變數間的差異情形感到興趣，您應該使用列主成份分析常態化。

© Copyright SPSS Inc. 1989, 2010

**行主成分分析。** 另一方面，您可能想要估算對應表格中行間的卡方距離。在此情況下，行分數應該是列分數的加權平均數。將列分數標準化，得到至重心 1 的加權距離平方和。此方法會最大化行類別間的距離，如果您主要是對行變數間的差異情形感到興趣，您應該使用列主成份分析常態化。

**對稱的。** 您也可以視列和行為對稱的。此常態化將慣量平均分配至列分數和行分數。注意，在此情況下，列點間的距離或行點間的距離都不是卡方距離的近似值。如果您主要是對兩個變數間的差異或相似性感到興趣，請使用這個方法。這是最常用來製作行列散佈圖的方法。

**主成份分析。** 第四個選項稱為主成份分析常態化，在此方法中，慣量進行兩次散佈—一次是到列分數，一次是到行分數。如果您分別對列點間的距離和行點間的距離感到興趣，但不清楚列點和行點間的相關如何時，您應該使用此方法。行列散佈圖不適用於此常態化選項，因此如果您已指定主成份分析常態化方法，將無法使用行列散佈圖。

# 範例：咖啡品牌的感覺

前一個範例討論到假設資料的小型表格。實際的應用通常會用到較大型的表格。在此範例中，您將使用配合六種冰咖啡品牌感覺印象的資料 (Kennedy, Riquier, 和 Sharp, 1996)。這個資料集存放在 coffee.sav 中。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

對 23 種冰咖啡中每一種的印象屬性，由群眾來選取依其屬性描述的所有品牌。該六種品牌已標示為 AA、BB、CC、DD、EE和 FF，保持機密。

**表格 12-1**
**冰咖啡屬性**

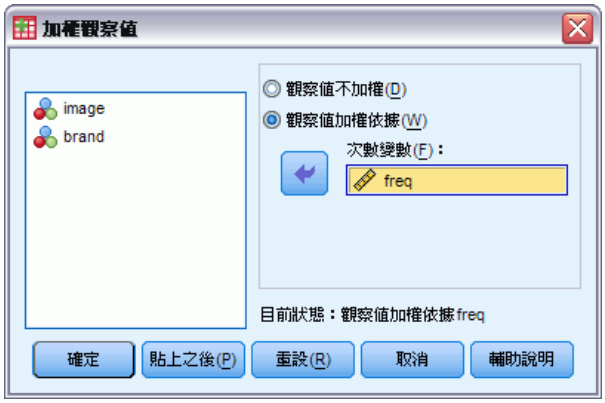
| 印象屬性 | 標記 | 印象屬性 | 標記 |
|---|---|---|---|
| 解宿醉效果佳 | 治療 | 增肥的品牌 | 增肥 |
| 低脂/卡洛里的品牌 | 低脂 | 對男人具有吸引力 | 男人 |
| 適合孩童的品牌 | 孩童 | 南澳的品牌 | 南澳 |
| 勞動階級的品牌 | 勞動 | 傳統/過時的品牌 | 傳統 |
| 濃郁/香甜的品牌 | 香甜 | 品質優良的品牌 | 優良 |
| 不受歡迎的品牌 | 不受歡迎 | 健康的品牌 | 健康 |
| 適合肥胖/其貌不揚者的品牌 | 其貌不揚者 | 高咖啡因的品牌 | 咖啡因 |
| 非常新鮮 | 新鮮 | 新的品牌 | 新增 |
| 適合雅痞的品牌 | 雅痞 | 適和有魅力者的品牌 | 有魅力者 |
| 營養豐富的品牌 | 營養豐富 | 屹立不搖的品牌 | 屹立不搖 |
| 適合女人的品牌 | 女人 | 受歡迎的品牌 | 受歡迎 |
| 非主流的品牌 | 非主流 | | |

一開始，您必須注意每個屬性間的相關性，以及品牌之間的關係。使用主成份分析常態化可將總慣量一次散佈至列和行。儘管如此一來會影響解讀行列散佈圖，但是可以檢驗每個變數的類別間的距離。

## 執行分析

- 資料設定需要以變數 freq 來加權觀察值。若要進行此功能，在功能表中選擇：
  資料 > 加權觀察值(W)...

圖表 12-1
「加權觀察值」對話方塊

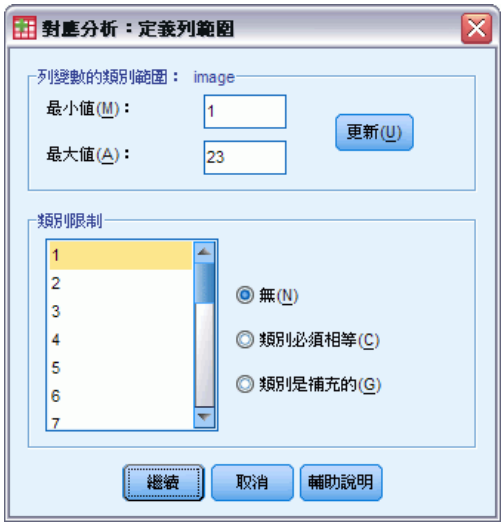



- 以 freq 加權觀察值。
- 按一下「確定」。
- 若要使用主成份分析常態化，取得五個維度的未轉軸之統計量，在功能表中選擇：
  分析(A) > 維度縮減(D) > 對應分析...

圖表 12-2
「對應分析」對話方塊。

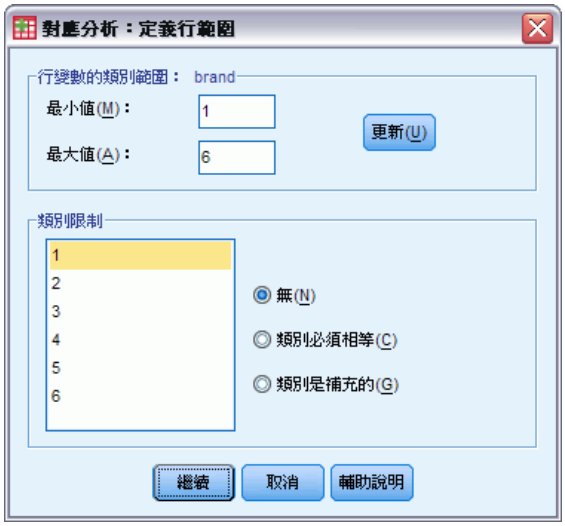
- 選取「印象」作為列變數。
- 按一下「定義範圍」。

圖表 12-3
「定義列範圍」對話方塊




- 輸入 1 作為最小值。
- 輸入 23 作為最大值。
- 按一下「更新」。
- 按一下「繼續」。
- 選取「品牌」作為行變數。
- 按一下「對應分析」對話方塊中的「定義範圍」。

圖表 12-4
「定義行範圍」對話方塊



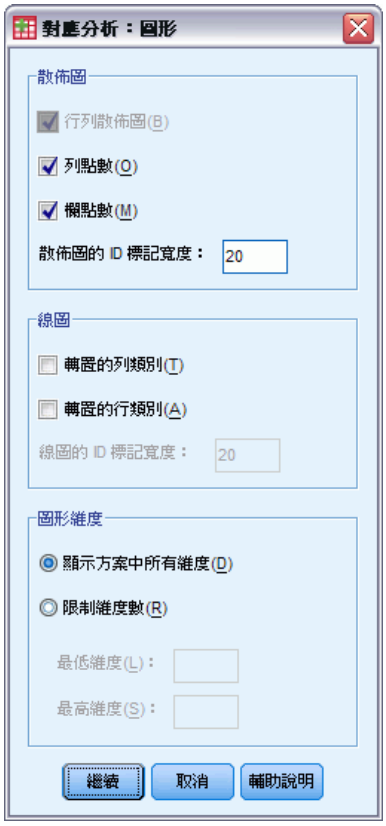

- 輸入 1 作為最小值。
- 輸入 6 作為最大值。

- 按一下「**更新**」。
- 按一下「**繼續**」。
- 在「對應分析」對話方塊中，按一下「**模式**」。

圖表 12-5
「模式」對話方塊

- 選取「**主成份分析**」作為常態化方法。
- 按一下「**繼續**」。
- 「對應分析」對話方塊中，按一下「**圖形**」。

圖表 12-6
「圖形」對話方塊




- 在「散佈圖」組別中，選取「列點」和「行點」。
- 按一下「繼續」。
- 按一下「對應分析」對話方塊中的「確定」。

## 維度

每個維度的慣量顯示出每個維度總慣量的分解。兩個維度佔總慣量 83%。加入第三個維度僅對所佔慣量再加 8.6%。因此，您可以選擇使用二維表示法。

圖表 12-7
每個維度的慣量

| 維度 | 奇異值 | 慣性 | 卡方 | 顯著性 | 概化變異數的比例 | | 信賴奇異值 | |
|---|---|---|---|---|---|---|---|---|
| | | | | | 計算 | 累積 | 標準差 | 相關 2 |
| 1 | .711 | .506 | | | .629 | .629 | .009 | .132 |
| 2 | .399 | .159 | | | .198 | .827 | .014 | |
| 3 | .263 | .069 | | | .086 | .913 | | |
| 4 | .234 | .055 | | | .068 | .982 | | |
| 5 | .121 | .015 | | | .018 | 1.000 | | |
| 總數 | | .804 | 3746.968 | .000[a] | 1.000 | 1.000 | | |

a. 110 自由度

## 貢獻

列點的概覽顯示列點對維度的貢獻，以及維度對列點慣量的貢獻。如果所有點對慣量的貢獻相等，則貢獻應是 0.043。健康和低脂兩者皆對第一個維度慣量有實質的貢獻。男人和屹立不搖對第二維度的慣量貢獻最大。其貌不揚者和新鮮則對兩個維度貢獻極少。

圖表 12-8
屬性貢獻

| image | 量 | 維度的分數 | | 慣性 | 貢獻 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 點對維度的概化變異數 | | 維度對點的概化變異數 | | |
| | | 1 | 2 | | 1 | 2 | 1 | 2 | 總數 |
| fattening | .080 | -.514 | -.265 | .033 | .042 | .035 | .652 | .173 | .825 |
| men | .051 | -.852 | .825 | .072 | .073 | .219 | .512 | .480 | .992 |
| South Australian | .057 | -.303 | -.350 | .046 | .010 | .044 | .114 | .152 | .266 |
| traditional | .040 | -.703 | -.532 | .043 | .039 | .071 | .454 | .260 | .715 |
| premium | .042 | -.444 | -.582 | .028 | .016 | .090 | .296 | .509 | .805 |
| healthy | .053 | 1.200 | .174 | .081 | .152 | .010 | .953 | .020 | .973 |
| caffeine | .047 | -.452 | .124 | .014 | .019 | .005 | .702 | .053 | .755 |
| new | .047 | .960 | .147 | .048 | .086 | .006 | .893 | .021 | .914 |
| attractive | .041 | .657 | -.056 | .019 | .035 | .001 | .911 | .007 | .918 |
| tough | .039 | -.850 | 1.002 | .070 | .056 | .246 | .404 | .560 | .964 |
| popular | .060 | -.697 | -.042 | .038 | .058 | .001 | .771 | .003 | .774 |
| cure | .026 | -.389 | .266 | .009 | .008 | .011 | .446 | .209 | .655 |
| low fat | .052 | 1.305 | .196 | .094 | .175 | .013 | .941 | .021 | .962 |
| children | .024 | -.352 | -.513 | .017 | .006 | .041 | .179 | .380 | .559 |
| working | .045 | -.785 | .477 | .040 | .055 | .064 | .693 | .255 | .948 |
| sweet | .038 | -.519 | -.683 | .048 | .020 | .112 | .212 | .368 | .580 |
| unpopular | .024 | .489 | .186 | .010 | .011 | .005 | .585 | .085 | .670 |
| ugly | .030 | .006 | -.109 | .003 | .000 | .002 | .000 | .131 | .131 |
| fresh | .036 | -.096 | -.100 | .002 | .001 | .002 | .196 | .214 | .410 |
| yuppies | .034 | .380 | -.301 | .012 | .010 | .019 | .392 | .246 | .637 |
| nutritious | .040 | .722 | .055 | .022 | .041 | .001 | .946 | .006 | .951 |
| women | .054 | .758 | -.063 | .032 | .062 | .001 | .965 | .007 | .972 |
| minor | .040 | .579 | .063 | .023 | .027 | .001 | .593 | .007 | .600 |
| 使用中總和 | 1.000 | | | .804 | 1.000 | 1.000 | | | |

兩個維度對大部分列點的慣量貢獻很大。第一維度對健康、新的、有魅力者、低脂、營養豐富和女人的大量貢獻表示：這些點在一個維度中顯示良好。所以，較高的維度對這些點的慣量貢獻較少，位置將會非常接近水平軸。第二維度對男人、優良和屹立不搖貢獻大部分。兩個維度對南澳和其貌不揚者的慣量貢獻極少。

行點的概覽顯示貢獻包含行點。品牌 CC 和 DD 大部分貢獻給第一個維度，相反地，EE 和 FF 說明了第二維度的慣量很大。AA 和 BB 則對兩個維度貢獻極少。

**圖表 12-9**
**品牌貢獻**

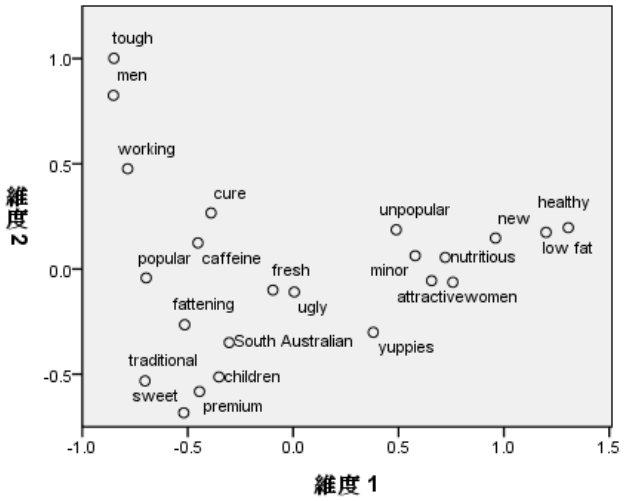
| | | 維度的分數 | | | 貢獻 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 點對維度的概化變異數 | | 維度對點的概化變異數 | | |
| brand | 量 | 1 | 2 | 慣性 | 1 | 2 | 1 | 2 | 總數 |
| AA | .217 | -.659 | .046 | .127 | .187 | .003 | .744 | .004 | .748 |
| BB | .131 | -.284 | -.404 | .078 | .021 | .134 | .135 | .272 | .407 |
| CC | .185 | .996 | .076 | .193 | .362 | .007 | .951 | .006 | .957 |
| DD | .162 | .915 | .101 | .146 | .267 | .010 | .928 | .011 | .939 |
| EE | .152 | -.651 | .706 | .153 | .127 | .477 | .420 | .494 | .914 |
| FF | .153 | -.343 | -.618 | .107 | .036 | .369 | .169 | .550 | .718 |
| 使用中總和 | 1.000 | | | .804 | 1.000 | 1.000 | | | |

在兩個維度中，除 BB 外，所有的品牌都顯示良好。CC 和 DD 在一個維度中顯示良好。第二維度對 EE 和 FF 貢獻最大。注意，AA 在第一個維度中顯示良好，但是對該維度卻沒有很高的貢獻。

## 圖形

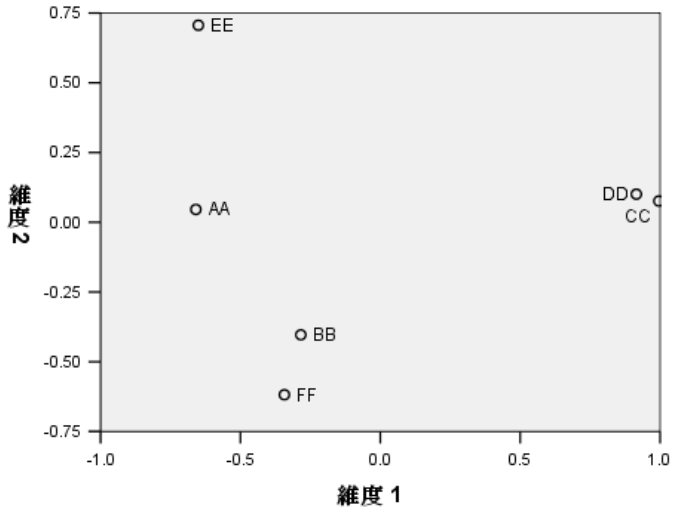
列分數圖形顯示新鮮和其貌不揚者兩者都非常接近原點，表示他們與平均列剖面圖的差異不大。會出現三種一般分類。在圖形左上角，屹立不搖、男人和勞動是彼此相似的。左下角有香甜、增肥、孩童和優良。相反地，健康、低脂、營養豐富和新的集群位於圖形右側。

圖表 12-10
印象屬性的圖形（主成份分析常態化）




注意，在行點圖形中，所有品牌都遠離原點，因此沒有品牌近似整體的重心。品牌 CC 和 DD 群聚在右側，相反地，品牌 BB 和 FF 集群在圖形的下半部。品牌 AA 和 EE 與其他任何品牌不相似。

圖表 12-11
品牌的圖形（主成份分析常態化）




## 對稱常態化

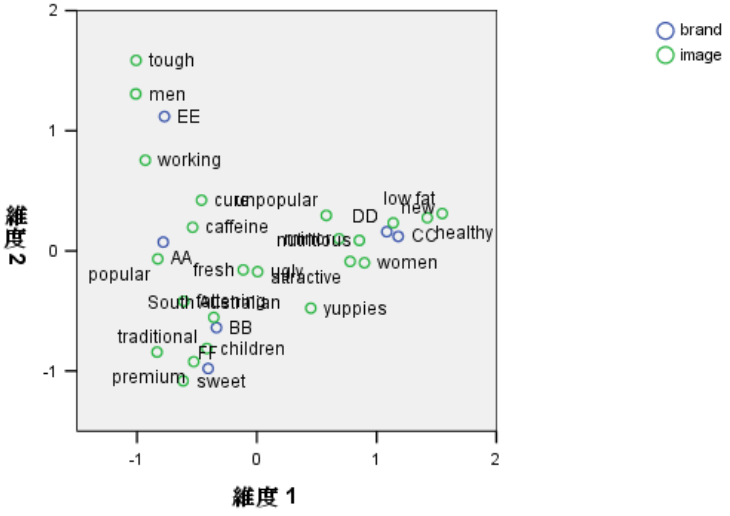
品牌與印象屬性的關係如何？主成份分析常態化無法提出它們之間的關係。若要瞭解變數之間的關係，請使用對稱常態化。不同於散佈慣量兩次（如同在主成份分析常態化之中），對稱常態化將慣量平均分配給列和行。單一變數的類別間距離無法被解讀，但是多個變數的類別間距離是有意義的。

圖表 12-12
「模式」對話方塊

- 若要使用對稱常態化產生下列方法，請叫回「對應分析」對話方塊，然後按一下「模式」。
- 選擇「對稱」作為常態化方法。
- 按一下「繼續」。
- 按一下「對應分析」對話方塊中的「確定」。

在結果行列散佈圖的左上角，品牌 EE 是屹立不搖的、是勞動階級的品牌，並且對男人具有吸引力。品牌 AA 是最受歡迎的，也是被視為含有最高咖啡因的。香甜、增肥的品牌有 BB 和 FF。當印象是新的、健康的，品牌 CC 和 DD 是最不受歡迎的。

**圖表 12-13**
**品牌和屬性的行列散佈圖（對稱常態化）**




若要進一步解讀，您可以在原點和兩個印象屬性男人和雅痞間畫一條直線，然後將品牌投射在此線上。那兩個屬性位在相對的位置，表示男人的品牌關聯形式與雅痞的形式相較起來是相反的。也就是男人最常與品牌 EE 相關，與品牌 CC 相關性最低；相反地，雅痞最常與品牌 CC 相關，與品牌 EE 相關性最低。

# 閱讀資料推薦

請參閱下列文字以取得對應分析的詳細資訊：

Fisher, R. A. 1938. Statistical methods for research workers (研究人員的統計方法). Edinburgh (愛丁堡): Oliver and Boyd.

Fisher, R. A. 1940. The precision of discriminant functions (判別函數之精確度). Annals of Eugenics (優生學年鑑), 10, .

Gilula, Z., 和 S. J. Haberman. 1988. The analysis of multivariate contingency tables by restricted canonical and restricted association models (有限典型與有限關聯模式之多變量列聯表分析). Journal of the American Statistical Association (美國統計協會彙報), 83, .

章節
13

# 多重對應分析

多重對應分析（也稱為均齊性分析）的目的，是用於尋找類別之間為盡可能互相分隔的最佳量化。這表示會將同類別的物件繪製為互相靠近，而會將不同類別的物件繪製為盡可能地互相分離。**均齊性**這個名詞也是表示當變數為同質時，也就是當物件分為含有相同或相似類別的集群時，分析會非常順利。

## 範例：硬體的特性

探索多重對應分析如何運作將使用 Hartigan (Hartigan, 1975)的資料，可以在 screws.sav 中找到資料。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案. 這個資料集包含螺絲釘、螺栓、螺帽和圖釘之特色的資訊。以下表格會顯示在 Hartigan 硬體資料集中的變數，以及其變數標記，和指定給各變數類別的數值標記。

**表格 13-1**
**Hartigan 硬體資料集**

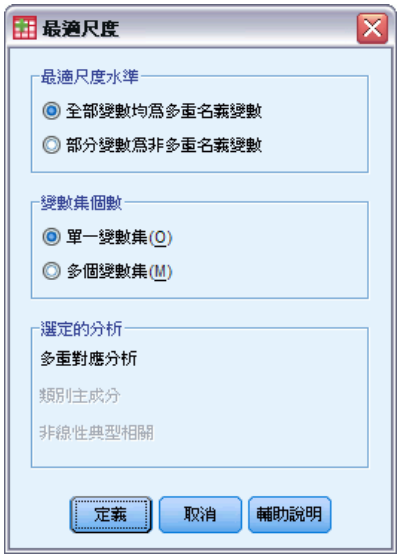
| 變數名稱 | 變數標記(L) | 數值標記 |
|---|---|---|
| 執行緒 | 執行緒 | Yes_Thread、No_Thread |
| 頂部 | 頂部形式 | 平的、杯狀、圓錐形、圓的、圓柱形 |
| 頂部刻痕 | 頂部刻痕 | 無、星狀、裂縫 |
| 底部 | 底部形狀 | 尖的、平的 |
| 長度 | 半英吋的長度 | 1/2_in、1_in、1_1/2_ in、2_in、2_1/2_in |
| 黃銅 | 黃銅 | Yes_Br、Not_Br |
| 物件 | 物件 | tack、nail1、nail2、nail3、nail4、nail5、nail6、nail7、nail8、screw1、screw2、screw3、screw4、screw5、bolt1、bolt2、bolt3、bolt4、bolt5、bolt6、tack1、tack2、nailb、screwb |

## 執行分析

- 若要取得「多重對應分析」，請從功能表選擇：
  **分析(A) > 維度縮減(D) > 最適尺度...**

© Copyright SPSS Inc. 1989, 2010

圖表 13-1
「最適尺度」對話方塊




▶ 請確定已經選擇「全部變數均為複名義量數」以及「單一變數集」，並且按一下「定義」。

圖表 13-2
「多重對應分析」對話方塊

▶ 選擇「執行緒」到「半英呎的長度」，做為分析變數。

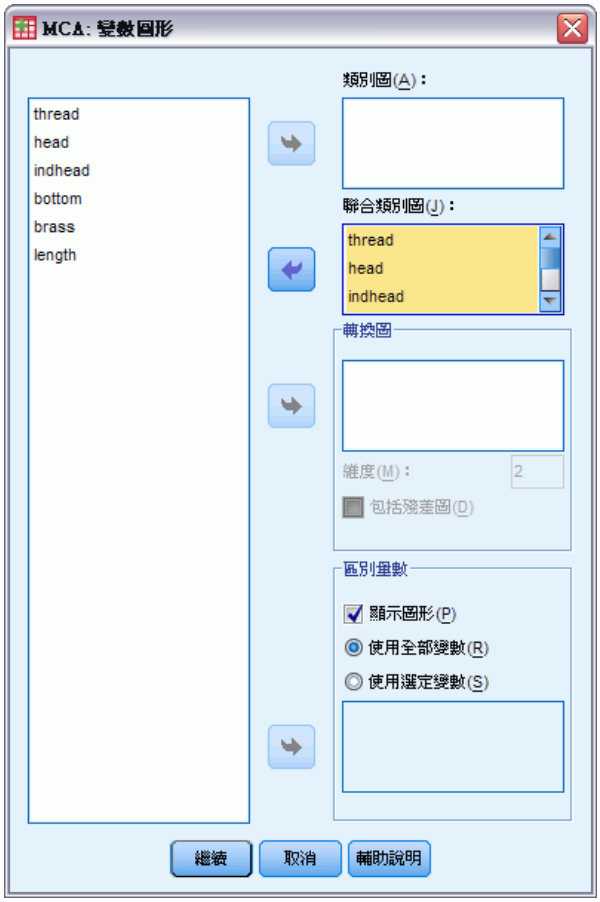
▶ 選擇「物件」做為標記變數。

▶ 在「圖形」組別中按一下「物件」。

圖表 13-3
「個體圖」對話方塊

- 選擇按照變數來標記個數。
- 選擇「執行緒」到「物件」做為標記變數。
- 按一下「繼續」，並且按一下「多重對應分析」對話方塊中「圖形」組別的「變數」。

圖表 13-4
「變數圖」對話方塊




- ▶ 選擇以產生「執行緒」到「長度」的聯合類別圖。
- ▶ 按一下「繼續」。
- ▶ 在「多重對應分析」對話方塊中按一下「確定」。

## 模式摘要

均齊性分析可計算出數個維度的解答。該程序所使用的維度最大個數，會等於類別的總個數減去不具遺漏值之變數個數，或者是等於觀察值個數減 1，請從這兩值中取較小的值。但是，您應該少用維度的最大個數。較小的維度個數較易解讀，而且在維度的某個個數之後，其他關連的數量會變成微不足道。一個、二個或三個維度之均齊性分析的解答是很常見的。

圖表 13-5
模式摘要(M)

| 維度 | Cronbach's Alpha 值 | 變異數歸因於 | | |
|---|---|---|---|---|
| | | 總和 (特徵值) | 慣性 | 變異數的 % |
| 1 | .878 | 3.727 | .621 | 62.123 |
| 2 | .657 | 2.209 | .368 | 36.809 |
| 總數 | | 5.936 | .989 | |
| 平均數 | .796[a] | 2.968 | .495 | 49.466 |

a. Cronbach's Alpha 平均值是以平均特徵值爲準。

幾乎所有資料中的變異數都可由解答來說明，62.1% 由第一維度，而 36.8% 由第二維度。

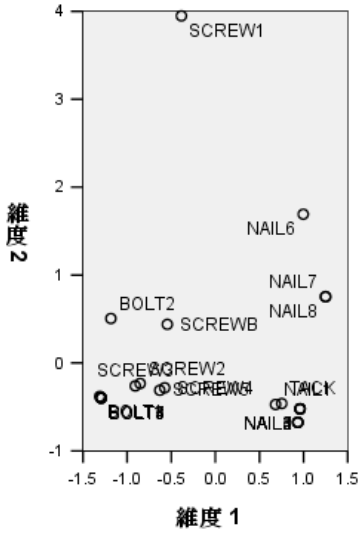
兩個維度會一起提供距離的解讀。如果變數能妥善區分，物件就會和所屬的類別很接近。理想的情形是相同類別的物件會互相接近（也就是說，它們應該都有相似的分數），而且當不同變數屬於相同物件時，不同變數的類別也會互相接近（也就是說，如果兩個物件對某一個變數而言具有相似的分數，則對解答中其他的變數而言，這兩個物件的分數也應該是彼此接近的）。

## 物件分數

在檢查模式摘要之後，您應該查看物件分數。您可以指定一個或多個變數以標記物件分數圖。各標記變數都會產生以該變數值標示之個別的圖。我們’將會查看以變數物件標示的物件分數圖。這只是觀察值識別變數，不會用來進行任何計算。

從物件到原點的距離會反映「平均」反應樣式的變化。這個平均反應樣式會對應到各變數次數最多的類別。有很多特色的物件會對應到接近原點次數最多的類別。相反的，有獨特特色的物件會離原點較遠。

圖表 13-6
以物件標示的物件分數圖



檢查圖形時，您會看到第一維度（水平軸）將螺絲釘和螺栓（含有執行緒）以及釘子和圖釘（不’含執行緒）區分開來。您可以很輕易在圖中看到這個情形，螺絲釘和螺栓在水平軸的其中一端，而圖釘和釘子則在另一端。若是較小的範圍，第一維度也會將螺栓（含有平的底部）與所有其他物件分隔開來。

第二維度（垂直軸）看來會將 SCREW1 和 NAIL6 與其他物件分隔開來。SCREW1 和 NAIL6 的共同點是其變數的長度值 — 它們都是資料中最長的物件。此外，SCREW1 位於比其他物件離原點更遠的地方，表示整體來說，這個物件沒有和其他物件共用很多特色。

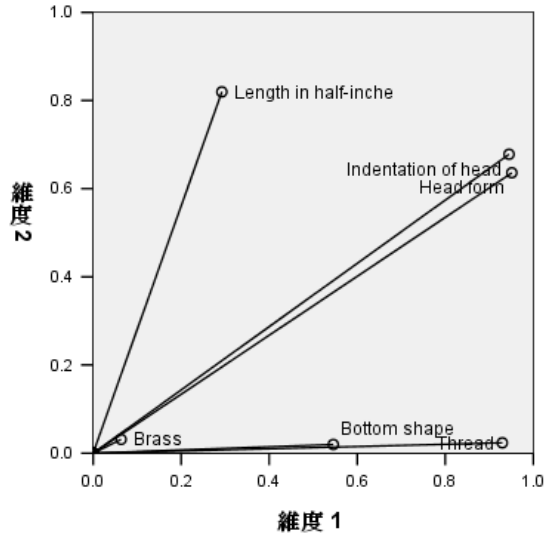
物件分數圖對於繪製偏離值特別實用。可能會將 SCREW1 視為偏離值。如果您捨棄這個物件，稍後我們’將會思考所發生的情況。

## 區別量數

在檢查剩下的物件分數圖之前，讓我們先察看區別量數是否符合到目前為止所做的說明。對於各變數而言，會為各維度計算區別量數（可以看作「平方成分負荷量」）。這個量數也是維度中量化變數的變異數。如果物件分數落入互相排斥的組別中，而且類別中所有的物件分數都相同時，此量數為最大值 1。(注意：如果有遺漏值，量數可能會是大於 1 的數值。）大區別量數會對應到變數類別中的大分佈，因此能指出該維度內的變數類別間有高度的區別。

任何維度的平均區別量數會等於說明該維度的變異數百分比。因此，會依照平均區別來排列維度。對於解答中之所有維度而言，第一維度含有最大的平均區別，第二維度含有第二大的平均區別，依此類推。

圖表 13-7
區別量數圖




如物件分數圖中所記錄，區別量數圖顯示第一維度與變數的執行緒和底部形狀相關連。這些變數在第一維度中含有大的區別量數，而在第二維度中含有小的區別量數。如此一來，對於這兩個變數，類別只會在第一維度內大範圍地散開。 半英吋的長度在第二維度中含有大數值，但是在第一維度中則含有小數值。如此一來，長度會最接近第二維度，也會符合物件分數圖的觀察值，此物件分數圖上的第二維度看來會將最長物件與其餘物件分隔開來。頂部刻痕和頂部形式在兩個維度中都含有相對上比較大的數值，表示同時

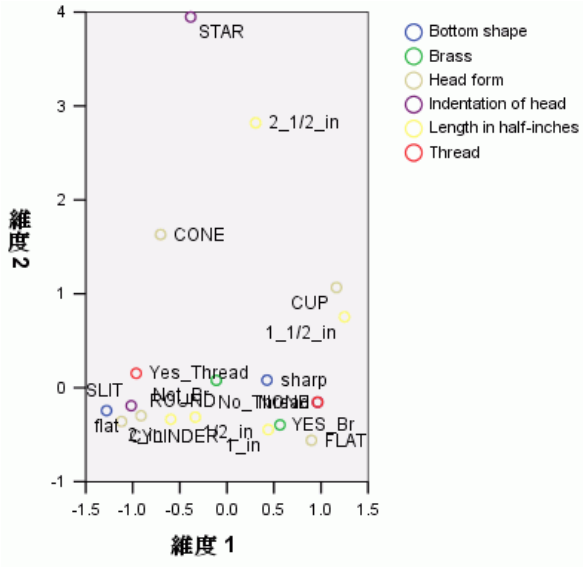
在第一與第二維度中的區別。與原點非常接近的黃銅變數在前兩個維度中完全沒有區別。這是合理的，因為所有的物件可以使用黃銅製作也可以不使用黃銅製作。

## 類別量化

請記住，區別量數就是特定維度內之量化變數的變異數。區別量數圖會包含這些變異數，並指出會在哪一個維度內區別哪些變數。但是，相同的變異數可以對應到所有適度散開的類別，或是對應到互相接近且有一些類別不同於這個組別的大部分類別。區別圖無法區分這兩種情況。

類別量化圖能提供另一個顯示變數區別的方法，可識別類別間的關係。在這個圖中，會顯示各維度中之各類別的座標。如此一來，您可以區別各變數的哪些類別是相似的。

圖表 13-8
類別變數的量化 (Q)




半英呎的長度含有五個類別，其中三類聚集在接近圖頂的地方。剩下的兩類位於圖的下半部，且 2_1/2_in 類別則位於離組別很遠的位置。維度 2 內的長度有很大的區別是因為這個類別和其他長度類別非常不同。同樣地，針對頂部形式而言，「星狀」類別距離其他類別很遠，且在維度 2 內有大區別量數。無法在區別量數圖中說明這些樣式。

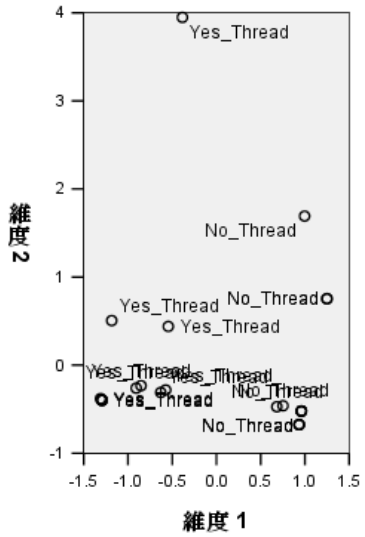
變數之類別量化的分佈會反映變異數，因此會指出變數在各維度中的區分效果。先觀察維度 1，會發現執行緒的類別互相距離很遠。但再觀察維度 2 時，便會發現這個變數的類別互相非常接近。因此，在維度 1 內比在維度 2 內更能妥善區分執行緒，相反的，兩個維度內的頂部形式皆會大範圍地散開，表示在兩個維度內都能妥善區分這個變數。

除了決定要在哪個維度內區分變數，以及此如何區分此變數外，類別量化圖也會比較變數區別。含有互相距離很遠之類別的變數，比含有互相接近之類別的變數，能進行較好的區分。 例如，維度 1 內黃銅的兩個類別比執行緒的兩個類別還要更為互相接近，這表示此維度內執行緒比黃銅能進行較好的區分。但是，維度 2 內的距離很相似，表示此維度內的變數會區別為相同的程度。透過使用變異數來反映類別的分佈情形，先前所討論的區別量數圖可識別這些相同的關係。

# 詳細查看物件分數

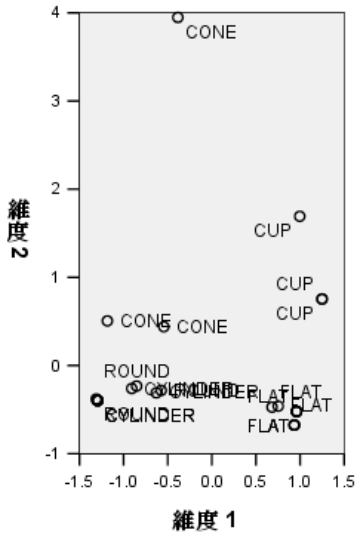
透過檢查以各變數標示的物件分數圖，可對資料有更深入的理解。理想的情況是，相似物件應該形成互相排斥的組別，且這些組別互相距離很遠。

圖表 13-9
標示為執行緒的物件分數



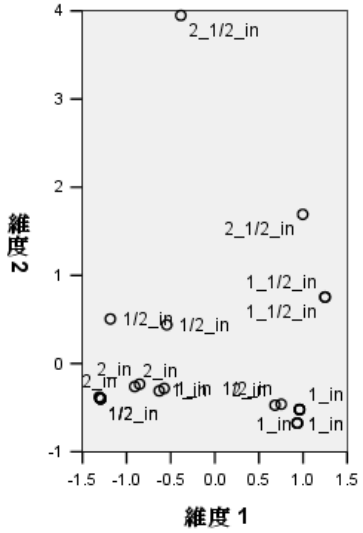

標示為執行緒的圖會顯示第一維度將 Yes_Thread 和 No_Thread 完美地區隔開來。所有具執行緒的物件分數皆為負數，而所有不具執行緒的物件分數皆為正數。雖然這兩個類別沒有形成緊密的組別，但類別間的完美區分整體來說可視為好的結果。

圖表 13-10
標示為頂部形式的物件分數



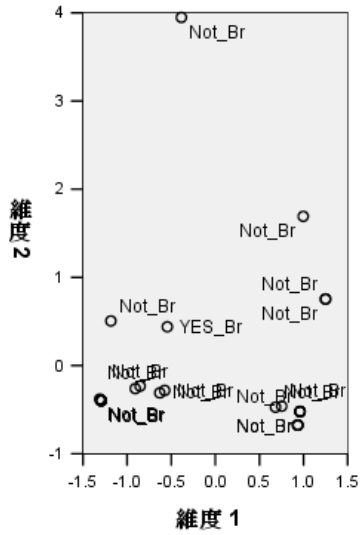

標示為頂部形式的圖顯示這個變數同時在兩個維度中進行區分。「平的」物件聚集在圖的右下方，而「杯狀」物件則聚集在右上方。「圓錐形」物件全部位於左上方。但是，這些物件比其他組別散佈的範圍更廣，因此同質性較低。最後，無法分隔「圓柱形」物件和「圓的」物件，兩者都位於圖的左下角。

圖表 13-11
標示為半英呎長度的物件分數




標示為半英呎長度的圖顯示這個變數無法在第一維度內進行區分。當投影在水平線上時，其類別顯示為沒有分組。但是，半英吋的長度會在第二維度內進行區別。較短的物件會對應到正分數，而較長的物件會對應到大的負分數。

圖表 13-12
標示為黃銅的物件分數



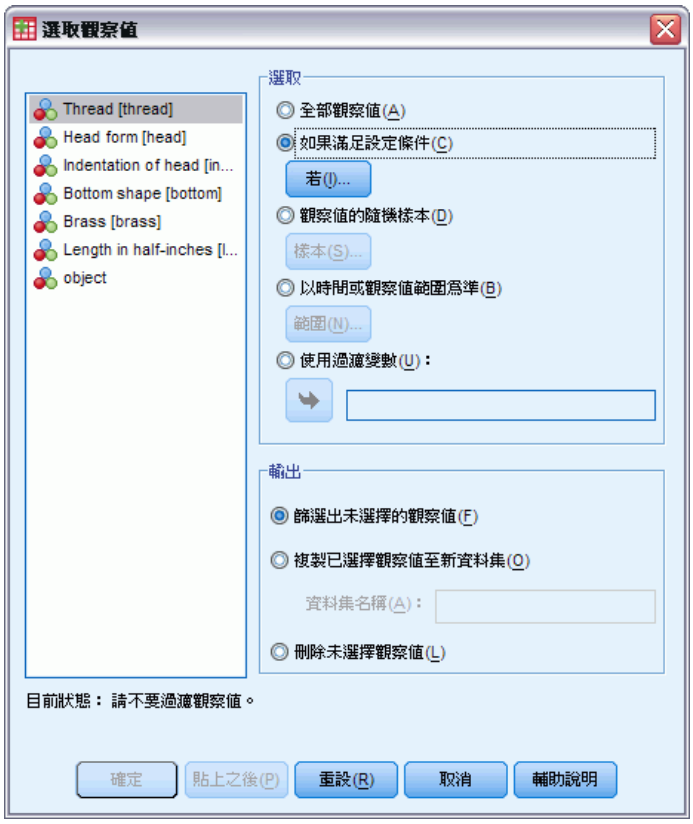

標示為黃銅的圖顯示無法在第一或第二維度內非常妥善的分隔此變數的類別。物件分數會在整個空間中廣泛散佈開來。無法區分黃銅物件和非黃銅物件。

## 遺漏偏離值

在均齊性分析中，偏離值為含有太多獨特功能的物件。如同之前所記錄，可能會將 SCREW1 視為偏離值。

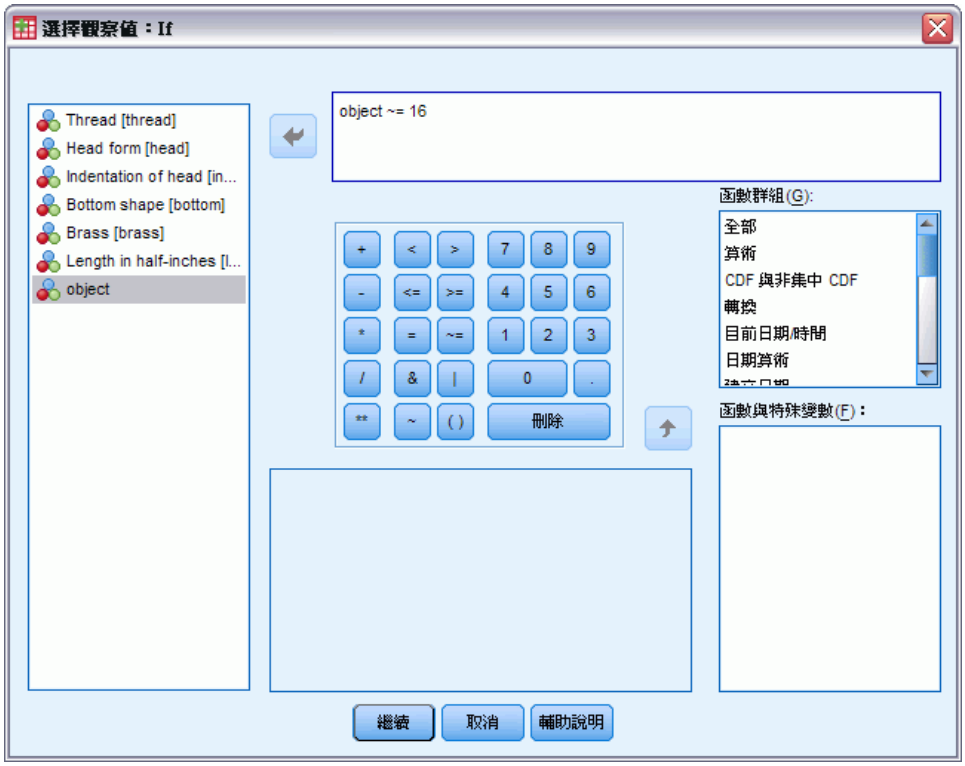
若要刪除此物件並再次執行分析，請從功能表中選擇：
**資料 > 選擇觀察值(S)...**

圖表 13-13
「選擇觀察值」對話方塊




- 選擇「如果滿足設定條件」。
- 按一下「如果」。

圖表 13-14
「如果」對話方塊




- 輸入 object ~= 16 做為條件。
- 按一下「繼續」。
- 按一下「選擇觀察值」對話方塊中的「確定」。
- 最後，叫回「多重對應分析」對話方塊，並按一下「確定」。

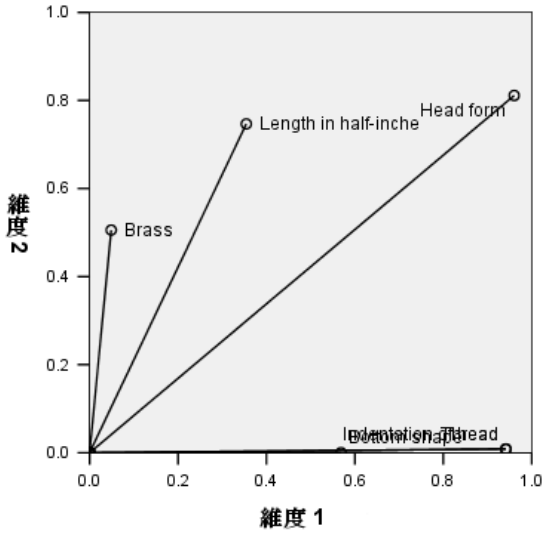
圖表 13-15
模式摘要（已移除偏離值）

| 維度 | Cronbach's Alpha 值 | 變異數歸因於 | | |
|---|---|---|---|---|
| | | 總和 (特徵值) | 慣性 | 變異數的 % |
| 1 | .885 | 3.815 | .636 | 63.591 |
| 2 | .623 | 2.081 | .347 | 34.676 |
| 總數 | | 5.896 | .983 | |
| 平均數 | .793[a] | 2.948 | .491 | 49.133 |

a. Cronbach's Alpha 平均值是以平均特徵值爲準。

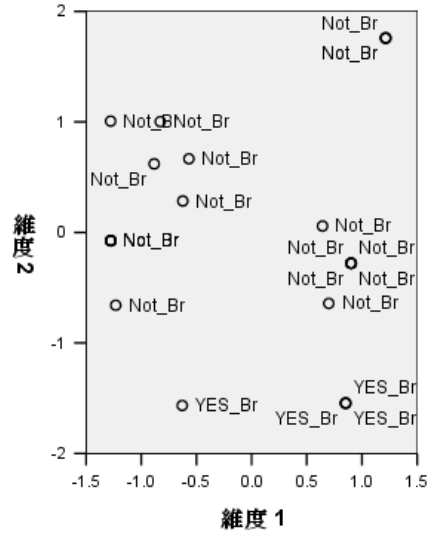
特徵值會稍微移位。現在第一維度會說明稍微多一些的變異數。

圖表 13-16
鑑別量測(S)



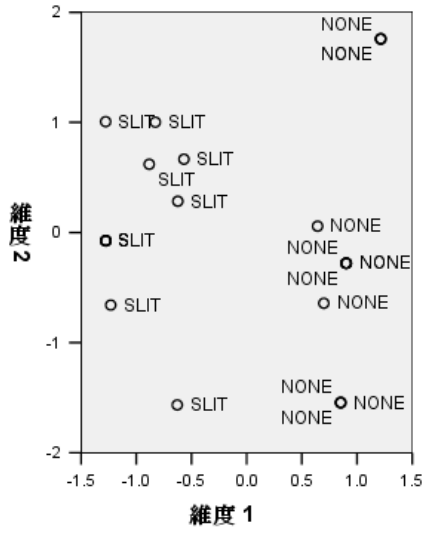

如區別圖所示，已不再區別第二維度內的頂部刻痕，而將原本在兩維度內皆不區別的黃銅，變更為在第二維度內區別。其他變數的區別則大部分皆保持原狀。

圖表 13-17
標示為黃銅的物件分數（已移除偏離值）




標示為黃銅的物件分數圖顯示四個黃銅物件都出現在接近圖的底部（三個物件位於相同位置），表示第二維度內有高度的區別。做為前一個分析中執行緒觀察值的物件雖然沒有形成緊密的組別，但已完美地依照類別來區分物件。

圖表 13-18
標示為頂部刻痕的物件分數（已移除偏離值）




如同前一個分析，標示為頂部刻痕的物件分數圖顯示可在第一維度內完美地區分無刻痕物件和有刻痕物件。但是，與前一個分析的相反情況是，現在第二維度無法區分兩個類別。

如此一來，遺漏唯一頂部為星狀的物件 SCREW1 會大幅地影響第二維度的解讀。現在維度會根據黃銅、頂部形式，和半英呎長來區分物件。

# 閱讀資料推薦

請參閱下文以取得多重對應分析的詳細資訊：

Benzécri, J. P. 1992. Correspondence analysis handbook (對應分析手冊). 紐約: Marcel Dekker.

Guttman 值, L. 1941. The quantification of a class of attributes: (屬性等級量化)A theory and method of scale construction (尺度建構之理論與方法). 於： The Prediction of Personal Adjustment (個人調整預測), P. Horst，ed. 紐約: Social Science Research Council (社會學研究協會).

Meulman, J. J. 1982. Homogeneity analysis of incomplete data (不完整資料之均齊性分析). Leiden (萊頓): DSWO Press (DSWO 出版社).

Meulman, J. J. 1996. Fitting a distance model to homogeneous subsets of variables: (在同質性變數子集上配適距離模式)Points of view analysis of categorical data (類別資料之觀點分析). Journal of Classification (分類雜誌), 13, .

Meulman, J. J., 和 W. J. Heiser. 1997. Graphical display of interaction in multiway contingency tables by use of homogeneity analysis (使用同質性分析之多因子列聯表顯示互動圖形). 於： Visual Display of Categorical Data (類別資料視覺顯示), M. Greenacre, 和 J. Blasius，eds. 紐約: Academic Press (學術出版社).

Nishisato, S. 1984. Forced classification: (強制分類)A simple application of a quantification method (量化法簡易應用). Psychometrika (心理學計量報導), 49, .

Tenenhaus, M., 和 F. W. Young. 1985. An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis, and other methods for quantifying categorical multivariate data (量化類別多變量資料的多重對應分析、最適尺度、雙尺度、均質性分析及其他方法之分析與結合). Psychometrika (心理學計量報導), 50, .

Van Rijckevorsel, J. 1987. The application of fuzzy coding and horseshoes in multiple correspondence analysis (多重對應分析之模糊編碼與馬蹄應用). Leiden (萊頓): DSWO Press (DSWO 出版社).

章節
14

# 多元尺度方法

如果指定了一組物件，多元尺度方法的目標就是找出物件在低維度空間中的表示法。這是透過使用物件之間的**相似性**找到解答。此程序將原始、可能經過轉換後、物件相似性，及其在低維度空間中的歐基里得直線距離之間的離差平方減到最低。

低維度空間的目的就是找出物件之間的關係。透過將解答限制為自變數的線性組合，您可能可以依這些變數解釋解答的維度。在下面範例中，您將看到 15 個親屬稱謂在三維度中的表示方式，以及依性別、世代和各稱謂間的分離度來解釋該空間的方式。

## 範例：關係稱呼的檢查

Rosenberg 與 Kim (Rosenberg 和 Kim, 1975) 致力於分析 15 個親屬關係稱呼 (aunt(姑/姨)、brother (兄弟)、cousin (堂/表兄弟姐妹)、daughter (女兒)、father (父親)、granddaughter (孫女)、grandfather (祖父)、grandmother (祖母)、grandson (孫子)、mother (母親)、nephew (姪子/外甥)、niece (姪女/外甥女)、sister (姐妹)、son (兒子)、uncle (叔/舅父))。他們請四組大學生 (兩組女性、兩組男性) 根據其相似性來分類整理這些稱謂。他們請其中兩組 (一組女性、一組男性) 作兩次分類整理，第二次要根據與第一次不同的準則進行分類整理。總共取得了六組“來源”，請看下表的概要。

**表格 14-1**
**親屬關係資料的來源結構**

| 來源 | 性別 | 條件 | 樣本大小 |
|---|---|---|---|
| 1 | 女性 | 單次分類整理 | 85 |
| 2 | 男性 | 單次分類整理 | 85 |
| 3 | 女性 | 第一次分類整理 | 80 |
| 4 | 女性 | 第二次分類整理 | 80 |
| 5 | 男性 | 第一次分類整理 | 80 |
| 6 | 男性 | 第二次分類整理 | 80 |

每個來源對應於 15 × 15 個相似性矩陣，其儲存格等於來源中人數減去物件在該來源中分為同組的次數。可以在 kinship_dat.sav 中找到此資料集。 如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

© Copyright SPSS Inc. 1989, 2010

## 選擇維度數

由您決定解答應該有的維度數。陡坡圖可以協助您作決定。

▶ 若要建立陡坡圖，請從功能表選擇：
分析 > 尺度 > 多元尺度方法 (PROXSCAL)...

圖表 14-1
資料格式對話方塊

▶ 在「來源數」組別中選擇「多重矩陣來源」。

▶ 按一下「定義」。

圖表 14-2
「多元尺度方法」對話方塊

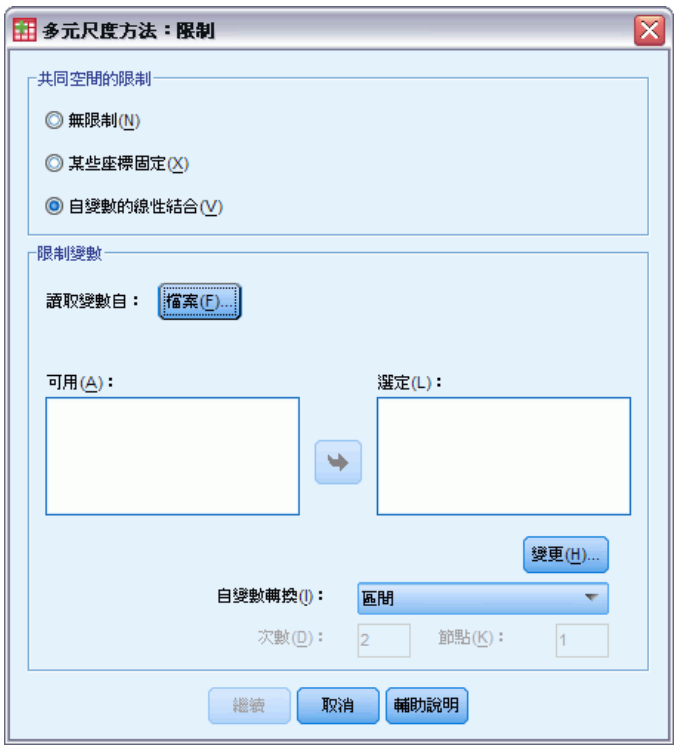
- 選擇 Aunt 一直到 Uncle 作為相似性變數。
- 選擇 sourceid 作為識別來源的變數。
- 按一下「模式」。

圖表 14-3
「模式」對話方塊

- 輸入 10 作為最大維度數。
- 按一下「繼續」。

▶ 在「多元尺度方法」對話方塊中按一下「限制」。

圖表 14-4
限制對話方塊




▶ 選擇「自變數的線性組合」。

▶ 按一下「檔案」，選擇自變數的來源。

▶ 選擇 kinship_var.sav。

圖表 14-5
限制對話方塊

▶ 選擇 性別、世代和度作為限制變數。

請注意，變數性別有使用者定義的遺漏值 — 9 = 遺漏值（代表 cousin(堂/表兄弟姐妹)）。此程序將此視為有效類別。因此，預設線性轉換不可能合適。改為使用名義轉換。

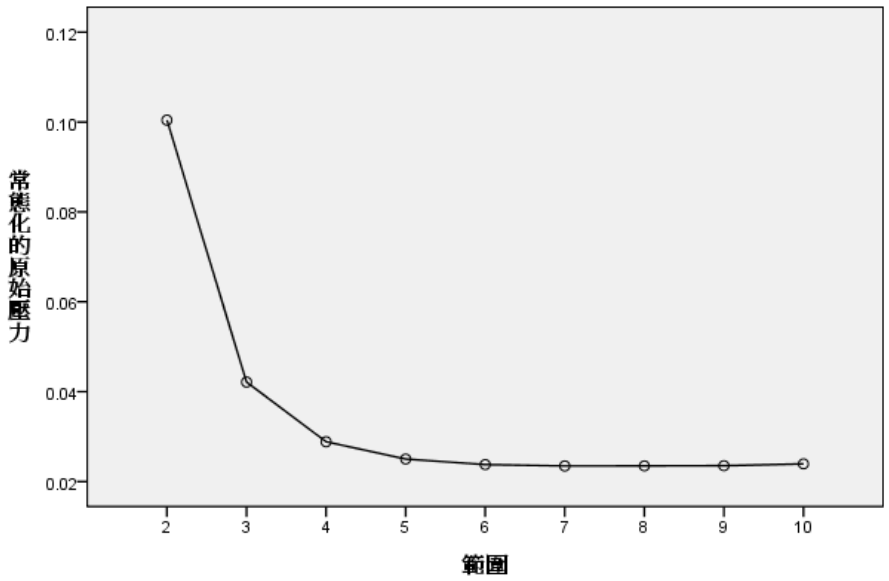
圖表 14-6
限制對話方塊

- 選擇「性別」。
- 從「自變數轉換」下拉式清單中選擇「名義」。
- 按一下「變更」。
- 按一下「繼續」。
- 在「多元尺度方法」對話方塊中按一下「圖形」。

圖表 14-7
「圖形」對話方塊

- 在「圖形」組別中選擇「**壓力係數**」。
- 按一下「**繼續**」。
- 在「多元尺度方法」對話方塊中按一下「**確定**」。

圖表 14-8
陡坡圖




此程序是以十維解開始，一直往下做到二維解。陡坡圖顯示在各維度中解答的常態化原始壓力。您可以從圖形中看出，將維度從 2 增加到 3 以及從 3 增加到 4 可以提供壓力係數大規模的改善。在 4 之後，改善情形相當小。您將選擇使用 3 維解來分析資料，因為結果比較容易解釋。

## 三維解

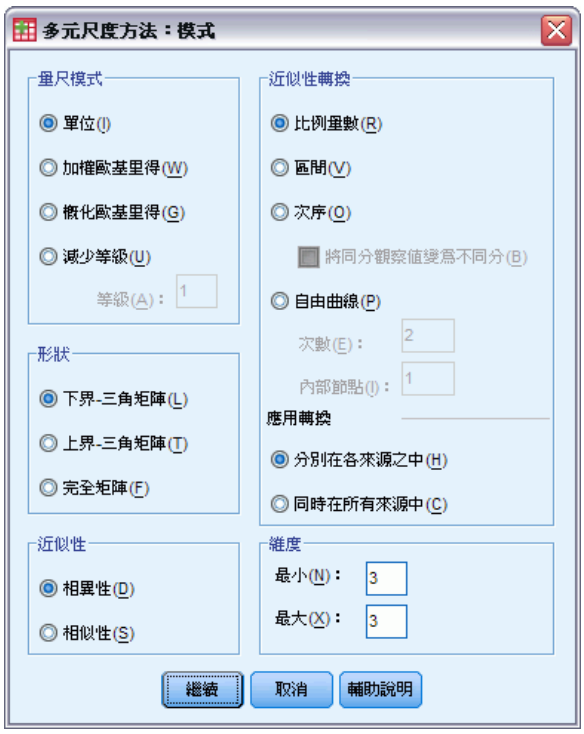
自變數性別、世代和（分離）度在建構時原本的用意是要用來解釋解答的維度。自變數建構如下：

| | |
|---|---|
| 性別 | 1 = 男性、2 = 女性、9 = 遺漏值（代表 cousin(堂/表兄弟姐妹)） |
| 世代 | 以您的親屬關係來看，就是從您開始的世代數，數字較低對應於較老世代。因此祖母是 - 2，孫子是 2，而兄弟姐妹是 0。 |
| 度 | 在家族樹上的分離度數。所以，您的父母是向上 1 個節點，而您的子女則是向下 1 個節點。您的兄弟姐妹則是向上 1 個節點到您的父母，然後再向下 1 個節點到他們，成為 2 度分離。您的堂/表兄弟姐妹則是離 4 度 — 2 度向上到您的祖父母，然後再向下 2 度通過您的姑/姨/叔/舅到他們。 |

外部變數可以在 kinship_var.sav 中找到。此外，從較早分析取得的起始組態在 kinship_ini.sav 中提供。 如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

## 執行分析

圖表 14-9
「模式」對話方塊




- 若要取得三維解，請叫回「多元尺度方法」對話方塊，然後按一下「**模式**」。
- 輸入 3 作為最小和最大維度數。
- 按一下「**繼續**」。
- 在「多元尺度方法」對話方塊中按一下「**選項**」。

圖表 14-10
「選項」對話方塊

- 選擇「自訂」作為起始組態。
- 選擇 kinship_ini.sav 作為要從其中讀取變數的檔案。
- 選擇 dim01、dim02 和 dim03 作為變數。
- 按一下「繼續」。
- 在「多元尺度方法」對話方塊中按一下「圖形」。

圖表 14-11
「圖形」對話方塊

- 選擇「原始對轉換後的相似性」和「轉換後的自變數」。
- 按一下「繼續」。
- 在「多元尺度方法」對話方塊中按一下「輸出」。

圖表 14-12
「輸出」對話方塊

- 選擇「輸入資料」、「壓力係數分解」，以及「變數與維度之相關」。
- 按一下「繼續」。
- 在「多元尺度方法」對話方塊中按一下「確定」。

## 壓力係數測度

壓力係數與適合度量數可以指出，解答中的距離與原始距離的相似性。

圖表 14-13
壓力係數與適合度量數

| | |
|---|---|
| 常態化的 原始壓力 | .06234 |
| 壓力-I | .24968[a] |
| 壓力-II | .87849[a] |
| S-壓力 | .14716[b] |
| 離散歸因於 (D.A.F.) | .93766 |
| Tucker's 全等係數 | .96833 |

PROXSCAL 把常態化原始壓力減到最小。

a. 最適尺度因子 = 1.066。

b. 最適尺度因子 = .984。

四個壓力係數的每個統計量可以測量資料的不適合度，而分散情形歸因於及 Tucker’s 一致的係數值可以測量適合度。較低的壓力係數量數（到最小值 0）和較高適合度量數（到最大值 1）表示較佳的解答。

**圖表 14-14**
**常態化原始壓力分解**

| | | 來源 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | SRC_1 | SRC_2 | SRC_3 | SRC_4 | SRC_5 | SRC_6 | 平均數 |
| 個體 | Aunt | .0991 | .0754 | .0629 | .0468 | .0391 | .0489 | .0620 |
| | Brother | .1351 | .0974 | .0496 | .0813 | .0613 | .0597 | .0807 |
| | Cousin | .0325 | .0336 | .0480 | .0290 | .0327 | .0463 | .0370 |
| | Daughter | .0700 | .0370 | .0516 | .0229 | .0326 | .0207 | .0391 |
| | Father | .0751 | .0482 | .0521 | .0225 | .0272 | .0298 | .0425 |
| | Granddaughter | .1410 | .0736 | .0801 | .0707 | .0790 | .0366 | .0802 |
| | Grandfather | .1549 | .1057 | .0858 | .0821 | .0851 | .0576 | .0952 |
| | Grandmother | .1550 | .0979 | .0858 | .0844 | .0816 | .0627 | .0946 |
| | Grandson | .1374 | .0772 | .0793 | .0719 | .0791 | .0382 | .0805 |
| | Mother | .0813 | .0482 | .0526 | .0229 | .0260 | .0227 | .0423 |
| | Nephew | .0843 | .0619 | .0580 | .0375 | .0317 | .0273 | .0501 |
| | Niece | .0850 | .0577 | .0503 | .0353 | .0337 | .0260 | .0480 |
| | Sister | .1361 | .0946 | .0496 | .0816 | .0629 | .0588 | .0806 |
| | Son | .0689 | .0373 | .0456 | .0242 | .0337 | .0253 | .0392 |
| | Uncle | .0977 | .0761 | .0678 | .0489 | .0383 | .0498 | .0631 |
| 平均數 | | .1035 | .0681 | .0613 | .0508 | .0496 | .0407 | .0623 |

壓力係數分解可協助您識別對解答之整體壓力係數貢獻最多的來源和物件。在此範例中，來源中的大部份壓力可歸因於來源 1 和 2，而在物件之間，大部份壓力可歸因於 Brother（兄弟姐妹）、Granddaughter（孫女）、Grandfather（祖父）、Grandmother（祖母）、Grandson（孫子）和 Sister（姐妹）。

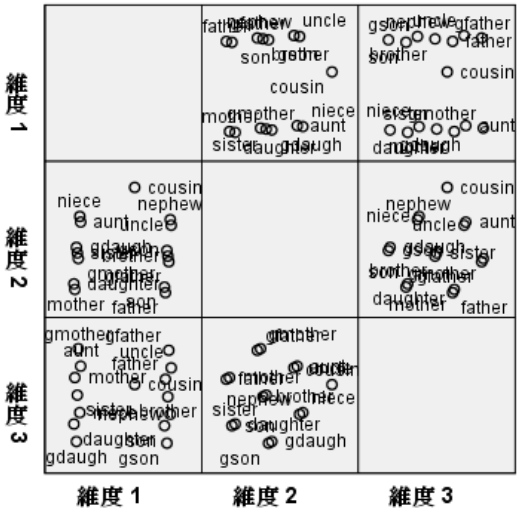
兩個可解釋大部份壓力的來源是對親屬稱謂只作一次分類整理的兩組。這項資訊表示：學生在分類親屬稱謂時考慮多重因子，而允許作兩次分類整理的學生則專注於第一次分類的部份因子，然後在第二次分類時考慮其餘因子。

能夠解釋大部份壓力的物件是分離 2 度的物件。這些人是與“核心”家庭（Mother（母親）、Father（父親）、但 Daughter（女兒）、Son（兒子））無關的，但還是比其他關係要來得親近。這個中間位置很容易造成這些親屬稱謂的排序差異。

## 共同空間的最終座標

共同空間圖形提供物件間關係的視覺呈現。

圖表 14-15
共同空間座標



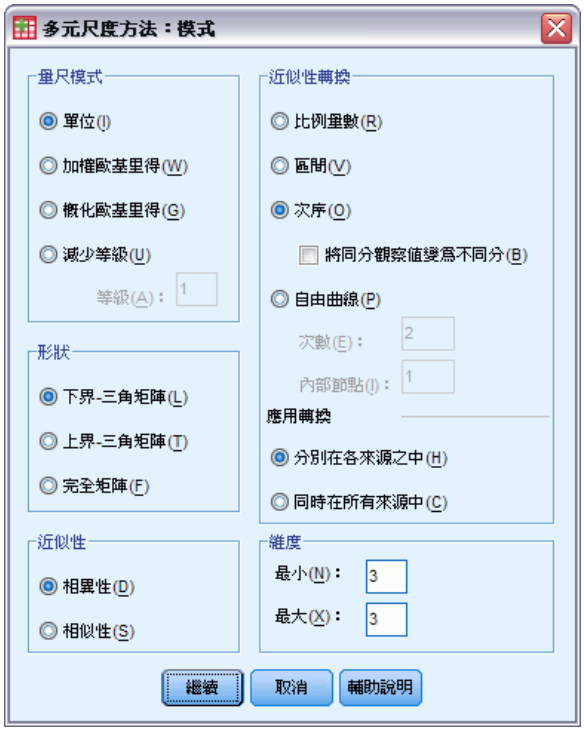

查看一下維度 1 和 3 中物件的最終座標，這是散佈圖矩陣左下角的圖形。這個圖形顯示：維度 1（在 x 軸上）是與變數性別 相關，而維度 3（在 y 軸上）則與世代相關。您可以看到由左至右，維度 1 分隔了女性與男性親屬稱謂，而無性別區分的親屬稱謂 Cousin（堂/表兄弟姐妹）則位於中間。圖形中由底端至頂端，沿軸增加的值對應於年紀較大的親屬稱謂。

現在查看一下維度 2 和 3 物件的最終座標，這是位於散佈圖矩陣中間的右半邊圖形。您可以從這個圖形看出來，第二維度（沿著 y 軸）對應於變數度，其較大值沿軸對應於遠離「核心」家庭的親屬稱謂。

## 非預設轉換的三維解

上一個解答是利用自變數世代和度之相似性與間隔轉換的預設比例轉換進行計算。結果相當不錯，但您可能還可以使用其他轉換，取得更好的結果。例如，相似性、世代和度都有自然的順序，但次序轉換的模式化作業可能會優於線性轉換。

圖表 14-16
「模式」對話方塊




- 若要重新執行分析，在次序水準（維持關係）使用相似性、世代和度尺度，叫回「多元尺度方法」對話方塊，然後按一下「**模式**」。
- 選擇「**次序**」作為相似性轉換。
- 按一下「**繼續**」。
- 在「多元尺度方法」對話方塊中按一下「**限制**」。

圖表 14-17
限制對話方塊

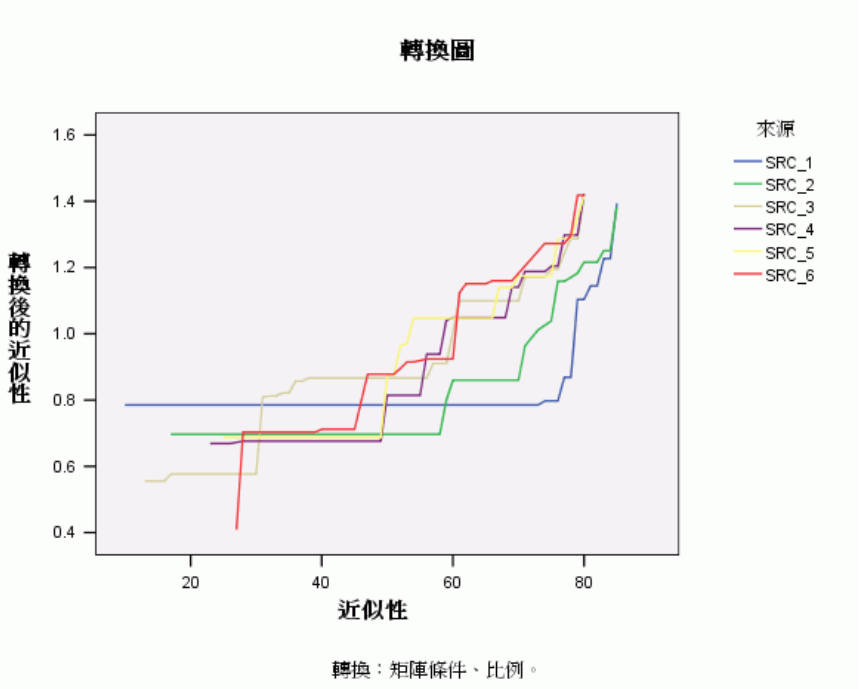
- 選擇「世代」和「度」。
- 從「自變數轉換」下拉式清單中選擇「**次序（維持關係）**」。
- 按一下「**變更**」。
- 按一下「**繼續**」。
- 在「多元尺度方法」對話方塊中按一下「**確定**」。

## 轉換圖

轉換圖很適合用來做第一次檢查，查看原始轉換是否適當。如果圖形看來近似線性，則表示線性假設是合適的。若不似線性就要檢查壓力係數量度，查看配適度是否有改善，同時檢查共同空間圖形，查看解釋是否更有用。

每個自變數都各自取得近似線性轉換，因此將其解釋為數值可能較適合。但是相似性不會取得線性轉換，因此次序轉換可能比相似性更適合。

圖表 14-18
轉換後的相似性




## 壓力係數測度

目前解答的壓力係數支持在次序水準使用相似性尺度的論點。

圖表 14-19
壓力係數與適合度量數

| | |
|---|---|
| 常態化的 原始壓力 | .03137 |
| 壓力-I | .17712[a] |
| 壓力-II | .61987[a] |
| S-壓力 | .07953[b] |
| 離散歸因於 (D.A.F.) | .96863 |
| Tucker's 全等係數 | .98419 |

PROXSCAL 把常態化原始壓力減到最小。

a. 最適尺度因子 = 1.032。

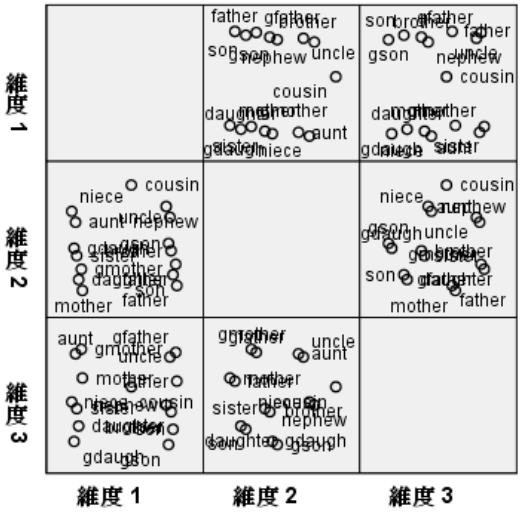
b. 最適尺度因子 = .980。

上一個解答的常態化原始壓力係數是 0.06234。使用非預設轉換的變數尺度將壓力係數減半為 0.03137。

## 共同空間的最終座標

共同空間圖形基本上提供了與上一個解答相同的維度解釋。

圖表 14-20
共同空間座標




## 討論

將相似性視為次序變數是最佳作法，因為在壓力係數測度上有極大的改善。因此下一個步驟，您可能要將次序變數“解除關係” — 也就是說，允許原始變數具有相等的值，以取得不同的轉換後值。例如，在第一個來源中，Aunt (姑/姨) 與 Son (兒子)，以及 Aunt (姑/姨) 與 Grandson (孫子) 之間的相似性是 85；以“維持關係”手法處理次序變數迫使這些相似性的轉換後值相等，但您並沒有特別的理由做此假設。在此情形下，允許相似性解除關係，可以讓您解除不必要的限制。

# 閱讀資料推薦

請參閱下列文字以取得有關多元尺度方法的詳細資訊：

Commandeur, J. J. F., 和 W. J. Heiser. 1993. Mathematical derivations in the proximity scaling (PROXSCAL) of symmetric data matrices (對稱資料矩陣相似性尺度 (PROXSCAL) 的數學應用). Leiden (萊頓): Department of Data Theory, University of Leiden (萊頓大學資料理論系).

De Leeuw, J., 和 W. J. Heiser. 1980. Multidimensional scaling with restrictions on the configuration (限制組態下多元尺度方法). 於： Multivariate Analysis, Vol. V (多變量分析第五冊), P. R. Krishnaiah, ed. Amsterdam (阿姆斯特丹): North-Holland (北荷蘭).

Heiser, W. J. 1981. Unfolding analysis of proximity data (相似性資料的展開分析). Leiden (萊頓): Department of Data Theory, University of Leiden (萊頓大學資料理論系).

Heiser, W. J., 和 F. M. T. A. Busing. 2004. Multidimensional scaling and unfolding of symmetric and asymmetric proximity relations (對稱與非對稱相似關係之多元尺度與展開方法). 於： Handbook of Quantitative Methodology for the Social Sciences (社會學量化方法論手冊), D. Kaplan，ed. Thousand Oaks, Calif. (加州千橡): Sage Publications, Inc..

Kruskal, J. B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis (將適合度優化為非計量假設之多元尺度方法). Psychometrika (心理學計量報導), 29, .

Kruskal, J. B. 1964. Nonmetric multidimensional scaling: (非計量多元尺度方法)A numerical method (數值方法). Psychometrika (心理學計量報導), 29, .

Shepard, R. N. 1962. The analysis of proximities: (相似性分析)Multidimensional scaling with an unknown distance function I (未知距離函數多元尺度法 I). Psychometrika (心理學計量報導), 27, .

Shepard, R. N. 1962. The analysis of proximities: (相似性分析)Multidimensional scaling with an unknown distance function I (未知距離函數多元尺度法 II). Psychometrika (心理學計量報導), 27, .

章節
15

# 多維度展開

「多維度展開」程序會嘗試搜尋一般數值尺度，讓您以目視的方式檢查兩組個體之間的關係。

## 範例：喜愛的早餐食品

在經典研究中(Green 和 Rao, 1972)，21 名 Wharton 學院 MBA 學生及其配偶被要求為 15 項早餐食品按喜愛程度分出等級：從 1 =「最喜愛」到 15 =「最不喜愛」。此資訊收集於 breakfast_overall.sav 之中。 如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

研究結果提供了一般多維度展開演算法經常產生之簡併問題的典型範例，這種演算法是以懲罰轉換後相似性的變異係數求解(Busing, Groenen, 和 Heiser, 2005)。您將會看到簡併解，也會看到如何使用「多維度展開」解決問題，讓您判斷個人如何區分各項早餐食品。可以在 prefscal_breakfast-overall.sps 找到重製這些分析的語法。

### 產生簡併解

- 若要執行「多維度展開」分析，請從功能表中選擇：
  **分析 > 尺度 > 多維度展開 (PREFSCAL)...**

© Copyright SPSS Inc. 1989, 2010

圖表 15-1
「多維度展開」主對話方塊

- 選擇「現烤吐司」到「玉米奶油鬆餅」作為相似性變數。
- 按一下「選項」。

圖表 15-2
「選項」對話方塊

- 選擇「Spearman」作為「典型」開始的插補方法。
- 在「懲罰項」組別中，輸入 1.0 作為 Strength 參數的值，並輸入 0.0 作為 Range 參數的值。這個動作會關閉懲罰項。
- 按一下「繼續」。
- 按一下「多維度展開」對話方塊中的「確定」。

下面是在這些選擇下所產生的指令語法：

```
PREFSCAL
  VARIABLES=TP BT EMM JD CT BMM HRB TMd BTJ TMn CB DP GD CC CMB
  /INITIAL=CLASSICAL (SPEARMAN)
  /TRANSFORMATION=NONE
  /PROXIMITIES=DISSIMILARITIES
  /CRITERIA=DIMENSIONS(2,2) DIFFSTRESS(.000001) MINSTRESS(.0001)
  MAXITER(5000)
  /PENALTY=LAMBDA(1.0) OMEGA(0.0)
  /PRINT=MEASURES COMMON
  /PLOT=COMMON .
```

- 這個語法會指定從 tp（現烤土司）到 cmb（玉米奶油鬆餅）之變數分析。
- INITIAL 次指令會指定使用「Spearman 距離」來計算起始值。

- 在 PENALTY 次指令上指定的值主要是關閉懲罰項，其結果是，程序會將 Kruskal 壓力-I 減到最小，如此一來，就會形成簡併解。
- PLOT 次指令會要求共同空間的圖形。
- 所有其他參數會回復為其預設值。

## 量數

圖表 15-3
簡併解的量數

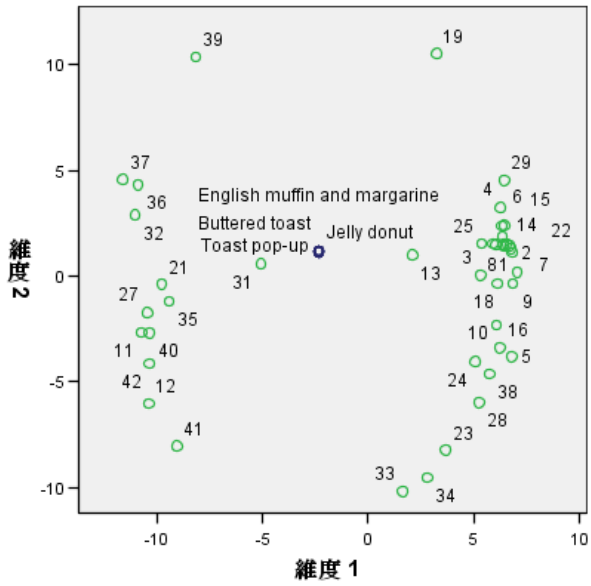
| | | |
|---|---|---|
| 疊代 | | 154 |
| 最終函數值 | | .0000990 |
| 函數值部份 | 壓力部份 | .0000990 |
| | 處罰部份 | 1.0000000 |
| 不適合度 | 正常化的壓力 | .0000000 |
| | Kruskal's Stress-I | .0000990 |
| | Kruskal's Stress-II | .6129749 |
| | Young's S-Stress-I | .0001980 |
| | Young's S-Stress-II | .7703817 |
| 適合度 | 分散歸因於 | 1.0000000 |
| | 變異歸因於 | .6230788 |
| | 復原的喜好設定順序 | .7074830 |
| | Spearman's Rho | .7450748 |
| | Kendall's Tau-b | .6218729 |
| 變異係數 | 變異相似性 | .5590170 |
| | 變異轉換相似性 | .0000924 |
| | 變異距離 | .1808765 |
| 退化指數 | DeSarbo 混雜指數的平方和 | 1.1731154E2 |
| | Shepard's 粗略非退化指數 | .0000000 |

演算法在 154 次疊代後收斂成解答，其懲罰應力（標示的最終函數值）為 0.0000990。由於懲罰項已經關閉，懲罰應力等於 Kruskal 壓力-I（函數值的應力部份相當於 Kruskal 不適合度量數）。低應力值一般都表示解答很適合該資料，但也有幾個簡併解的警告訊號：

- 轉換後相似性的變異係數與原始相似性的變異係數相比，顯得非常小。這種情形暗示各列的轉換後相似性是幾近常數，因此解答將無法提供任何物件之間的區別。
- DeSarbo 互相混合指數的平方和是代表不同組各點互相混合情況的量數。如果並未互相混合，就是解答可能為簡併的警告訊號。越趨近 0，解答的互相混合情況越好。所報告的值很大，表示解答並未互相混合。
- Shepard 粗略非簡併指數（以差異距離的百分比報告）等於 0，這個數值清楚地指出，差異距離不足，而解答大概是簡併。

# 共同空間

圖表 15-4
簡併解的共同空間聯合圖形




視覺化確認解答為簡併可在列及欄物件的共同空間聯合圖形中看出。列物件（個人）位於圍繞在欄物件（早餐食品）外的圓周上，其座標收合至單一的點上。

# 執行非簡併分析

圖表 15-5
「選項」對話方塊

- ▶ 若要產生非簡併解，請按一下「叫回對話」工具，然後選擇「**多維度展開**」。
- ▶ 按一下「多維度展開」對話方塊中的「**選項**」。
- ▶ 在「懲罰項」組別中，輸入 **0.5** 作為 Strength 參數的值，並輸入 **1.0** 作為 Range 參數的值。這個動作會關閉懲罰項。
- ▶ 按一下「**繼續**」。
- ▶ 按一下「多維度展開」對話方塊中的「**確定**」。

下面是在這些選擇下所產生的指令語法：

```
PREFSCAL
  VARIABLES=TP BT EMM JD CT BMM HRB TMd BTJ TMn CB DP GD CC CMB
  /INITIAL=CLASSICAL (SPEARMAN)
  /TRANSFORMATION=NONE
  /PROXIMITIES=DISSIMILARITIES
  /CRITERIA=DIMENSIONS(2,2) DIFFSTRESS(.000001) MINSTRESS(.0001)
  MAXITER(5000)
  /PENALTY=LAMBDA(0.5) OMEGA(1.0)
  /PRINT=MEASURES COMMON
  /PLOT=COMMON .
```

- 唯一的改變在 **PENALTY** 次指令上。**LAMBDA** 設定為 0.5，而 **OMEGA** 則設定為 1.0，都是原來的預設值。

## 量數

圖表 15-6
非簡併解的量數

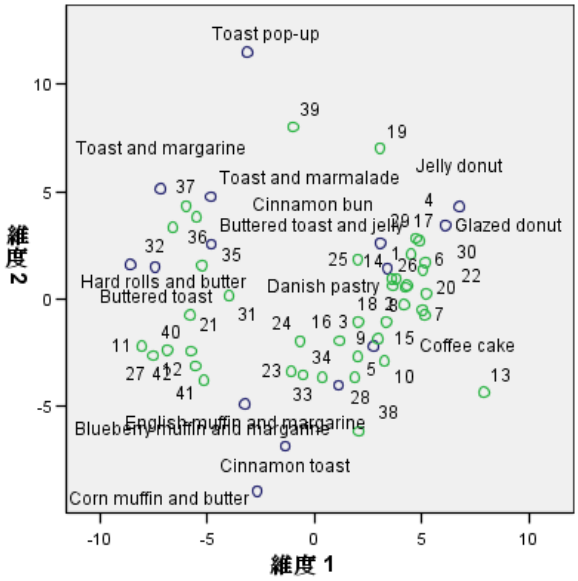
| | | |
|---|---|---|
| 疊代 | | 157 |
| 最終函數值 | | .6848930 |
| 函數值部份 | 壓力部份 | .2428268 |
| | 處罰部份 | 1.9317409 |
| 不適合度 | 正常化的壓力 | .0583589 |
| | Kruskal's Stress-I | .2415758 |
| | Kruskal's Stress-II | .5875599 |
| | Young's S-Stress-I | .3446361 |
| | Young's S-Stress-II | .5030127 |
| 適合度 | 分散歸因於 | .9416411 |
| | 變異歸因於 | .7651552 |
| | 復原的喜好設定順序 | .7818594 |
| | Spearman's Rho | .8179181 |
| | Kendall's Tau-b | .6916725 |
| 變異係數 | 變異相似性 | .5590170 |
| | 變異轉換相似性 | .6006156 |
| | 變異距離 | .4833617 |
| 退化指數 | DeSarbo 混雜指數的平方和 | .1590979 |
| | Shepard's 粗略非退化指數 | .7895692 |

簡併解量數中看到的問題此處已經校正。

- 常態化壓力不再為 0。
- 現在，轉換後相似性的變異係數值與原始相似性的變異係數值相近。
- DeSarbo 互相混合指數更接近 0，表示解答互相混合的情形好得多。
- Shepard 粗略非簡併指數（以差異距離的百分比報告）現在是將近 80%。現在有足夠的差異距離，而解答大概是非簡併。

## 共同空間

圖表 15-7
非簡併解的共同空間聯合圖形




共同空間的聯合圖形能夠解釋維度。水平維度顯示出軟、硬麵包或吐司之間的區別，沿軸向右移動時，越來越軟。垂直維度則沒有清楚的解釋，不過也許可以根據方便性作區別，沿軸往下移，越來越「正式」。

這樣會形成幾種早餐食品的集群。例如，甜甜圈、肉桂麵包和丹麥酥組成柔軟而較不正式食品的集群。鬆餅和肉桂吐司組成較硬而較正式食品的集群。其他吐司和硬麵包捲則組成硬而較不正式食品的集群。現烤吐司是硬而極不正式的食品。

由列物件代表的個人很清楚地依偏好硬或軟食品而分成不同集群，在垂直維度上有相當大的集群內變異。

# 範例：喜愛的早餐食品的三因子展開

在經典研究中(Green et al。, 1972)，21 名 Wharton 學院 MBA 學生及其配偶被要求為 15 項早餐食品按喜愛程度分出等級：從 1 =「最喜愛」到 15 =「最不喜愛」。他們的喜愛程度分六種不同情況記錄，從「整體喜愛」到「點心，僅配飲料」。此資訊收集於 breakfast.sav 之中。 如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.

六種情況可以視為不同的來源。使用 PREFSCAL 來執行列、欄和來源的三因子展開。可以在 prefscal_breakfast.sps 找到重製這些分析的語法。

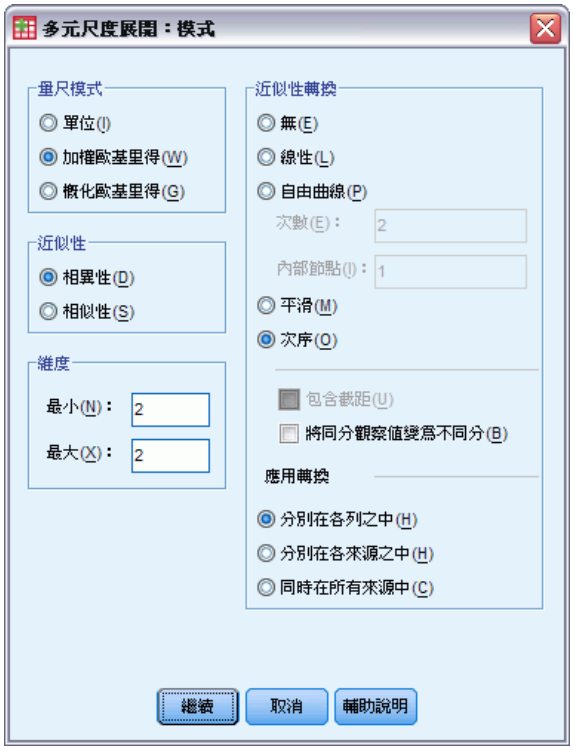
## 執行分析

▶ 若要執行「多維度展開」分析，請從功能表中選擇：
**分析 > 尺度 > 多維度展開 (PREFSCAL)...**

圖表 15-8
「多維度展開」主對話方塊

- 選擇「現烤吐司」到「玉米奶油鬆餅」作為相似性變數。
- 選擇「菜單花樣」作為來源變數。
- 按一下「模式」。

圖表 15-9
「模式」對話方塊




- 選擇「加權歐基里得」作為尺度模式。
- 按一下「繼續」。
- 按一下「多維度展開」對話方塊中的「選項」。

圖表 15-10
「選項」對話方塊

- 選擇「Spearman」作為「典型」開始的插補方法。
- 按一下「繼續」。
- 按一下「多維度展開」對話方塊中的「圖形」。

圖表 15-11
「圖形」對話方塊

- 在「圖形」組別中選擇「個別空間」。
- 按一下「繼續」。
- 按一下「多維度展開」對話方塊中的「確定」。

下面是在這些選擇下所產生的指令語法：

```
PREFSCAL
  VARIABLES=TP BT EMM JD CT BMM HRB TMd BTJ TMn CB DP GD CC CMB
  /INPUT=SOURCES(srcid )
  /INITIAL=CLASSICAL (SPEARMAN)
  /CONDITION=ROW
  /TRANSFORMATION=NONE
  /PROXIMITIES=DISSIMILARITIES
  /MODEL=WEIGHTED
  /CRITERIA=DIMENSIONS(2,2) DIFFSTRESS(.000001) MINSTRESS(.0001)
  MAXITER(5000)
```

```
/PENALTY=LAMBDA(0.5) OMEGA(1.0)
/PRINT=MEASURES COMMON
/PLOT=COMMON WEIGHTS INDIVIDUAL ( ALL ) .
```

- 這個語法會指定從 tp（現烤土司）到 cmb（玉米奶油鬆餅）之變數分析。會使用變數 srcid 來辨識來源。
- INITIAL 次指令會指定使用「Spearman 距離」來計算起始值。
- MODEL 次指令會指定加權的歐基里得模式，此模式可以讓各個個別空間對共同空間的維度進行不同的加權。
- PLOT 次指令會要求共同空間、個別空間和個別空間權重圖形。
- 所有其他參數會回復為其預設值。

## 量數

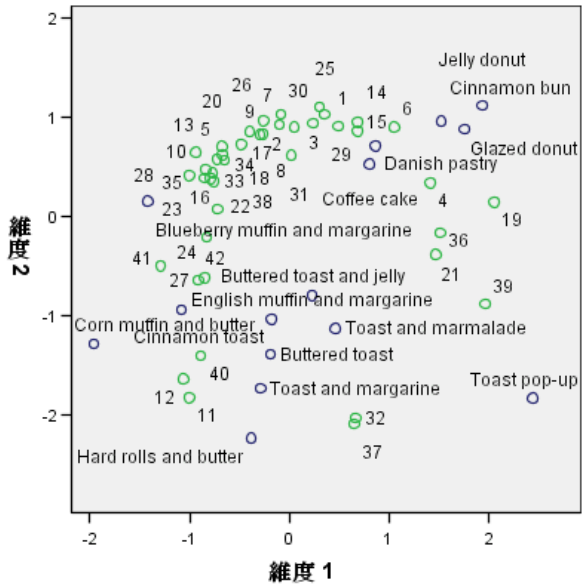
圖表 15-12
量數

| | | |
|---|---|---|
| 疊代 | | 481 |
| 最終函數值 | | .8199642 |
| 函數值部份 | 壓力部份 | .3680994 |
| | 處罰部份 | 1.8265211 |
| 不適合度 | 正常化的壓力 | .1335343 |
| | Kruskal's Stress-I | .3654234 |
| | Kruskal's Stress-II | .9780824 |
| | Young's S-Stress-I | .4938016 |
| | Young's S-Stress-II | .6912352 |
| 適合度 | 分散歸因於 | .8664657 |
| | 變異歸因於 | .5024853 |
| | 復原的喜好設定順序 | .7025321 |
| | Spearman's Rho | .6271702 |
| | Kendall's Tau-b | .4991188 |
| 變異係數 | 變異相似性 | .5590170 |
| | 變異轉換相似性 | .6378878 |
| | 變異距離 | .4484515 |
| 退化指數 | DeSarbo 混雜指數的平方和 | .2199287 |
| | Shepard's 粗略非退化指數 | .7643613 |

演算法在 481 次疊代後收斂至最終懲罰應力 0.8199642。變異係數和 Shepard 指數夠大，而 DeSarbo 指數則夠低，暗示沒有簡併的問題。

# 共同空間

圖表 15-13
共同空間的聯合圖形




共同空間的聯合圖形顯示了非常類似整體偏好二因子分析的最終組態，其解答翻越 45 度線。因此，現在垂直維度顯示出軟、硬麵包或吐司之間的區別，沿軸向上移動時，越來越軟。現在水平維度倒反而沒有清楚的解釋，不過也許可以根據方便性作區別，沿軸往左移，越來越「正式」。

由列物件代表的個人仍然清楚地依偏好「硬」或「軟」食品而分成不同集群，但在水平維度上有相當大的集群內變異。

## 個別空間

圖表 15-14
維度加權

| | | 維度 | | |
|---|---|---|---|---|
| | | 1 | 2 | 專一性[a] |
| 來源 | Overall preference | 3.235 | 4.297 | .186 |
| | Breakfast, with juice, bacon and eggs, and beverage | 4.883 | 2.193 | .457 |
| | Breakfast, with juice, cold cereal, and beverage | 4.131 | 3.438 | .109 |
| | Breakfast, with juice, pancakes, sausage, and beverage | 4.291 | 3.267 | .164 |
| | Breakfast, with beverage only | 3.124 | 4.413 | .223 |
| | Snack, with beverage only | 2.750 | 4.541 | .313 |
| | 重要性[b] | .504 | .496 | |

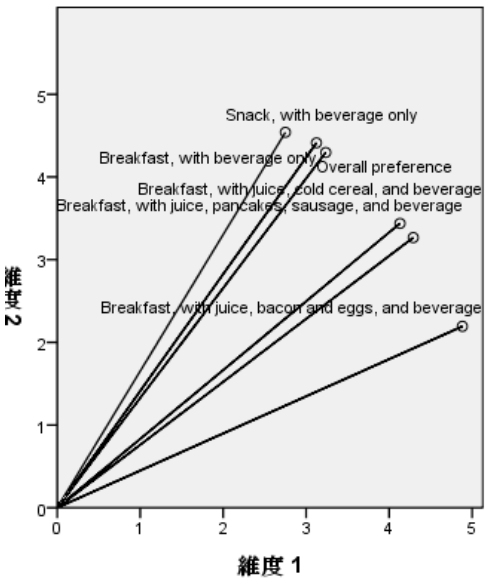
a. 專一性指示出來源的典型性。專一性的範圍是介在零和一之間，其中零表示具有相同維度加權的平均來源，而一表示具有一個異常、大維度加權及其他幾近於零的加權之特定來源。

b. 每個維度的相對重要性，即是一個維度的平方和與總平方和之間的比率。

會為每個來源計算個別空間。維度加權則顯示將個別空間載入共同空間維度中的情形。較大加權表示個別空間中的距離較大，因此在該個別空間之維度上的物件間有較大區別。

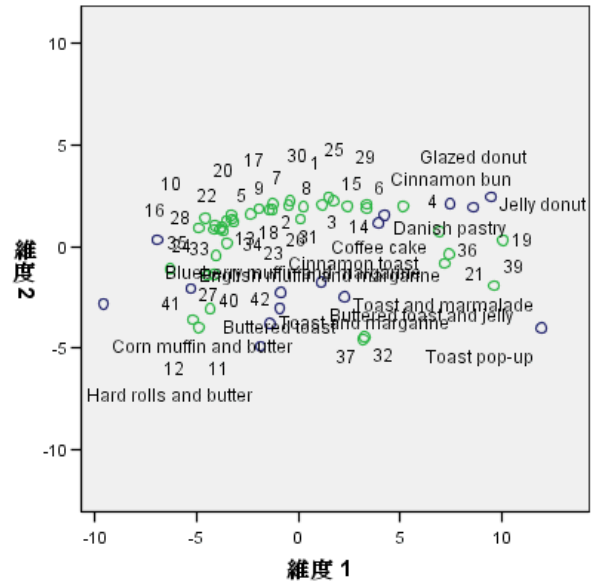
- **特異性**是個別空間與共同空間差異情況的量數。與共同空間完全相同的個別空間會有完全相同的維度加權和特異性 0，但個別空間若屬特定維度則將有單一大維度加權和特異性 1。在此情況下，最不同的來源是「早餐、果汁、培根和蛋，以及飲料」與「點心，僅配飲料」。
- **重要性**是各維度對解答之相對貢獻的量數。在此情況下，所有維度都一樣重要。

圖表 15-15
維度加權



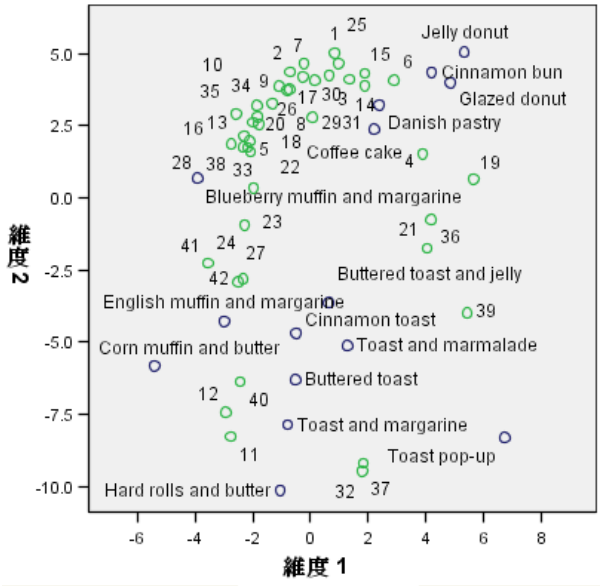

維度加權圖表提供視覺化的加權表。「早餐、果汁、培根和蛋，以及飲料」與「點心，僅配飲料」兩項最接近維度軸，但兩者對特定維度都沒有強烈的特異性。

圖表 15-16
個別空間「早餐、果汁、培根和蛋，以及飲料」的聯合圖形




個別空間「早餐、果汁、培根和蛋，以及飲料」的聯合圖形顯示這種情況對喜愛程度的影響。這個來源在第一維度上的載入負荷量較重，因此食品項之間的差異多半是來自第一維度。

圖表 15-17
個別空間「點心，僅配飲料」的聯合圖形




個別空間「點心，僅配飲料」的聯合圖形顯示這種情況對喜愛程度的影響。這個來源在第二維度上的載入負荷量較重，因此食品品項之間的差異主要是來自第二維度。但是，在第一維度上仍然還有相當大的差異，因為這個來源的特異性相當低。

## 使用不同的起始組態

最終組態可能會依指定給演算法的開始點而有所不同。最理想的情形是，解答的一般結構應該保持相同，否則可能很難確定何者為正確。不過當您嘗試不同的起始組態時，例如在早餐資料的三因子分析上使用對應的起點，詳細情況可能會更加明確。

- 若要用對應的起點產生解答，請按一下「叫回對話」工具，然後選擇「多維度展開」。

▶ 按一下「多維度展開」對話方塊中的「選項」。

圖表 15-18
「選項」對話方塊

▶ 在「起始組態」組別中選擇「對應」。

▶ 按一下「繼續」。

▶ 按一下「多維度展開」對話方塊中的「確定」。

下面是在這些選擇下所產生的指令語法：

```
PREFSCAL
  VARIABLES=TP BT EMM JD CT BMM HRB TMd BTJ TMn CB DP GD CC CMB
  /INPUT=SOURCES(srcid)
  /INITIAL=CORRESPONDENCE
  /TRANSFORMATION=NONE
  /PROXIMITIES=DISSIMILARITIES
  /CRITERIA=DIMENSIONS(2,2) DIFFSTRESS(.000001) MINSTRESS(.0001)
  MAXITER(5000)
  /PENALTY=LAMBDA(0.5) OMEGA(1.0)
  /PRINT=MEASURES COMMON
  /PLOT=COMMON WEIGHTS INDIVIDUAL (ALL).
```

■ 唯一的改變在 INITIAL 次指令上。起始組態設定為 CORRESPONDENCE，會搭配列與行分數的對稱常態化，使用在相反資料（是相似性而不是相異性）上之對應分析的結果。

# 量數

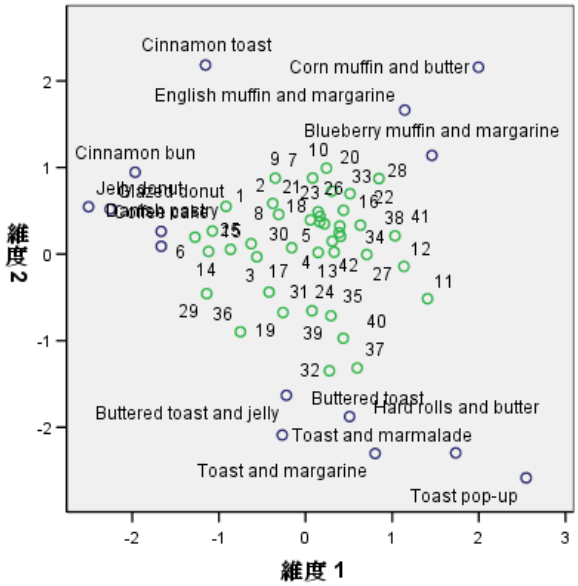
圖表 15-19
對應起始組態的量數

| | | |
|---|---|---|
| 疊代 | | 385 |
| 最終函數值 | | .8140741 |
| 函數值部份 | 壓力部份 | .3493640 |
| | 處罰部份 | 1.8969229 |
| 不適合度 | 正常化的壓力 | .1212145 |
| | Kruskal's Stress-I | .3481587 |
| | Kruskal's Stress-II | 1.0770522 |
| | Young's S-Stress-I | .4812632 |
| | Young's S-Stress-II | .6871733 |
| 適合度 | 分散歸因於 | .8787855 |
| | 變異歸因於 | .5183498 |
| | 復原的喜好設定順序 | .7174981 |
| | Spearman's Rho | .6446272 |
| | Kendall's Tau-b | .5165230 |
| 變異係數 | 變異相似性 | .5590170 |
| | 變異轉換相似性 | .6122308 |
| | 變異距離 | .4043695 |
| 退化指數 | DeSarbo 混雜指數的平方和 | 1.7571887 |
| | Shepard's 粗略非退化指數 | .7532124 |

演算法在 385 次疊代後收斂至最終懲罰應力 0.8140741。這項統計量、不適合度、適合度、變異係數，以及 Shepard 指數皆與使用典型 Spearman 開始來求解所得之數值非常類似。DeSarbo 指數則有些差異，值為 1.7571887 而不是 0.2199287，這暗示使用對應起點的解答混合情況並不是很好。若要檢視這種情形如何影響解答，請看共同空間的聯合圖形。

# 共同空間

圖表 15-20
對應起始組態的共同空間聯合圖形




共同空間的聯合圖形顯示出類似以典型 Spearman 起始組態來分析的最終組態，但是欄物件（早餐食品品項）位於列物件（個人）周遭而不是與列物件互相混合。

# 個別空間

圖表 15-21
對應起始組態的維度加權

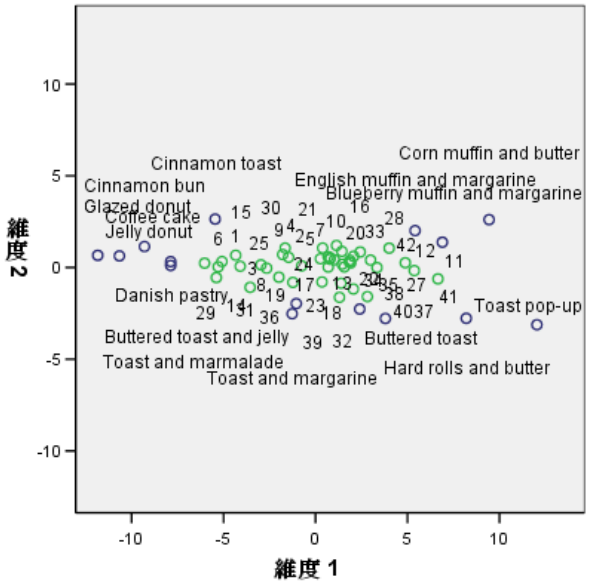
| | | 維度 | | |
|---|---|---|---|---|
| | | 1 | 2 | 專一性[a] |
| 來源 | Overall preference | 2.836 | 3.877 | .279 |
| | Breakfast, with juice, bacon and eggs, and beverage | 4.727 | 1.207 | .636 |
| | Breakfast, with juice, cold cereal, and beverage | 4.183 | 2.377 | .263 |
| | Breakfast, with juice, pancakes, sausage, and beverage | 4.412 | 1.993 | .389 |
| | Breakfast, with beverage only | 2.605 | 4.050 | .351 |
| | Snack, with beverage only | 1.864 | 4.415 | .552 |
| | 重要性[b] | .556 | .444 | |

a. 專一性指示出來源的典型性。專一性的範圍是介在零和一之間，其中零表示具有相同維度加權的平均來源，而一表示具有一個異常、大維度加權及其他幾近於零的加權之特定來源。

b. 每個維度的相對重要性，即是一個維度的平方和與總平方和之間的比率。

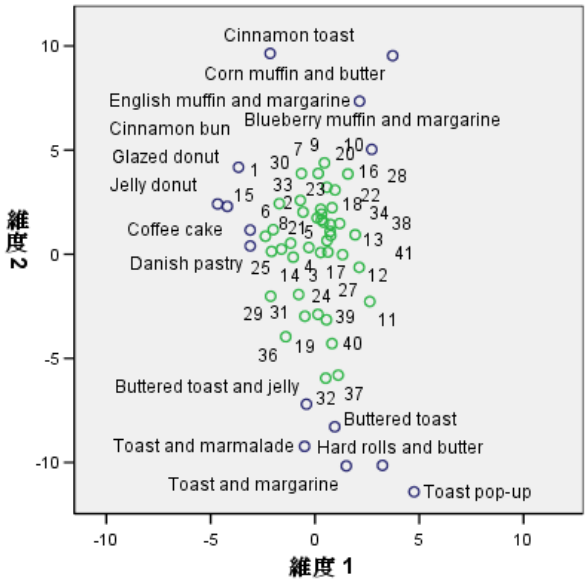
在對應起始組態下，各個別空間都有較高的特異性，也就是說，受訪者個人對早餐食品評等的各種情境與特定維度的關聯更強。差別最大的來源仍然是「早餐、果汁、培根和蛋，以及飲料」與「點心，僅配飲料」。

**圖表 15-22**
**對應起始組態下個別空間「早餐、果汁、培根和蛋，以及飲料」的聯合圖形**




較高特異性在個別空間「早餐、果汁、培根和蛋，以及飲料」的聯合圖形中很明顯。來源負荷量在第一維度甚至更重於典型 Spearman 開始，因此列和欄物件在垂直軸上顯示的變異比較少，而在水平軸上顯示的變異比較多。

圖表 15-23
對應起始組態下個別空間「點心，僅配飲料」的聯合圖形




個別空間「點心，僅配飲料」的聯合圖形顯示出列和欄物件的位置比使用典型 Spearman 開始更接近垂直線。

# 範例：檢查行為 - 情境適切性

在典型範例 (Price 和 Bouffard, 1974) 之中，52 名學生被要求為 15 種情境與 15 種行為組合評等，等級共分為 10 點，從 0 =「非常適當」 到 9 =「非常不適當」。平均值超過個別值，值會被視為相異性。

此資訊收集於 behavior.sav 之中。如需詳細資訊，請參閱第 273 頁附錄 A 中的範例檔案.使用「多維度展開」，您就可以尋找相似情境的集群，以及這些集群最為密切相關的行為。可以在 prefscal_behavior.sps 找到重製這些分析的語法。

## 執行分析

▶ 若要執行「多維度展開」分析，請從功能表中選擇：
分析 > 尺度 > 多維度展開 (PREFSCAL)...

圖表 15-24
「多維度展開」主對話方塊

- 選擇 Run 到 Shout 作為相似性變數。
- 選擇 ROWID 作為列變數。
- 按一下「模式」。

圖表 15-25
「模式」對話方塊

- 選擇「線性」作為相似性轉換，然後選擇「包含截距」。

- 選擇「同時在所有來源上」套用轉換。
- 按一下「繼續」。
- 按一下「多維度展開」對話方塊中的「選項」。

圖表 15-26
「選項」對話方塊

- 在「起始組態」組別中選擇「自訂」。
- 瀏覽並選擇 behavior_ini.sav 作為包含自訂起始組態的檔案。 如需詳細資訊，請參閱 第 273 頁附錄 A 中的範例檔案.
- 選擇 dim1 和 dim2 作為變數，指定起始組態。
- 按一下「繼續」。
- 按一下「多維度展開」對話方塊中的「圖形」。

圖表 15-27
「圖形」對話方塊

- 在「圖形」組別中選擇「**轉換圖形**」。
- 按一下「**繼續**」。
- 按一下「多維度展開」對話方塊中的「**確定**」。

下面是在這些選擇下所產生的指令語法：

```
PREFSCAL
  VARIABLES=Run Talk Kiss Write Eat Sleep Mumble Read Fight Belch Argue Jump
  Cry Laugh Shout
  /INPUT=ROWS(ROWID )
  /INITIAL=('samplesDirectory/behavior_ini.sav' )
  dim1 dim2
  /CONDITION=UNCONDITIONAL
  /TRANSFORMATION=LINEAR (INTERCEPT)
  /PROXIMITIES=DISSIMILARITIES
```

```
/MODEL=IDENTITY
/CRITERIA=DIMENSIONS(2,2) DIFFSTRESS(.000001) MINSTRESS(.0001)
MAXITER(5000)
/PENALTY=LAMBDA(0.5) OMEGA(1.0)
/PRINT=MEASURES COMMON
/PLOT=COMMON TRANSFORMATIONS .
```

- 這條語法指定對變數 run 到 shout 進行分析。會使用變數 rowid 來辨識列。
- INITIAL 次指令指定起始值要取自 behavior_ini.sav 檔。列和欄座標堆疊起來，行座標接著列座標。
- CONDITION 次指令指定所有相似性都可以彼此互相比較。在此分析中確實是這樣，因為我們應該能夠比較在公園中跑步與在教堂中跑步的相似性，而可以看出一種行為跟可能會比另外一種行為更不適當。
- TRANSFORMATION 次指令指定有截距之相似性的線性轉換。如果相似性的 1 點差異在全部 10 點等級的範圍中都相等，這種作法沒有問題。也就是說，如果學生們指定了分數，而讓 0 與 1 之間的差異與 5 與 6 之間的差異相同，則線性轉換很合適。
- PLOT 次指令會要求共同空間的圖形和轉換圖形。
- 所有其他參數會回復為其預設值。

## 量數

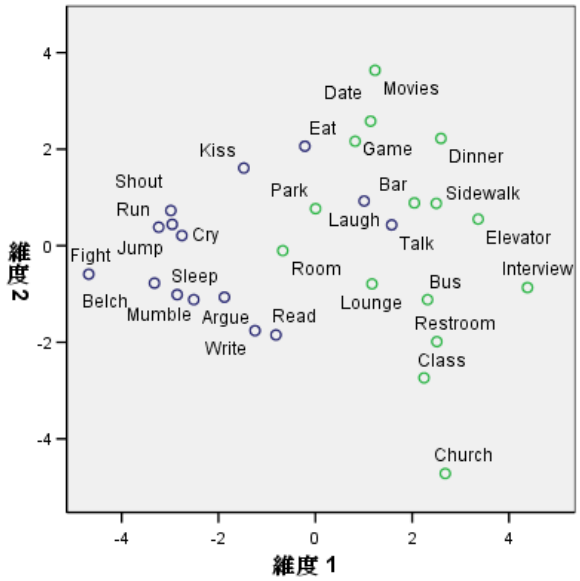
圖表 15-28
量數

| | | |
|---|---|---|
| 疊代 | | 169 |
| 最終函數值 | | .6427725 |
| 函數值部份 | 壓力部份 | .1900001 |
| | 處罰部份 | 2.1745069 |
| 不適合度 | 正常化的壓力 | .0361000 |
| | Kruskal's Stress-I | .1900001 |
| | Kruskal's Stress-II | .5224668 |
| | Young's S-Stress-I | .2760971 |
| | Young's S-Stress-II | .4525933 |
| 適合度 | 分散歸因於 | .9639000 |
| | 變異歸因於 | .8082862 |
| | 復原的喜好設定順序 | .8608333 |
| | Spearman's Rho | .8981120 |
| | Kendall's Tau-b | .7202452 |
| 變異係數 | 變異相似性 | .5138436 |
| | 變異轉換相似性 | .4751934 |
| | 變異距離 | .3912592 |
| 退化指數 | DeSarbo 混雜指數的平方和 | .4957969 |
| | Shepard's 粗略非退化指數 | .7173810 |

演算法在 169 次疊代後收斂至最終懲罰應力 0.6427725。變異係數和 Shepard 指數夠大，而 DeSarbo 指數則夠低，暗示沒有簡併的問題。

# 共同空間

圖表 15-29
共同空間的聯合圖形

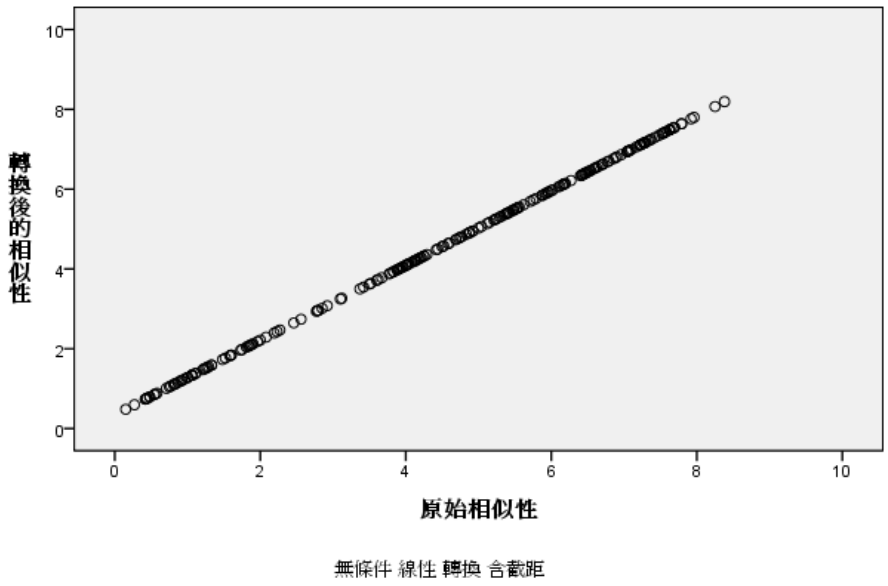



水平維度看起來與欄物件（行為）更密切相關，並區別出不適當行為（打架、抱怨）與較適當行為。垂直維度看起來與列物件（情境）更密切相關，且定義不同的情境－行為限制。

- 接近垂直維度底端的是情境（教堂、教室），將行為限制於更安靜/更內省的行為（閱讀、寫作）。因此，會將這些行為向下拉近垂直軸。
- 接近垂直維度頂端的是情境（電影、遊戲、約會），將行為限制於社交/外向行為（吃、親吻、笑）。因此，會將這些行為向上拉近垂直軸。
- 垂直維度中央的情境會依一般情境限制在水平維度上做區分。距離行為（訪問）越遠的最受限制，而距離行為（房間、公園）較近的一般限制比較少。

## 相似性轉換

圖表 15-30
轉換圖




無條件 線性 轉換 含截距

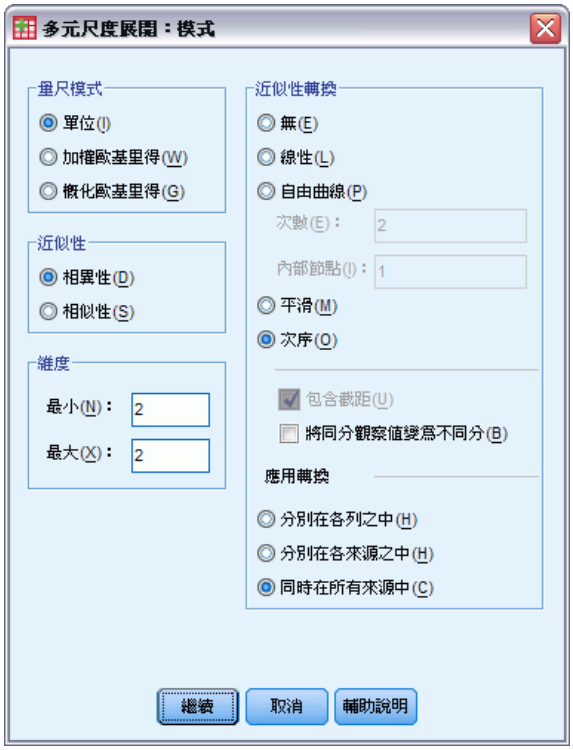
在此分析中，相似性以線性方式處理，因此轉換後值與原始相似性相對的圖形形成直線。此解的適合度相當好，但也許用相似性的另一種轉換會達到更好的適合度。

## 變更相似性轉換（次序）

- 若要產生相似性次序轉換解，請按一下「叫回對話」工具，然後選擇「多維度展開」。

▶ 按一下「多維度展開」對話方塊中的「模式」。

圖表 15-31
「模式」對話方塊




▶ 選擇「次序」作為相似性轉換。

▶ 按一下「繼續」。

▶ 按一下「多維度展開」對話方塊中的「確定」。

下面是在這些選擇下所產生的指令語法：

```
PREFSCAL
 VARIABLES=Run Talk Kiss Write Eat Sleep Mumble Read Fight Belch Argue Jump
 Cry Laugh Shout
 /INPUT=ROWS(ROWID )
 /INITIAL=('samplesDirectory/behavior_ini.sav' )
 dim1 dim2
 /CONDITION=UNCONDITIONAL
 /TRANSFORMATION=ORDINAL (KEEPTIES)
 /PROXIMITIES=DISSIMILARITIES
 /MODEL=IDENTITY
 /CRITERIA=DIMENSIONS(2,2) DIFFSTRESS(.000001) MINSTRESS(.0001)
 MAXITER(5000)
 /PENALTY=LAMBDA(0.5) OMEGA(1.0)
 /PRINT=MEASURES COMMON
 /PLOT=COMMON TRANSFORMATIONS .
```

■ 唯一的改變在 TRANSFORMATION 次指令上。已將轉換設定為 ORDINAL，保留了相似性的次序，但轉換後的值並不需要與原始值成比例。

# 量數

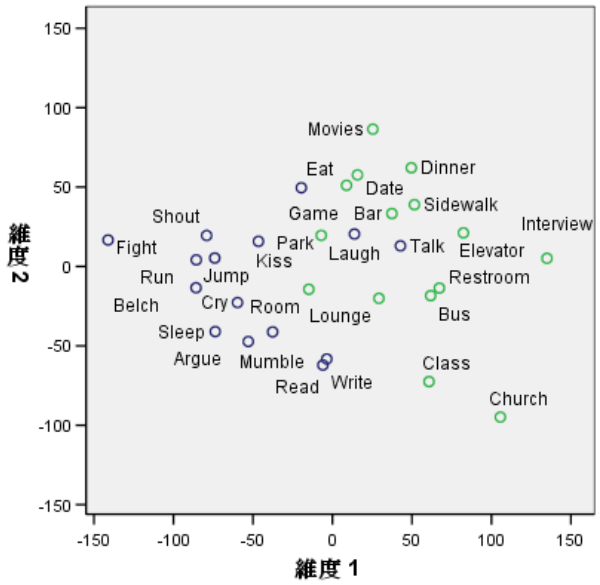
圖表 15-32
次序轉換解的量數

| | | |
|---|---|---|
| 疊代 | | 268 |
| 最終函數值 | | .6044671 |
| 函數值部份 | 壓力部份 | .1747239 |
| | 處罰部份 | 2.0911875 |
| 不適合度 | 正常化的壓力 | .0305285 |
| | Kruskal's Stress-I | .1747239 |
| | Kruskal's Stress-II | .4444641 |
| | Young's S-Stress-I | .2707147 |
| | Young's S-Stress-II | .3978003 |
| 適合度 | 分散歸因於 | .9694715 |
| | 變異歸因於 | .8454488 |
| | 復原的喜好設定順序 | .8574206 |
| | Spearman's Rho | .9032676 |
| | Kendall's Tau-b | .7532788 |
| 變異係數 | 變異相似性 | .5138436 |
| | 變異轉換相似性 | .4930018 |
| | 變異距離 | .4284849 |
| 退化指數 | DeSarbo 混雜指數的平方和 | .3610680 |
| | Shepard's 粗略非退化指數 | .7469048 |

演算法在 268 次疊代後收斂至最終懲罰應力 0.6044671。此解的這個統計量及其他量數比使用相似性線性轉換所得的解稍微好一點。

# 共同空間

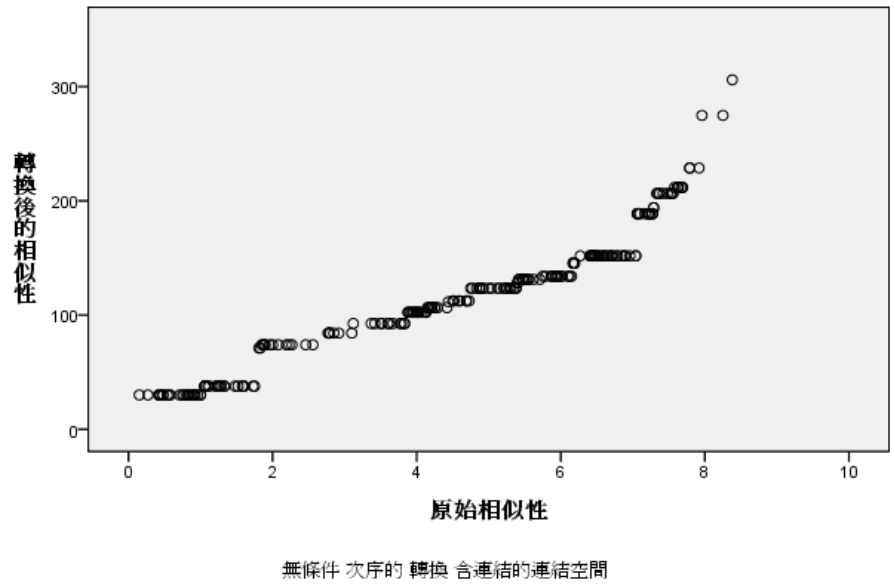
圖表 15-33
次序轉換解的共同空間聯合圖形




兩種解答的共同空間解釋是相同的。也許此解（次序轉換）的變異在垂直維度上相對地比水平維度小，也比線性轉換解小。

## 相似性轉換

圖表 15-34
次序轉換解的轉換圖形




除了最大相似性值（此值與其他值相異彎離）以外，相似性次序轉換線性相當強。這些相似性很可能解釋了次序與線性解之間絕大部分的差異，但是此處並沒有足夠資訊，可判斷這種較高值的非線性趨勢是真正的趨勢或是異常狀況。

# 閱讀資料推薦

請參閱下列文字以取得更多資訊。

Busing, F. M. T. A., P. J. F. Groenen, 和 W. J. Heiser. 2005. Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation (透過變異係數懲罰避免多維度展開簡併). Psychometrika (心理學計量報導), 70, .

Green, P. E., 和 V. Rao. 1972. Applied multidimensional scaling (應用多元尺度方法). Hinsdale, Ill.: Dryden Press (Dryden 出版社).

Price, R. H., 和 D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior (行為適切性與情境限制作為社會行為維度). Journal of Personality and Social Psychology (人格與社會心理學雜誌), 30, .

附錄
A

# 範例檔案

與產品同時安裝的範例檔存放在安裝目錄的範例子目錄中。在下列每種語言的「範例」子目錄中存有個別資料夾：英文、法文、德文、義大利文、日文、韓文、波蘭文、俄文、簡體中文、西班牙文和繁體中文。

並非所有範例檔案皆提供各種語言。如果範例檔案沒提供您需要的語言，語言資料夾有英文版的範例檔案。

## 說明

以下是使用於本文件中不同範例的範例檔之簡要描述。

- **accidents.sav。** 這是有關某保險公司研究年齡和性別風險因子對給定地區汽車意外事件的假設資料檔。每一個觀察值對應至一個年齡類別和性別的交叉分類。
- **adl.sav。** 這是有關致力於確定一個建議中風病患治療類型之效益的假設資料檔。醫師隨機指定女性中風病患至兩個組別之一。第一組接受標準的物理治療，而第二組則接受額外的情緒治療。在治療了三個月後，將每一個病患進行日常活動的能力記分為次序變數。
- **advert.sav。** 這是有關一家零售商致力於調查廣告費與廣告後銷售情形之間的關係的假設資料檔。為了這個目的，他們收集了過往銷售數字和相關的廣告費用。
- **aflatoxin.sav。** 這是有關檢定玉米作物是否有黃麴毒素（一種毒物，其濃度在介於和處於作物產量中都有很大的差異）的假設資料檔。一名穀物加工者收到來自 8 個作物產量各 16 個樣本，並以十億當量 (PPB) 來測量黃麴毒素的水準。
- **aflatoxin20.sav。** 這個資料檔包含由 aflatoxin.sav 取得，來自 4 和 8 作物產量的 16 個樣本，每一個樣本的黃麴毒素測量。
- **anorectic.sav。** 在將厭食/暴食行為症狀學標準化的過程中，研究人員 (Van der Ham, Meulman, Van Strien, 和 Van Engeland, 1997) 研究了 55 個飲食失調的青少年。每個病患在四年之中被訪問四個回合，所以得到總數為 220 的觀察值。在每次觀察中，為病患在 16 種症狀上逐一評分。目前遺漏了第二次訪察的病患 71，第二次訪察的病患 76，以及第三次訪察的病患 47 的症狀分數，因此只剩下 217 個有效觀察值。
- **autoaccidents.sav。** 這是有關一位保險分析師致力於為每個駕駛的汽車意外事件次數建立模式，同時考量駕駛的年齡和性別的假設資料檔。每一個觀察值代表一位不同的駕駛，記錄了駕駛的性別、年齡、和近五年內的汽車意外事故次數。
- **band.sav。** 本資料檔包含某樂團音樂 CD 假設性的每週銷售數字。也包含三個可能預測變數的資料。
- **bankloan.sav。** 這是有關一家銀行致力於減少放款利率預設值的假設資料檔。本檔包含 850 位以前的客戶與現在的準客戶的財務和人口資料。前 700 個觀察值為以前有借貸的客戶。最後 150 個觀察值是銀行需要作信用風險優良與不良分類的準客戶。

© Copyright SPSS Inc. 1989, 2010

- **bankloan_binning.sav。** 這是包含 500 位以前客戶的財務和人口資料的假設資料檔。
- **behavior.sav。**在典型範例 (Price 和 Bouffard, 1974) 中，52 名學生被要求為 15 種情境與 15 種行為組合評等，等級共分為 10 點，從 0 =「非常適當」到 9 =「 非常不適當」。平均值超過個別值，值會被視為相異性。
- **behavior_ini.sav。** 本資料檔包含 behavior.sav 之二維解的起始組態。
- **brakes.sav。** 這是有關一間生產高性能汽車碟型煞車片工廠中品質管制的假設資料檔。資料檔包含由 8 個生產機器分別取得 16 個碟片的直徑測量。煞車的目標直徑是 322 公釐。
- **breakfast.sav。**在經典研究中(Green 和 Rao, 1972)，21 名 Wharton 學院 MBA 學生及其配偶被要求為 15 項早餐食品按喜愛程度分出等級：從 1 = 「最喜愛」到 15 = 「最不喜愛」。他們的喜愛程度分六種不同情況記錄，從「整體喜愛」到「點心，僅配飲料」。
- **breakfast-overall.sav。** 本資料檔只包含第一種情況－「整體喜愛」－所喜愛的早餐項目。
- **broadband_1.sav。** 這是包含全國性寬頻服務地區用戶數目的假設資料檔。本資料檔包含四年期間 85 個地區每月的用戶數目。
- **broadband_2.sav。** 本資料檔與 broadband_1.sav 相同，但多了三個月的資料。
- **car_insurance_claims.sav。** 一個在別處 (McCullagh 和 Nelder, 1989) 出現和分析過，有關汽車損害理賠的資料集。理賠金額的平均數可建立模式為具有 gamma 分配，使用反連結函數將依變數的平均數相關至一被保險人年齡、車輛類型、和車齡的線性組合。提出理賠的數量可以用作尺度權重。
- **car_sales.sav。** 本資料檔包含假設性的銷售估計、定價、和不同的品牌與車輛型式的實體規格。定價和實體規格是由 edmunds.com 和製造商處輪流取得。
- **car_sales_uprepared.sav。** 這是 car_sales.sav 的修改版本，其中不包含任何欄位的轉換版本。
- **carpet.sav。**在一個普遍的範例 (Green 和 Wind, 1973) 中，計劃銷售全新地毯清潔機的公司想要檢驗影響消費者偏好的五個因子—包裝設計、品牌名稱、價格、「優秀家用品」獎章及退費保證。包裝設計有三個因子水準，每個水準中的清潔刷位置都不相同；三個品牌名稱 (K2R、Glory、及 Bissell)；三個價格水準；且最後兩個因子各有兩個水準 (無論無或有)。十名消費者將這些因子所定義的 22 種組合分級。「偏好」變數包含每個組合平均排名的等級。排名數值較小者會對應高偏好程度。這個變數反映每個組合偏好的整體量數。
- **carpet_prefs.sav。** 本資料檔是根據 carpet.sav 所描述的相同範例，但它包含 10 個消費者每一個人的實際等級。消費者被要求將 22 個產品組合從最喜歡排列到最不喜歡。變數「PREF1」到「PREF22」包含相關組合的識別碼，如 carpet_plan.sav 中所定義。
- **catalog.sav。** 本資料檔包含郵購公司銷售三項產品的每月假設銷售數字。也包含五個可能預測變數的資料。
- **catalog_seasfac.sav。** 本資料檔與 catalog.sav 相同，不過多了一組由「週期性分解」程序所計算的週期性因子以及隨附的資料變數。
- **cellular.sav。** 這是有關一家手機公司致力於減少顧客不忠的假設資料檔。顧客不忠傾向分數套用於帳戶，範圍由 0 至 100。帳戶分數 50 或以上有可能正尋求變更供應商。

- **ceramics.sav。** 這是有關一家製造商致力於確定一種新的優良合金是否較標準的合金有較大的耐熱性的假設資料檔。每一個觀察值代表對合金之一的不同檢定；記錄了讓軸承失效的溫度。
- **cereal.sav。** 這是有關對 880 人的早餐喜好進行訪談的假設資料檔，也記下他們的年齡、性別、婚姻狀況、和是否有活躍的生活型態 (根據他們是否一週運動兩次)。每一個觀察值代表一位不同的應答者。
- **clothing_defects.sav。** 這是有關一家服裝工廠品質管制過程的假設資料檔。由該工廠所生產的每一批產品中，檢查員取出一件服裝的樣本並計算不合格的服裝個數。
- **coffee.sav。** 本資料檔是關於六種冰咖啡品牌的感覺印象(Kennedy, Riquier, 和 Sharp, 1996)。對 23 種冰咖啡中每一種的印象屬性，由群眾來選取依其屬性描述的所有品牌。該六種品牌已標示為 AA、BB、CC、DD、EE、和 FF，以保持機密。
- **contacts.sav。** 這是有關一群公司電腦銷售代表聯絡清單的假設資料檔。每一個聯絡人依他們在公司所服務的部門及其公司的等級而分類。最後一次銷售的金額、到最後一次銷售的時間、和該聯絡人公司的規模也都被列入記錄。
- **creditpromo.sav。** 這是有關一家百貨公司致力於評估近期信用卡促銷活動效果的假設資料檔。為達此目標，隨機選取了 500 位持卡人。有半數收到廣告，促銷在未來三個月購買將獲得降低利率的優惠。半數收到標準的週期性廣告。
- **customer_dbase.sav。** 這是有關一家公司致力於使用其資料倉庫的資訊來對最有可能回應的客戶提供優惠的假設資料檔。隨機選取客戶庫的子集，提供優惠，再將他們的回應記錄下來。
- **customer_information.sav。** 本檔案是包含客戶郵寄資訊的假設資料檔，例如姓名和地址。
- **customer_subset.sav。** 80 個 customer_dbase.sav 的觀察值子集。
- **customers_model.sav。** 本檔案包含一市場行銷活動所鎖定之個人的假設資料。這些資料包含人口資訊、購買歷史摘要、和每一個人是否對該活動有回應。每一個觀察值代表一位不同的個人。
- **customers_new.sav。** 本檔案包含一市場行銷活動潛在候選人之個人的假設資料。這些資料包含每一位個人的人口資訊和購買歷史摘要。每一個觀察值代表一位不同的個人。
- **debate.sav。** 這是有關一項政治辯論會參與者辯論前和辯論後接受調查之成對反應的假設資料檔。每一個觀察值對應至一位不同的應答者。
- **debate_aggregate.sav。** 這是將 debate.sav 中之反應作整合的假設資料檔。每一個觀察值對應至辯論前和辯論後對偏好之交叉分類的反應。
- **demo.sav。** 這是有關提供郵寄每月優惠之購買客戶資料庫的假設資料檔。記錄了客戶是否對該優惠回應，以及各種的人口資訊。
- **demo_cs_1.sav。** 這是有關一家公司致力於匯編調查資訊資料庫之第一步的假設資料檔。每一個觀察值對應至一個不同的城市，也記錄了其地區、省、區、和城市識別。
- **demo_cs_2.sav。** 這是有關一家公司致力於匯編調查資訊資料庫之第二步的假設資料檔。每一個觀察值對應至在第一步中選取的城市中的一個不同的家庭單位，也記錄了其地區、省、區、分區、和單位識別。也納入了由該設計的前兩階段所得之取樣資訊。
- **demo_cs.sav。** 這是包含以複合取樣設計所收集之調查資訊的假設資料檔。每一個觀察值對應至一個不同的家庭單位，也記錄了各種的人口和取樣資訊。

- **dmdata.sav。** 這是包含直效行銷公司之人口和購買資訊的假設資料檔。dmdata2.sav 包含收到測試郵件的連絡人子集資訊，而 dmdata3.sav 則包含剩下未收到測試郵件的連絡人資訊。
- **dietstudy.sav。** 本假設資料檔包含對「Stillman 飲食法」(Rickman, Mitchell, Dingman, 和 Dalen, 1974) 研究的結果。每一個觀察值對應至一個不同的受試者，並記錄下他或她飲食法前、後之體重 (磅) 和三酸甘油酯水準 (毫克/100 毫升)。
- **dvdplayer.sav。** 這是有關新 DVD 播放器開發的假設資料檔。市場行銷團隊使用原型收集了焦點組別資料。每一個觀察值對應至不同調查到的使用者，並記錄下一些有關他們的人口資訊和他們對有關原型問題的回應。
- **german_credit.sav。** 本資料檔取自 (Blake 和 Merz, 1998) 艾文 (Irvine) 在加州大學機器學習資料庫儲存器的「德國信用」資料集。
- **grocery_1month.sav。** 本假設資料檔是將 grocery_coupons.sav 資料檔和每週購買的「彙總」，因此每一個觀察值對應至一個不同的客戶。結果部份每週變更的變數消失了，而目前所記錄的銷售量是在研究的四週期間銷售量之總和。
- **grocery_coupons.sav。** 這是包含某連鎖雜貨店想要知道他們客戶購買習慣所收集之調查資料的假設資料檔。每一個客戶被追蹤了四週，每一個觀察值對應至一個不同的客戶-週，並記錄有關客戶在何處及如何購物的資訊，包含那一週在雜貨店花了多少錢。
- **guttman.sav。**Bell(Bell, 1961) 以此表說明可能的社會團體。Guttman (Guttman 值, 1968) 過去曾使用此表的一部分，在這部分中有 5 個變數，分別說明 7 個理論社會團體的社會互動、團體歸屬感、成員實際接觸和關係正式性，而這 7 個群組包括：群眾 (例如，足球場上的人)、觀眾 (例如在戲院中和課堂上的人)、公眾 (例如，報紙讀者和電視觀眾)、暴民 (和群眾相似，但互動較為激烈)、原級團體 (親密性)、次級團體 (自願性) 和現代社群 (因親密的身體接近而導致鬆散的結盟和特殊服務的需求)。
- **health_funding.sav。** 這是包含醫療保健基金 (每 100 個人口的金額)、疾病率 (每 10,000 個人口的比率)、造訪醫療保健機構的比例 (每 10,000 個人口的比率) 的假設資料檔。每一個觀察值代表一個不同的城市。
- **hivassay.sav。** 這是有關一家製藥實驗室致力於開發一種偵測 HIV 感染快速檢驗的假設資料檔。檢驗結果是八個紅色加深的陰影，陰影愈深表示感染的可能性愈大。進行了一項實驗室的試驗，在 2,000 個血液樣本中，有半數遭到 HIV 的感染，而半數則未感染。
- **hourlywagedata.sav。** 這是有關在辦公室和醫院任職的護士依經驗水準不同之鐘點費的假設資料檔。
- **insurance_claims.sav.** 這是有關一家保險公司想要建立模式來標示可疑及可能的詐欺理賠之假設資料檔。每一個觀察值代表個不同的理賠。
- **insure.sav。** 這是有關一家保險公司正在研究表示客戶是否必定理賠 10 年壽險合約之風險因子的假設資料檔。在資料檔中的每一個觀察值代表二份合約，其一記錄了理賠而另一則否，二者的年齡和性別相符。
- **judges.sav。** 這是有關受過訓練的裁判 (加上一位熱心人士) 為 300 個體操表演評分的假設資料檔。每一列代表一個不同的表演；裁判們觀看相同的表演。
- **kinship_dat.sav。**Rosenberg 與 Kim (Rosenberg 和 Kim, 1975) 致力於分析 15 個親屬關係稱呼 (姑/姨、兄弟、堂/表兄弟姐妹、女兒、父親、孫女、祖父、祖母、孫子、母親、姪子/外甥、姪女/外甥女、姐妹、兒子、叔/舅父)。他們請四組大學生 (兩組女性、兩組男性) 根據其相似性來分類整理這些稱謂。他們請其中兩組 (一組

女性、一組男性）作兩次分類整理，第二次要根據與第一次不同的準則進行分類整理。因此，總共得到六個「來源」。每一個來源對應至一個 15 × 15 的相似性矩陣，其儲存格等於來源中人數減去物件在該來源中分為同組的次數。

- **kinship_ini.sav。** 本資料檔包含 kinship_dat.sav 之三維解的起始組態。
- **kinship_var.sav。** 本資料檔包含自變數「性別」、「世代」、和可用來解讀 kinship_dat.sav 解答維度的（分離）「度」。尤其，它們可用來將解答空間限制為這些變數的線性組合。
- **marketvalues.sav。** 本資料檔有關於一項在伊立諾州阿爾岡京 (Algonquin, Ill.) 的新屋開發案自 1999 年至 2000 年之房屋銷售情況。這些銷售與公共記錄有關。
- **nhis2000_subset.sav。** 「國民健康訪問調查 (NHIS)」為美國民間人口的一大型民眾調查。其以具全國代表性的家庭為樣本，面對面的完成訪問。而取得各家庭中成員的人口統計學資訊及健康行為、健康狀態方面等觀察報告。本資料檔包含一個 2000 年調查資訊的子集。國家衛生統計中心。2000 年「國民健康訪問調查 (NHIS)」。公用資料檔案和文件。ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/。2003 年曾存取。
- **ozone.sav。** 本資料包含對六個氣象變數所作的 330 個觀察值，以自其餘的變數中預測臭氧濃度。先前研究人員中，(Breiman 和 Friedman 檢定(F), 1985)、(Hastie 和 Tibshirani, 1990) 在這些會阻礙標準迴歸方式的變數中發現非線性。
- **pain_medication.sav。** 本假設資料檔包含治療慢性關節炎疼痛之消炎藥物臨床試驗的結果。特別關注於藥物發生作用的時間以及它是如何與現用藥物作比較。
- **patient_los.sav。** 本假設資料檔包含對因可能為心肌梗塞 (MI，或「心臟病」）入院病患的治療記錄。每一個觀察值對應至一個不同的病患並記錄許多與其留院期間有關的變數。
- **patlos_sample.sav。** 本假設資料檔包含病患在為心肌梗塞 (MI，或「心臟病」）治療期間接受血栓溶解治療的治療記錄樣本。每一個觀察值對應至一個不同的病患並記錄許多與其留院期間有關的變數。
- **polishing.sav。** 這是取自「資料和故事圖書館」的「Nambeware 打磨時間」資料檔。它是有關一家金屬餐具製造商 (Nambe Mills, 聖塔非, 新墨西哥州）致力於規劃其生產排程。每一個觀察值代表生產線上一個不同的產品。每一個產品都記錄下直徑、打磨時間、價格、和產品類別。
- **poll_cs.sav。** 這是有關民意測驗專家致力於確定交付立法之前公眾對法案支持水準的假設資料檔。觀察值對應至登記選民。每一個觀察值記錄下選民的郡、鎮、和他居住的鄰近範圍。
- **poll_cs_sample.sav。** 本假設資料檔包含列於 poll_cs.sav 中的選民樣本。樣本是根據在 poll.csplan 計劃檔中指定的設計來取得，而本資料檔記錄了包含機率和樣本權重。不過，請注意，由於取樣計劃採用到機率 - 比例 - 大小 (PPS) 方法，也用到一個包含聯合選擇機率的檔案 (poll_jointprob.sav)。其他與選民人口及其對提議法案之意見有關的變數都在取樣後收集並加入資料檔中。
- **property_assess.sav。** 這是有關郡財產估價人員致力於對限定資源保持財產價值評估維持最新的假設資料檔。觀察值對應至郡內過去一年銷售的財產。資料檔中的每一個觀察值記錄了財產所在的鎮、上次訪查該財產的估價人員、自那次評估後經過的時間、當時定的估價、和該財產銷售價值。

- **property_assess_cs.sav。** 這是有關州財產估價人員致力於對限定資源保持財產價值評估維持最新的假設資料檔。觀察值對應至州中的財產。資料檔中的每一個觀察值記錄了郡、鎮、和財產所在的鄰近範圍、自最後一次評估後經過的時間、和當時定的估價。
- **property_assess_cs_sample.sav。** 本假設資料檔包含列於 property_assess_cs.sav 中的財產樣本。樣本是根據在 property_assess.csplan 計劃檔中指定的設計來取得，而本資料檔記錄了包含機率和樣本權重。另外的變數「目前價值」是在取樣後收集並加入資料檔中。
- **recidivism.sav。** 這是有關政府法令執行機構致力於瞭解其轄區內之再犯率的假設資料檔。每一個觀察值對應至一個先前的違法者並記錄其人口資訊、第一次犯罪的一些細節、然後是直到第二次被捕的時間（如果它發生在第一次被捕的兩年之內）。
- **recidivism_cs_sample.sav。** 這是有關政府法令執行機構致力於瞭解其轄區內之再犯率的假設資料檔。每一個觀察值對應到一個先前的違法者，在 2003 年六月第一次被捕後釋放，並記錄其人口資訊、第一次犯罪的一些細節、和第二次被捕日期（如果它發生在 2006 年六月之前）。違法者是根據在 recidivism_cs.csplan 中所指定的取樣計劃之樣本部門來選取；由於取樣計劃採用到機率 - 比例 - 大小 (PPS) 方法，也用到一個包含聯合選擇機率的檔案 (recidivism_cs_jointprob.sav)。
- **rfm_transactions.sav。** 本檔案是包含購買交易資料的假設資料檔，包括購買日期、購買項目及每一項交易的金額。
- **salesperformance.sav。** 這是有關評估兩個新售貨員訓練課程的假設資料檔。六十個員工，分成三個組別，全部接受標準訓練。此外，組別二得到技術訓練；組別三則是實務輔導簡介。每一個員工在訓練課程結束時接受測驗並記錄他們的分數。在資料檔中每一個觀察值代表一個不同的訓員，並記錄他們所分派的組別和他們在測驗中得到的分數。
- **satisf.sav。** 這是有關一家零售公司在 4 個商店位置所作之滿意度調查的假設資料檔。總共有 582 位客戶接受調查，每一個觀察值代表一位客戶的反應。
- **screws.sav。** 這個資料檔包含螺絲釘、螺栓、螺帽和圖釘之特色的資訊(Hartigan, 1975)。
- **shampoo_ph.sav。** 這是有關一家美髮產品工廠品質管制過程的假設資料檔。在固定的時間間隔，記錄下六個不同輸出批次的測量和它們的 pH 值。目標範圍是 4.5 - 5.5。
- **ships.sav。** 一個在別處 (McCullagh et al。, 1989) 出現和分析過，有關商船因風浪所造成損壞的資料集。事件次數可建立模式為以 Poisson 率發生，給定船型、建造期間、和服務期間。以因子交叉分類所形成的表格的每一個儲存格服務月數的整合，提供了暴露於風險之值。
- **site.sav。** 這是有關一家公司致力於為事業擴展選擇新地點的假設資料檔。您們僱請兩位顧問分別評估該地點，除了一份廣泛的報告之外，他們還要將每個地點摘要為前景「佳」、「可」、或「差」。
- **smokers.sav。** 本資料檔是由「1998 年全國家庭毒品濫用調查」中摘錄，且是美國家庭的機率樣本。(http://dx.doi.org/10.3886/ICPSR02934) 因此，在分析本資料檔的第一步應該是將資料加權以反映母群體傾向。
- **stroke_clean.sav。** 本假設資料檔包含一個醫療資料庫，其在以「資料準備」選項中的程序清理之後的狀態。
- **stroke_invalid.sav。** 本假設資料檔包含一個醫療資料庫的起始狀態並包含幾個資料輸入錯誤。

- **stroke_survival。** 本假設資料檔是有關缺血性中風的病患，其在結束康復計畫後存活時間方面，面臨許多挑戰。中風後，記載了心肌梗塞、缺血性中風、或出血性中風的發生，以及事件記錄的時間。由於它只包含在康復計劃所管制的中風存活的病患，此樣本的左側被截斷。
- **stroke_valid.sav。** 本假設資料檔包含一個醫療資料庫，在其值以「驗證資料」程序檢查之後的狀態。它仍包含可能的異常觀察值。
- **survey_sample.sav。** 本資料檔包含調查資料，包括人口資料和各種態度測量。雖然已修改一些資料數值，且為人口資料之目的新增了一些額外的虛構變數，但是資料仍是以「1998 NORC 基本社會調查」的變數子集為基礎。
- **telco.sav。** 這是有關一家電信公司致力於在客戶庫中減少顧客不忠的假設資料檔。每一個觀察值對應至一位不同的客戶並記錄不同的人口資料和服務使用方式資訊。
- **telco_extra.sav。** 本資料檔類似於 telco.sav 資料檔，但「任期」的對數轉換客戶花費變數已予刪除，並更換為標準的對數轉換客戶花費變數。
- **telco_missing.sav。** 本資料檔是 telco.sav 資料檔的子集，不過某些人口資料值已更換為遺漏值。
- **testmarket.sav。** 本假設資料檔有關於一家速食連鎖店計劃在菜單中加入新的項目。有三個可能的活動來促銷此新產品，所以該新項目在幾個隨機選取市場中的地點作介紹。在每一個地點使用不同的促銷，並記錄該新項目前四週的每週銷售量。每一個觀察值對應至一個不同的地點-週。
- **testmarket_1month.sav。** 本假設資料檔是將 testmarket.sav 資料檔和每週購買的「彙總」，因此每一個觀察值對應至一個不同的客戶。結果部份每週變更的變數消失了，而目前所記錄的銷售量是在研究的四週期間銷售量之總和。
- **tree_car.sav。** 這是包含人口資料和車輛購買價格資料的假設資料檔。
- **tree_credit.sav。** 這是包含人口資料和銀行放款歷史資料的假設資料檔。
- **tree_missing_data.sav** 這是包含有大量遺漏值的人口資料和銀行放款歷史資料的假設資料檔。
- **tree_score_car.sav。** 這是包含人口資料和車輛購買價格資料的假設資料檔。
- **tree_textdata.sav。** 一個只有兩個變數的簡單資料檔，主要目的在顯示變數預設狀態（在指定量測水準和數值標記之前）。
- **tv-survey.sav。** 這是有關一家電視製片廠考量是否要延長一個成功節目的播送所作之調查的假設資料檔。有 906 位應答者被問到在不同的狀況下他們是否願意觀看這個節目。每一列代表一個不同的應答者；每一行為一個不同的狀況。
- **ulcer_recurrence.sav。** 本檔案包含一項用來比較兩種防止潰瘍復發治療法功效之研究的部分資訊。它是很好的區間受限資料範例，且已在別處 (Collett, 2003) 出現和分析過。
- **ulcer_recurrence_recoded.sav。** 本檔案是將 ulcer_recurrence.sav 的資訊重新組織，以讓您為此研究的每一個區間事件機率而非只是研究目的事件機率建立模式。它已在別處 (Collett et al。, 2003) 出現和分析過。
- **verd1985.sav。** 本資料檔有關於一項調查 (Verdegaal, 1985)。在調查中記錄了來自 15 個受訪者對 8 個變數的回應。所需的變數被分成三組。集 1 包括 age 和 marital，集 2 包括 pet 和 news，集 3 包括 music 和 live。Pet 調整為多重名義量數，age 調整為次序量數，其他的變數調整為單一名義量數。

- **virus.sav。** 這是有關一家網際網路服務提供者致力於在其網路上判斷病毒之影響的假設資料檔。他們在其網路上追蹤從發現病毒直到控制威脅的這段時間，被病毒感染之電子郵件的流量（約略）百分比。
- **wheeze_steubenville.sav。** 這是空氣污染對兒童健康之影響 (Ware, Dockery, Spiro III, Speizer, 和 Ferris Jr., 1984) 縱向研究的子集。本資料包含來自俄亥俄州 Steubenville，年齡 7、8、9 和 10 歲兒童的氣喘聲狀態之重複二元測量，以及其母親在本研究的第一年是否抽煙的固定記錄。
- **workprog.sav。** 這是有關一項政府職業計劃，設法將弱勢民眾安置到較好之工作的假設資料檔。一個樣本的可能計劃參與者被追蹤，他們之中某些被選取加入本計劃，而其他的則否。每一個觀察值代表一位不同的計劃參與者。

附錄
B

# Notices

Licensed Materials - Property of SPSS Inc., an IBM Company. © Copyright SPSS Inc. 1989, 2010.

Patent No. 7,023,453

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** SPSS INC., AN IBM COMPANY, PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. SPSS Inc. may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-SPSS and non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this SPSS Inc. product and use of those Web sites is at your own risk.

When you send information to IBM or SPSS, you grant IBM and SPSS a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Information concerning non-SPSS products was obtained from the suppliers of those products, their published announcements or other publicly available sources. SPSS has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-SPSS products. Questions on the capabilities of non-SPSS products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to SPSS Inc., for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. SPSS Inc., therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided “AS IS”, without warranty of any kind. SPSS Inc. shall not be liable for any damages arising out of your use of the sample programs.

**Trademarks**

IBM, the IBM logo, and ibm.com are trademarks of IBM Corporation, registered in many jurisdictions worldwide. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shmtl.

SPSS is a trademark of SPSS Inc., an IBM Company, registered in many jurisdictions worldwide.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

This product uses WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, http://www.winwrap.com.

Other product and service names might be trademarks of IBM, SPSS, or other companies.

Adobe product screenshot(s) reprinted with permission from Adobe Systems Incorporated.

Microsoft product screenshot(s) reprinted with permission from Microsoft Corporation.

# 參考書目


Barlow, R. E., D. J. Bartholomew, D. J. Bremner, 和 H. D. Brunk. 1972. Statistical inference under order restrictions (次序限制下之統計推論). 紐約: John Wiley and Sons.

Bell, E. H. 1961. Social foundations of human behavior: (人類行為之社會基礎)Introduction to the study of sociology (社會學研究簡介). 紐約: Harper & Row.

Benzécri, J. P. 1969. Statistical analysis as a tool to make patterns emerge from data (透過統計分析從資料中找出樣式). 於: Methodologies of Pattern Recognition (樣式辨識方法論), S. Watanabe，ed. 紐約: Academic Press (學術出版社).

Benzécri, J. P. 1992. Correspondence analysis handbook (對應分析手冊). 紐約: Marcel Dekker.

Bishop, Y. M., S. E. Feinberg, 和 P. W. Holland. 1975. Discrete multivariate analysis: (離散多變量分析：)Theory and Practice (理論與實踐). Cambridge, Mass. (麻州劍橋): MIT Press (MIT 出版部).

Blake, C. L., 和 C. J. Merz. 1998. "UCI Repository of machine learning databases (機器學習資料庫 UCI 儲存器)." Available at http://www.ics.uci.edu/~mlearn/MLRepository.html.

Breiman, L., 和 J. H. Friedman 檢定(F). 1985. Estimating optimal transformations for multiple regression and correlation (估計多重迴歸與相關之最適轉換). Journal of the American Statistical Association (美國統計協會彙報), 80, .

Buja, A. 1990. Remarks on functional canonical variates, alternating least squares methods and ACE (最小平方法與 ACE 交替使用下函數典型變異之淺見). Annals of Statistics (統計學年鑑), 18, .

Busing, F. M. T. A., P. J. F. Groenen, 和 W. J. Heiser. 2005. Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation (透過變異係數懲罰避免多維度展開簡併). Psychometrika (心理學計量報導), 70, .

Carroll, J. D. 1968. Generalization of canonical correlation analysis to three or more sets of variables (三個或以上變數集合之典型相關分析泛論). 於: Proceedings of the 76th Annual Convention of the American Psychological Association, 3 (第 76 屆美國心理學協會大會記錄，第 3 冊), 華盛頓特區: American Psychological Association (美國心理學協會).

Collett, D. 2003. Modelling survival data in medical research (模式化醫學研究中的存活資料), 2 ed. Boca Raton: Chapman & Hall/CRC.

Commandeur, J. J. F., 和 W. J. Heiser. 1993. Mathematical derivations in the proximity scaling (PROXSCAL) of symmetric data matrices (對稱資料矩陣相似性尺度(PROXSCAL) 的數學應用). Leiden (萊頓): Department of Data Theory, University of Leiden (萊頓大學資料理論系).

De Haas, M., J. A. Algera, H. F. J. M. Van Tuijl, 和 J. J. Meulman. 2000. Macro and micro goal setting: (微觀及宏觀目標設定)In search of coherence (尋求一貫性). Applied Psychology (應用心理學), 49, .


© Copyright SPSS Inc. 1989, 2010

De Leeuw, J. 1982. Nonlinear principal components analysis (非線性主成份分析). 於： COMPSTAT Proceedings in Computational Statistics (計算統計學 COMPSTAT 會議記錄), Vienna (維也納): Physica Verlag.

De Leeuw, J. 1984. Canonical analysis of categorical data (類別資料之典型分析), 第二版 ed. Leiden (萊頓): DSWO Press (DSWO 出版社).

De Leeuw, J. 1984. The Gifi system of nonlinear multivariate analysis (非線性多變量分析 Gifi 系統). 於： Data Analysis and Informatics III (資料分析與資訊學 III), E. Diday, et al. (Diday 與其他)，ed..

De Leeuw, J., 和 W. J. Heiser. 1980. Multidimensional scaling with restrictions on the configuration (限制組態下多元尺度方法). 於： Multivariate Analysis, Vol. V (多變量分析第五冊), P. R. Krishnaiah，ed. Amsterdam (阿姆斯特丹): North-Holland (北荷蘭).

De Leeuw, J., 和 J. Van Rijckevorsel. 1980. HOMALS and PRINCALS—Some generalizations of principal components analysis (HOMALS 與 PRINCALS — 主成份分析泛論). 於： Data Analysis and Informatics (資料分析與資訊學), E. Diday, et al. (Diday 與其他)，ed. Amsterdam (阿姆斯特丹): North-Holland (北荷蘭).

De Leeuw, J., F. W. Young, 和 Y. Takane. 1976. Additive structure in qualitative data: (量化資料之附加結構)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 41, .

De Leeuw, J. 1990. Multivariate analysis with optimal scaling (最適尺度多變量分析). 於： Progress in Multivariate Analysis (多變量分析進度), S. Das Gupta, 和 J. Sethuraman，eds. Calcutta (加爾各答): Indian Statistical Institute (印度統計中心).

Eckart, C., 和 G. Young. 1936. The approximation of one matrix by another one of lower rank (一個矩陣及另一個低等級矩陣的近似法). Psychometrika (心理學計量報導), 1, .

Fisher, R. A. 1938. Statistical methods for research workers (研究人員的統計方法). Edinburgh (愛丁堡): Oliver and Boyd.

Fisher, R. A. 1940. The precision of discriminant functions (判別函數之精確度). Annals of Eugenics (優生學年鑑), 10, .

Gabriel 檢定(G), K. R. 1971. The biplot graphic display of matrices with application to principal components analysis (矩陣行列散佈圖形顯示與其在主成份分析之應用). Biometrika (生物統計學雜誌), 58, .

Gifi, A. 1985. PRINCALS。Research Report UG-85-02 (研究報告 UG-85-02). Leiden (萊頓): Department of Data Theory, University of Leiden (萊頓大學資料理論系).

Gifi, A. 1990. Nonlinear multivariate analysis (非線性多變量分析). Chichester: John Wiley and Sons.

Gilula, Z., 和 S. J. Haberman. 1988. The analysis of multivariate contingency tables by restricted canonical and restricted association models (有限典型與有限關聯模式之多變量列聯表分析). Journal of the American Statistical Association (美國統計協會彙報), 83, .

Gower, J. C., 和 J. J. Meulman. 1993. The treatment of categorical information in physical anthropology (體質人類學之類別資訊處理). International Journal of Anthropology (國際人類學刊), 8, .

Green, P. E., 和 V. Rao. 1972. Applied multidimensional scaling (應用多元尺度方法). Hinsdale, Ill.: Dryden Press (Dryden 出版社).

Green, P. E., 和 Y. Wind. 1973. Multiattribute decisions in marketing: (市場行銷之多重屬性決策)A measurement approach (測量方法). Hinsdale, Ill.: Dryden Press (Dryden 出版社).

Guttman 值, L. 1941. The quantification of a class of attributes: (屬性等級量化)A theory and method of scale construction (尺度建構之理論與方法). 於： The Prediction of Personal Adjustment (個人調整預測), P. Horst，ed. 紐約: Social Science Research Council (社會學研究協會).

Guttman 值, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points (尋找組態點最小座標空間之一般非計量技巧). Psychometrika (心理學計量報導), 33, .

Hartigan, J. A. 1975. Clustering algorithms (集群演算法). 紐約: John Wiley and Sons.

Hastie, T., 和 R. Tibshirani. 1990. Generalized additive models (概化附加模式). 倫敦: Chapman and Hall.

Hastie, T., R. Tibshirani, 和 A. Buja. 1994. Flexible discriminant analysis (彈性判別分析). Journal of the American Statistical Association (美國統計協會彙報), 89, .

Hayashi, C. 1952. On the prediction of phenomena from qualitative data and the quantification of qualitative data from the mathematico-statistical point of view (從數學統計觀點論性質資料與性質資料量化之現象預測). Annals of the Institute of Statitical Mathematics (統計數學所年鑑), 2, .

Heiser, W. J. 1981. Unfolding analysis of proximity data (相似性資料的展開分析). Leiden (萊頓): Department of Data Theory, University of Leiden (萊頓大學資料理論系).

Heiser, W. J., 和 F. M. T. A. Busing. 2004. Multidimensional scaling and unfolding of symmetric and asymmetric proximity relations (對稱與非對稱相似關係之多元尺度與展開方法). 於： Handbook of Quantitative Methodology for the Social Sciences (社會學量化方法論手冊), D. Kaplan，ed. Thousand Oaks, Calif. (加州千橡): Sage Publications, Inc..

Heiser, W. J., 和 J. J. Meulman. 1994. Homogeneity analysis: (均齊性分析)Exploring the distribution of variables and their nonlinear relationships (探索變數分配及其非線性關係). 於： Correspondence Analysis in the Social Sciences: (社會學對應分析)Recent Developments and Applications (近期發展與應用), M. Greenacre, 和 J. Blasius，eds. 紐約: Academic Press (學術出版社).

Heiser, W. J., 和 J. J. Meulman. 1995. Nonlinear methods for the analysis of homogeneity and heterogeneity (均齊性與異質性非線性分析). 於： Recent Advances in Descriptive Multivariate Analysis (敘述多變量分析最新發展), W. J. Krzanowski，ed. Oxford (牛津): Oxford University Press (牛津大學出版部).

Horst, P. 1961. Generalized canonical correlations and their applications to experimental data (概化典型相關及其實驗性資料應用). Journal of Clinical Psychology (臨床心理學雜誌), 17, .

Horst, P. 1961. Relations among m sets of measures (m 集量測之間關係 ). Psychometrika (心理學計量報導), 26, .

Israëls, A. 1987. Eigenvalue techniques for qualitative data (性質資料的特徵值技巧). Leiden (萊頓): DSWO Press (DSWO 出版社).

Kennedy, R., C. Riquier, 和 B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research (市場研究類別資料之對應分析實際應用). Journal of Targeting, Measurement, and Analysis for Marketing (市場行銷之目標訂定、測量與分析雜誌), 5, .

Kettenring, J. R. 1971. Canonical analysis of several sets of variables (多個變數集合的典型分析). Biometrika (生物統計學雜誌), 58, .

Kruskal, J. B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis (將適合度優化為非計量假設之多元尺度方法). Psychometrika (心理學計量報導), 29, .

Kruskal, J. B. 1964. Nonmetric multidimensional scaling: (非計量多元尺度方法)A numerical method (數值方法). Psychometrika (心理學計量報導), 29, .

Kruskal, J. B. 1965. Analysis of factorial experiments by estimating monotone transformations of the data (資料單調轉換評估之因子實驗分析). Journal of the Royal Statistical Society Series B (皇家統計協會彙報，系列 B), 27, .

Kruskal, J. B. 1978. Factor analysis and principal components analysis: (因子分析與主成份分析)Bilinear methods (雙線性方法). 於： International Encyclopedia of Statistics (國際統計學百科), W. H. Kruskal, 和 J. M. Tanur，eds. 紐約: The Free Press (自由出版社).

Kruskal, J. B., 和 R. N. Shepard. 1974. A nonmetric variety of linear factor analysis (線性因子分析之非計量變化). Psychometrika (心理學計量報導), 39, .

Krzanowski, W. J., 和 F. H. C. Marriott. 1994. Multivariate analysis: (多變量分析)Part I, distributions, ordination and inference (第一部，分配、順序與推論). 倫敦: Edward Arnold.

Lebart, L., A. Morineau, 和 K. M. Warwick. 1984. Multivariate descriptive statistical analysis (多變量敘述統計分析). 紐約: John Wiley and Sons.

Lingoes, J. C. 1968. The multivariate analysis of qualitative data (性質資料之多變量分析). Multivariate Behavioral Research (多變量行為研究), 3, .

McCullagh, P., 和 J. A. Nelder. 1989. 概化線性模式, 第二版 ed. 倫敦: Chapman & Hall.

Meulman, J. J. 1982. Homogeneity analysis of incomplete data (不完整資料之均齊性分析). Leiden (萊頓): DSWO Press (DSWO 出版社).

Meulman, J. J. 1986. A distance approach to nonlinear multivariate analysis (非線性多變量分析距離法). Leiden (萊頓): DSWO Press (DSWO 出版社).

Meulman, J. J. 1992. The integration of multidimensional scaling and multivariate analysis with optimal transformations of the variables (以變數最適轉換整合多元尺度方法及多變量分析). Psychometrika (心理學計量報導), 57, .

Meulman, J. J. 1993. Principal coordinates analysis with optimal transformations of the variables: (最適變數轉換之主座標分析)Minimizing the sum of squares of the smallest eigenvalues (將最小特徵值平方和最小化). British Journal of Mathematical and Statistical Psychology (英國數學暨統計心理學雜誌), 46, .

Meulman, J. J. 1996. Fitting a distance model to homogeneous subsets of variables: (在同質性變數子集上配適距離模式)Points of view analysis of categorical data (類別資料之觀點分析). Journal of Classification (分類雜誌), 13, .

Meulman, J. J. 2003. Prediction and classification in nonlinear data analysis: (非線性資料分析之預測與分類)Something old, something new, something borrowed, something blue (舊的、新的、借的、藍的). Psychometrika (心理學計量報導), 4, .

Meulman, J. J., 和 W. J. Heiser. 1997. Graphical display of interaction in multiway contingency tables by use of homogeneity analysis (使用同質性分析之多因子列聯表顯示互動圖形). 於： Visual Display of Categorical Data (類別資料視覺顯示), M. Greenacre, 和 J. Blasius，eds. 紐約: Academic Press (學術出版社).

Meulman, J. J., 和 P. Verboon. 1993. Points of view analysis revisited: (再論觀點分析)Fitting multidimensional structures to optimal distance components with cluster restrictions on the variables (以集群限制的變數將多元結構配適至最適距離成份). Psychometrika (心理學計量報導), 58, .

Meulman, J. J., A. J. Van der Kooij, 和 A. Babinec. 2000. New features of categorical principal components analysis for complicated data sets, including data mining (複雜資料集類別主成份分析之新功能 - 含資料採礦). 於： Classification, Automation and New Media (分類、自動化與新媒體), W. Gaul, 和 G. Ritter，eds. 柏林: Springer-Verlag.

Meulman, J. J., A. J. Van der Kooij, 和 W. J. Heiser. 2004. Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data (次序及非次序資料非線性最適尺度轉換之主成分分析). 於： Handbook of Quantitative Methodology for the Social Sciences (社會學量化方法論手冊), D. Kaplan，ed. Thousand Oaks, Calif. (加州千橡): Sage Publications, Inc..

Nishisato, S. 1980. Analysis of categorical data: (類別資料分析)Dual scaling and its applications (雙尺度分析及其應用). Toronto (多倫多): University of Toronto Press (多倫多大學出版部).

Nishisato, S. 1984. Forced classification: (強制分類)A simple application of a quantification method (量化法簡易應用). Psychometrika (心理學計量報導), 49, .

Nishisato, S. 1994. Elements of dual scaling: (雙尺度方法元素)An introduction to practical data analysis (實際資料分析簡介). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

Pratt, J. W. 1987. Dividing the indivisible: (個體分組)Using simple symmetry to partition variance explained (使用簡單對稱區隔已說明變異). 於： Proceedings of the Second International Conference in Statistics (第二屆統計學國際會議記錄), T. Pukkila, 和 S. Puntanen，eds. Tampere, Finland (芬蘭坦佩雷): University of Tampere (坦佩雷大學).

Price, R. H., 和 D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior (行為適切性與情境限制作為社會行為維度). Journal of Personality and Social Psychology (人格與社會心理學雜誌), 30, .

Ramsay, J. O. 1989. Monotone regression splines in action (單調迴歸曲線實作). Statistical Science (統計學), 4, .

Rao, C. R. 1973. Linear statistical inference and its applications (線性統計推論及其應用), 第二版 ed. 紐約: John Wiley and Sons.

Rao, C. R. 1980. Matrix approximations and reduction of dimensionality in multivariate statistical analysis (多變量統計分析之矩陣近似法與維度縮減). 於： Multivariate Analysis, Vol. V (多變量分析第五冊), P. R. Krishnaiah，ed. Amsterdam (阿姆斯特丹): North-Holland (北荷蘭).

Rickman, R., N. Mitchell, J. Dingman, 和 J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet (實行 Stillman 飲食法期間血膽固醇改變情形). Journal of the American Medical Association (美國醫學協會彙報), 228, .

Rosenberg, S., 和 M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research (多變量研究中作為資料收集程序之排序方法). Multivariate Behavioral Research (多變量行為研究), 10, .

Roskam, E. E. 1968. Metric analysis of ordinal data in psychology (心理學次序資料之距陣分析). Voorschoten: VAM.

Shepard, R. N. 1962. The analysis of proximities: (相似性分析)Multidimensional scaling with an unknown distance function I (未知距離函數多元尺度法 I). Psychometrika (心理學計量報導), 27, .

Shepard, R. N. 1962. The analysis of proximities: (相似性分析)Multidimensional scaling with an unknown distance function I (未知距離函數多元尺度法 II). Psychometrika (心理學計量報導), 27, .

Shepard, R. N. 1966. Metric structures in ordinal data (次序資料之距陣結構). Journal of Mathematical Psychology (數學心理學雜誌), 3, .

Tenenhaus, M., 和 F. W. Young. 1985. An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis, and other methods for quantifying categorical multivariate data (量化類別多變量資料的多重對應分析、最適尺度、雙尺度、均質性分析及其他方法之分析與結合). Psychometrika (心理學計量報導), 50, .

Theunissen, N. C. M., J. J. Meulman, A. L. Den Ouden, H. M. Koopman, G. H. Verrips, S. P. Verloove-Vanhorick, 和 J. M. Wit. 2003. Changes can be studied when the measurement instrument is different at different time points (在不同時間點依量測工具改變來研究變更). Health Services and Outcomes Research Methodology (健康服務與結果研究方法論), 4, .

Tucker, L. R. 1960. Intra-individual and inter-individual multidimensionality (個體內與個體間多元性). 於： Psychological Scaling: (心理學尺度法)Theory & Applications (理論與應用), H. Gulliksen, 和 S. Messick，eds. 紐約: John Wiley and Sons.

Van der Burg, E. 1988. Nonlinear canonical correlation and some related techniques (非線性典型相關及相關技巧舉例). Leiden (萊頓): DSWO Press (DSWO 出版社).

Van der Burg, E., 和 J. De Leeuw. 1983. Nonlinear canonical correlation (非線性典型相關). British Journal of Mathematical and Statistical Psychology (英國數學暨統計心理學雜誌), 36, .

Van der Burg, E., J. De Leeuw, 和 R. Verdegaal. 1988. Homogeneity analysis with k sets of variables: (K 變數集之均齊性分析)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 53, .

Van der Ham, T., J. J. Meulman, D. C. Van Strien, 和 H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: (依經驗法將青少年飲食異常次組別化)A longitudinal perspective (縱向觀點). British Journal of Psychiatry (英國心理學雜誌), 170, .

Van der Kooij, A. J., 和 J. J. Meulman. 1997. MURALS:Multiple regression and optimal scaling using alternating least squares (使用替代最小平方之多重迴歸與最適尺度). 於: Softstat ' 97, F. Faulbaum, 和 W. Bandilla, eds. Stuttgart: Gustav Fisher.

Van Rijckevorsel, J. 1987. The application of fuzzy coding and horseshoes in multiple correspondence analysis (多重對應分析之模糊編碼與馬蹄應用). Leiden (萊頓): DSWO Press (DSWO 出版社).

Verboon, P., 和 I. A. Van der Lans. 1994. Robust canonical discriminant analysis (穩健典型判別分析). Psychometrika (心理學計量報導), 59, .

Verdegaal, R. 1985. Meer sets analyse voor kwalitatieve gegevens (in Dutch) (更多性質資料的集合分析 (荷蘭文)). Leiden (萊頓): Department of Data Theory, University of Leiden (萊頓大學資料理論系).

Vlek, C., 和 P. J. Stallen. 1981. Judging risks and benefits in the small and in the large (判斷小規模及大規模之風險與惠益). Organizational Behavior and Human Performance (組織行為與人員績效), 28, .

Wagenaar, W. A. 1988. Paradoxes of gambling behaviour (賭博行為之弔詭). 倫敦: Lawrence Erlbaum Associates, Inc.

Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, 和 B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities (六個城市中的兒童二手煙、氣體煮食與呼吸道健康). American Review of Respiratory Diseases (美國呼吸道疾病評論), 129, .

Winsberg, S., 和 J. O. Ramsay. 1980. Monotonic transformations to additivity using splines (利用曲線進行可加性單調轉換). Biometrika (生物統計學雜誌), 67, .

Winsberg, S., 和 J. O. Ramsay. 1983. Monotone spline transformations for dimension reduction (維度縮減之單調曲線轉換). Psychometrika (心理學計量報導), 48, .

Wolter, K. M. 1985. Introduction to variance estimation (變異數估計簡介). 柏林: Springer-Verlag.

Young, F. W. 1981. Quantitative analysis of qualitative data (性質資料之量化分析). Psychometrika (心理學計量報導), 46, .

Young, F. W., J. De Leeuw, 和 Y. Takane. 1976. Regression with qualitative and quantitative variables: (性質與數值變數迴歸)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 41, .

Young, F. W., Y. Takane, 和 J. De Leeuw. 1978. The principal components of mixed measurement level multivariate data: (混合測量水準多變量資料之主成份)An alternating least squares method with optimal scaling features (最適尺度特性之替代最小平方法). Psychometrika (心理學計量報導), 43, .

Zeijl, E., Y. te Poel, M. du Bois-Reymond, J. Ravesloot, 和 J. J. Meulman. 2000. The role of parents and peers in the leisure activities of young adolescents (青少年休閒活動中父母與友伴的角色). Journal of Leisure Research (休閒研究雜誌), 32, .

最大值, J. 1960. Quantizing for minimum distortion (最小誤差量化). Proceedings IEEE (Information Theory) (IEEE (資訊理論) 會議記錄), 6, .

# 索引


ANOVA
  在「類別迴歸」中, 19
Cronbach's alpha 值
  類別量化, 129
DeSarbo 互相混合指數
  在多維度展開中, 244, 247, 253, 259, 266
legal notices, 281
PREFSCAL, 72
$R^2$ 值
  在「類別迴歸」中, 94
Shepard 圖
  在多維度展開中, 77
Shepard 粗略非簡併指數
  在多維度展開中, 244, 247, 253, 259, 266
trademarks, 282

三因子展開
  在多維度展開中, 248
三角散佈圖
  類別量化, 31
不適值
  在非線性典型相關分析中, 180

主成份分析常態化
  在對應分析中, 196

係數
  在「類別迴歸」中, 94
信賴統計量
  在對應分析中, 45
個別空間
  在多維度展開中, 255, 260
個別空間加權
  在多元尺度方法中, 69
  在多維度展開中, 79
個別空間圖
  在多元尺度方法中, 68
  在多維度展開中, 77
個別空間座標。
  在多維度展開中, 79
個別空間權重圖
  在多元尺度方法中, 68
  在多維度展開中, 77
個別空間的聯合圖形
  在多維度展開中, 255, 260
個體分數
  在多重對應分析中, 53, 211, 214
  在非線性典型相關分析中, 39
  類別量化, 29, 131, 135, 153
個體點圖
  在多重對應分析中, 55
  類別量化, 31
偏相關
  在「類別迴歸」中, 95
偏離值
  在多重對應分析中, 216

共同空間
  在多元尺度方法中, 234, 238
  在多維度展開中, 245, 248, 254, 260, 267, 271
共同空間圖
  在多元尺度方法中, 68
  在多維度展開中, 77
共同空間座標
  在多元尺度方法中, 69
  在多維度展開中, 79
共同空間的聯合圖形
  在多維度展開中, 245, 248, 254, 260, 267, 271
共同空間的限制
  在多維度展開中, 75

列主成份分析常態化
  在對應分析中, 196
列分數圖形
  在對應分析中, 203
初始共同空間圖
  在多維度展開中, 77

加權值
  在非線性典型相關分析中, 39, 181
加權的歐基里得模式
  在多維度展開中, 73

區別量數
  在多重對應分析中, 53, 212
區別量數圖
  在多重對應分析中, 56

圖形
  在多元尺度方法中, 68–69
  在對應分析中, 46
  在非線性典型相關分析中, 39
  在「類別迴歸」中, 21

壓力係數圖
  在多元尺度方法中, 68


© Copyright SPSS Inc. 1989, 2010

在多維度展開中, 77
壓力係數測度
在多元尺度方法中, 69, 233, 238
在多維度展開中, 79

多元尺度方法, 58, 60–63, 222
共同空間, 234, 238
圖形, 58, 68–69
壓力係數測度, 233, 238
指令的其他功能, 71
模式, 64
統計量, 58
輸出, 69
轉換圖, 237
選項, 66
限制, 65
多維度展開, 72, 241, 262
三因子展開 , 248
個別空間, 255, 260
共同空間, 245, 248, 254, 260, 267, 271
共同空間的限制, 75
圖形, 72, 77
指令的其他功能, 80
模式, 73
相似性轉換, 268, 272
簡併解, 241
統計量, 72
輸出, 79
選項, 76
量數, 244, 247, 253, 259, 266, 270
多重對應分析, 48, 52, 207
個體分數, 211, 214
偏離值, 216
儲存變數, 55
區別量數, 212
指令的其他功能, 57
最適尺度水準, 50
模式摘要, 210
類別量化, 213
多重起點圖
在多維度展開中, 77

套索
在「類別迴歸」中, 18

對應分析, 41–46, 196–197
列分數圖形, 203
圖形, 41
常態化, 196
指令的其他功能, 47
統計量, 41
維度, 201
行分數圖形, 203
貢獻, 202
對稱常態化
在對應分析中, 196

尺度模式
在多維度展開中, 73

常態化
在對應分析中, 44, 196

彈性網絡
在「類別迴歸」中, 18

慣量
在對應分析中, 45
懲罰應力
在多維度展開中, 244, 253, 259, 266, 270
懲罰條款
在多維度展開中, 76

成份負荷量
在非線性典型相關分析中, 39, 183
類別量化, 29, 133, 136, 151
成份負荷量圖
類別量化, 33

投影重心
在非線性典型相關分析中, 187
投影重心圖
類別量化, 32

敘述統計
在「類別迴歸」中, 19

最終共同空間圖
在多維度展開中, 77
最適尺度水準
在多重對應分析中, 50
類別量化, 25

概化的歐基里得模式
在多維度展開中, 73
標準化
在對應分析中, 44
模式摘要
在多重對應分析中, 210

殘差
在「類別迴歸」中, 97
殘差圖
在多維度展開中, 77

特徵值
在非線性典型相關分析中, 180
類別量化, 129, 135, 151

疊代歷程
  在多元尺度方法中, 69
  在多維度展開中, 79
  在多重對應分析中, 53
  類別量化, 29, 129
疊代準則
  在多元尺度方法中, 66
  在多維度展開中, 76

相似性轉換
  在多維度展開中, 73
相關
  在多元尺度方法中, 69
相關圖
  在多元尺度方法中, 68
相關性
  在「類別迴歸」中, 93
相關矩陣
  在多重對應分析中, 53
  類別量化, 29

空間加權圖
  在多維度展開中, 77

範例檔案
  位置, 273

維度
  在對應分析中, 44, 201
維度加權
  在多維度展開中, 255, 260

聯合類別圖
  在多重對應分析中, 56
  類別量化, 32

脊迴歸
  在「類別迴歸」中, 18

行主成份分析常態化
  在對應分析中, 196
行分數圖形
  在對應分析中, 203
行列散佈圖
  在多重對應分析中, 55
  在對應分析中, 46
  類別量化, 31

補充性的分析標
  在「類別迴歸」中, 16
複相關係數 R
  在「類別迴歸」中, 19

識別模式
  在多維度展開中, 73
變數權重
  在多重對應分析中, 50
  類別量化, 25
變異係數
  在多維度展開中, 244, 247, 253, 259, 266
變異歸因於
  類別量化, 29, 129, 151

貢獻
  在對應分析中, 202

起始組態
  在多元尺度方法中, 66
  在多維度展開中, 76
  在非線性典型相關分析中, 39
  在「類別迴歸」中, 16

距離
  在多元尺度方法中, 69
  在多維度展開中, 79
距離測量
  在對應分析中, 44

轉換圖
  在多元尺度方法中, 68, 237
  在多維度展開中, 77, 268, 272
  在多重對應分析中, 56
  在「類別迴歸」中, 96
  類別量化, 32
轉換後的相似性
  在多元尺度方法中, 69
  在多維度展開中, 79
轉換後的自變數
  在多元尺度方法中, 69

迴歸係數
  在「類別迴歸」中, 19
適合
  在非線性典型相關分析中, 39
適合度值
  在非線性典型相關分析中, 180
適合度的散佈圖
  在多維度展開中, 77
遺漏值
  在多重對應分析中, 51
  在「類別迴歸」中, 15
  類別量化, 26

部分相關
  在「類別迴歸」中, 95

重心
  在非線性典型相關分析中, 39, 187
重要性
  在「類別迴歸」中, 95
量化
  在非線性典型相關分析中, 184
  類別量化, 130, 149

限制
  在多元尺度方法中, 65
隨意更新
  在多元尺度方法中, 66

離散化
  在多重對應分析中, 50
  在「類別迴歸」中, 14
  類別量化, 26

零階相關
  在「類別迴歸」中, 95

非線性典型相關分析, 35, 38, 173
  分析摘要, 180
  加權值, 181
  圖形, 35
  成份負荷量, 181, 183
  指令的其他功能, 40
  統計量, 35
  重心, 187
  量化, 184
  類別座標, 187

類別主成份分析, 23, 28, 124, 136
  個體分數, 131, 135, 153
  儲存變數, 31
  成份負荷量, 133, 136, 151
  指令的其他功能, 34
  最適尺度水準, 25
  模式摘要, 129, 135, 151
  疊代歷程, 129
  量化, 130, 149
  類別點, 154
類別圖
  在多重對應分析中, 56
  類別量化, 32
類別座標
  在非線性典型相關分析中, 187
類別迴歸, 12, 82
  儲存, 20
  圖形, 12
  指令的其他功能, 22
  最適尺度水準, 13
  模式適合度, 94
  殘差, 97
  相關, 94–95
  相關性, 93
  統計量, 12
  規則化, 18
  轉換圖, 96
  重要性, 95
類別量化
  在多重對應分析中, 53, 213
  在非線性典型相關分析中, 39
  在「類別迴歸」中, 19
  類別量化, 29
類別點
  類別量化, 154