

IBM SPSS Regression 22

IBM

附註

使用此資訊和支援的產品之前，請先閱讀 第 27 頁的『注意事項』中的資訊。

產品資訊

若新版本未聲明，則此版本便適用於 IBM® SPSS® Statistics 22.0.0 版以及之後發行的所有版本和修正。

目錄

第 1 章 選擇二元羅吉斯迴歸程序	1	非線性迴歸的損失函數	15
第 2 章 羅吉斯迴歸	3	非線性迴歸的參數限制	15
羅吉斯迴歸分析的設定規則	4	非線性迴歸的儲存新變數	16
線性迴歸的變數選取方法	4	非線性迴歸選項	16
羅吉斯迴歸分析定義類別變數	4	說明非線性迴歸的結果	16
羅吉斯迴歸分析儲存新變數	5	NLR 指令的其他功能	16
羅吉斯迴歸分析選項	5	第 6 章 加權估計	19
LOGISTIC REGRESSION 指令的其他功能	6	權重估計的選項	20
第 3 章 多項式羅吉斯迴歸	7	WLS 指令的其他功能	20
多項式羅吉斯迴歸	7	第 7 章 二階最小平方迴歸	21
建置項目	8	二階最小平方迴歸選項	22
多項式羅吉斯迴歸參考類別	8	2SLS 指令的其他功能	22
多項式羅吉斯迴歸統計量	8	第 8 章 類別變數編碼方法	23
多項式羅吉斯迴歸條件	9	偏差	23
多項式羅吉斯迴歸選項	9	簡單	23
多項式羅吉斯迴歸的儲存	10	Helmert	24
NOMREG 指令的其他功能	10	差異	24
第 4 章 機率值分析	11	多項式	24
機率值分析的定義範圍	12	重複	25
機率值分析的選項	12	特殊	25
PROBIT 指令的其他功能	12	指示燈	25
第 5 章 非線性迴歸	13	注意事項	27
條件邏輯 (非線性迴歸)	13	商標	28
非線性迴歸的參數	14	索引	29
非線性迴歸的常見模型	14		

第 1 章 選擇二元羅吉斯迴歸程序

二元羅吉斯迴歸模型適合使用羅吉斯迴歸程序或多元羅吉斯迴歸程序。每種程序都有其他程序未提供的選項。理論上很重要的差異是羅吉斯迴歸程序會產生所有的預測、殘差、影響統計量、以及在個別觀察值等級使用資料的適合度測試，而不管資料是如何輸入的，以及共變異數型樣的數量是否小於觀察值的總數量。但是多元羅吉斯迴歸程序會內部整合觀察值以形成預測變數相同的共變異數型樣的次母體，以產生預測、殘差、以及根據這些次母體的適合度測試。如果所有的預測變數都是類別變數，或是任何連續預測變數只具有有限的變數值以使每個共變異數型樣中都有數個個案 次母體方式可以產生有效的適合度檢定和情報殘差，但是個別個案等級方法則不能。

羅吉斯迴歸分析 提供下列獨特的功能：

- Hosmer-Lemeshow 適合度測試模型
- 逐步迴歸分析
- 對照定義模式 parameterization
- 其他分類選擇切割點
- 分類圖(C)
- 符合觀察值集合到保留觀察值集合的模式。
- 儲存預測、殘差、以及影響統計量

多元羅吉斯迴歸提供下列獨特的功能：

- 模式適合度的 Pearson 和離差卡方檢定
- 為適合度測試將資料分組的次母體的規格。
- 列出次母體的次數、預測次數、以及殘差
- 過於分散情形的變異數估計修正
- 參數估計的共變異數矩陣
- 參數的線性結合測試
- 巢狀模型的清楚規格
- 適合 1-1 使用差異變數的符合條件羅吉斯迴歸模型

第 2 章 羅吉斯迴歸

當您希望能夠根據預測值變數集的數值來預測特性或結果的出現或缺席時，羅吉斯迴歸分析就很有用。它和線性迴歸模型很相似，但是適合二分因變數的模型。羅吉斯迴歸係數可以用來估計模式中每一個自變數的勝算比。羅吉斯迴歸分析適用在較廣範圍的研究情況，而不是區別分析。

範例。 對冠狀動脈心臟疾病 (CHD) 而言，什麼樣的生活型態特性是風險因素？假定以病人樣本來測量抽煙狀況、飲食、運動、酒精使用情形 以及 CHD 狀況，您可以利用這四種生活型態變數來建置模型，並預測在病人樣本中 CHD 的出現或缺席。之後可以用這個模式得到每個因素的勝算比的預估，舉例來說，告訴您吸煙者比不吸煙者更容易得到 CHD 的可能性。

統計量。 對於每一個分析：總觀察值、選取的觀察值、有效觀察值。對每一個類別變數：參數編碼。對於每一個步驟：輸入或移除的變數、疊代歷程、2 對數概似、適合度、Hosmer-Lemeshow 適合度統計量、模式卡方分配、改良卡方分配、分類表、相關變數、觀察組和預測機率圖、殘差卡方。對於方程式中的每一個變數：係數 (B)、 B 的標準誤、Wald 統計、預估勝算比 ($\exp(B)$)、 $\exp(B)$ 的信賴區間、若從模式移除項的對數概似。對每一個不在方程式中的每個變數：統計量評分。對於每一個觀察值：觀察組、預測機率、預測組、殘差、標準化殘差。

方法。 您可以使用變數的區塊進入點，或使用任何下列的逐步迴歸分析法來預測模型：向前條件、向前 LR、向前 Wald、向後條件、向後 LR、向後 Wald。

羅吉斯迴歸分析的資料考量

資料。 因變數應該是分成兩個的。自變數可以是區間等級或是類別的。如果它是類別的，那麼就應該是虛擬或是編碼過的指標（程序中有個選項可以自動將類別變數重新編碼）。

假設。 羅吉斯迴歸分析依靠分配假設的方法與區別分析不太相同。但是，如果預測值是多變量常態分配，那麼您的解答可能會更穩定。另外，如同其他的迴歸形式，預測值中的多重共線性會導致偏差估計及誇大的標準誤。當組別成員為真的類別變數時，此程序將可發揮最高效能；但如果組別成員是以連續變數的值為基礎（例如高 IQ 對 低 IQ），那麼您應該考慮使用線性迴歸，以便利用連續變數本身所提供的豐富資訊。

相關程序。 使用「散佈圖」程序來審查資料的多重共線性。當多變量常態和相同變異數-共變異數矩陣的假設相符合時，使用區別分析程序也許可以更快得到解答。如果所有的預測值變數都是類別變數的話，您也可以使用「對數線性」程序。如果您的因變數是連續的，請考慮使用「線性迴歸」程序。您可以使用 ROC 曲線程序來繪製「羅吉斯迴歸」程序所儲存的機率。

取得羅吉斯迴歸分析

1. 在功能表上，選擇：

分析 > 迴歸 > 二進位邏輯...

2. 選取一個二分因變數。這個變數可以為數值或字串。

3. 選取一個或多個共變量。若要包括交互作用項，請選擇所有和交互作用有關的變數，再選取「>a*b>」。

若要輸入群組（**區塊**）中的變數，請選取區塊的共變量，並且按一下**下一步**以指定新的區塊。一直重複到指定完所有的區塊。

或者您可以選取觀察值進行分析。選取一個選擇變數，然後按一下**規則**。

羅吉斯迴歸分析的設定規則

由選擇規則所定義出來的觀察值會包含在模型預測中。例如，如果您選取一個變數和等式，並且指定數值 5，那麼預測模型中只會包括觀察值等於 5 的選定變數。

選取和未選取的觀察值會產生統計量和分類結果。這樣提供了根據之前存在的資料將新觀察值分類、或是將資料分隔成訓練或測試子集的機制，以便在產生模式上執行驗證。

線性迴歸的變數選取方法

您可以利用選擇方法來指定如何在分析中輸入自變數。您可以使用不同方法，從相同變數集來建構各種迴歸模型。

- **輸入。**一種變數選取程序，在其單一步驟中會輸入區塊中的所有變數。
- **向前選取法（條件式）(Forward Selection (Conditional))。**一種逐步選取方法，其輸入檢定是以分數統計量顯著性為基礎，移除檢定則是以基於條件式參數估計值的概似比統計量機率為基礎。
- **向前選取法（概似比）(Forward Selection (Likelihood Ratio))。**一種逐步選取方法，其輸入檢定是以分數統計量顯著性為基礎，移除檢定則是以基於最大偏概似估計值的概似比統計量機率為基礎。
- **向前選取法 (Wald) (Forward Selection (Wald))。**一種逐步選取方法，其輸入檢定是以分數統計量顯著性為基礎，移除檢定則是以 Wald 統計量機率為基礎。
- **向後消去法（條件式）(Backward Elimination (Conditional))。**向後逐步選取。移除檢定是以基於條件式參數估計值的概似比統計量機率為基礎。
- **向後消去法（概似比）(Backward Elimination (Likelihood Ratio))。**向後逐步選取。移除檢定是以基於最大偏概似估計值的概似比統計量機率為基礎。
- **向後消去法 (Wald) (Backward Elimination (Wald))。**向後逐步選取。移除檢定是以 Wald 統計量機率為基礎。

您輸出中的顯著值會符合某個單一模式。因此，使用逐步迴歸分析法時通常會產生無效的顯著數值。

所有選用的自變數，都會被加入單一迴歸模型中。然而，您可以替不同的變數子集，指定多種不同的輸入方式。例如，您可以透過「逐步選取」這種方式，將某個區塊中的變數輸入迴歸模型中，而第二個區塊則使用「向前選擇」方式。若要將第二個變數區塊加入迴歸模型中，請按一下下一個。

羅吉斯迴歸分析定義類別變數

您可以指定「羅吉斯迴歸」程序如何處理類別變數的細節：

共變量。包含一份清單，顯示使用者透過主要對話框（或者任何一層選單）所指定的每一個共變量。如果部分共變量是字串變數或類別變數的話，那麼您就只能把它們當成類別共變量使用。

類別共變量。列出所有被視成類別的變數。每個變數包含一個符號（放在括弧中），這個符號代表即將使用的對比編碼。字串變數（其名稱後面跟著符號 <）已經出現在「類別共變量」清單中。請從「共變量」清單選取別的類別共變量，然後再將他們移到「類別共變量」清單上。

變更對比。讓您可以變更對比方法。目前可用的對比方法如下：

- **指標。**對比指出該類別成員是否存在。對比矩陣中的參考類別是用一整列的零來表示。
- **簡單。**預測變數的所有類別（除了參考類別以外）都會跟參考類別相比較。
- **差異。**預測變數的所有類別（除了第一個類別以外）都會跟先前類別的平均效果相比較。這種對比也叫作反「赫爾莫特 (Helmert) 對比」。
- **Helmert。**預測變數的所有類別（除了最後一個類別以外）都會跟後續類別的平均效果相比較。

- **重複**。預測變數的所有類別（除了第一個類別以外）都會跟它前一個類別相比較。
- **多項式**。正交多項式對比。它假設類別間距都是相等的。多項式對比只適用於數值變數。
- **離差**。預測變數的所有類別（除了參考類別以外）都會跟整體效果相比較。

如果您選取「離差」、「簡單」或「指標」的話，請選取「第一個」或「最後一個」當成參考類別。請注意，只有在您按一下「變更」後，方法才會真的隨之改變。

字串共變量必須是類別共變量。若要從「類別共變量」清單中移除字串變數，您必須先從主要對話框中的「共變量」清單中，移除所有包含該變數的項目。

羅吉斯迴歸分析儲存新變數

您可以將羅吉斯迴歸的結果存成作用中資料集的新變數。

預測值。儲存模型預測值。可用的選項為「機率」和「各組成員」。

- **機率(P)**。為每一個觀察值儲存該事件發生的預測機率。輸出中的表格會顯示任何新變數的名稱及內容。「事件」是具有較高值的應變數類別；例如，如果應變數採用值 0 和 1，則會儲存類別的預測機率。
- **預測群組成員資格 (Predicted Group Membership)**。根據區別分數，事後機率最大的群組。也是模型預測觀察值所屬的群組。

影響。儲存統計量的值，它可以測量觀察值在預測值上的影響。可用的選項有 Cook's、槓桿值、以及 DfBeta (s)。

- **Cook's(K)**。Cook 影響統計量的 Logistic 迴歸類比。在從迴歸係數計算中排除特定觀察值時，所有觀察值殘差變更程度的量數。
- **槓桿值 (Leverage Value)**。模式適合度中之每個觀察值的相對影響。
- **DfBeta(B)**。beta 值的差異即為因排除特定觀察值而導致的迴歸係數變更。值是針對模型中的每一個項目計算（包含常數）。

殘差。儲存殘差。可用的選項為未標準化、Logit、Studentized、標準化、以及離差。

- **未標準化殘差 (Unstandardized Residuals)**。觀察值與模型所預測的值之間的差異。
- **Logit 殘差 (Logit Residual)**。若以 Logit 尺度進行預測，則此為觀察值的殘差。Logit 殘差等於除以 1 預測機率次數後，再減去預測機率所得出的殘差。
- **學生化殘差 (Studentized Residual)**。排除觀察值時，模型離差中的變更。
- **標準化殘差**。殘差除以其標準差的估計值。標準化殘差（也稱為 Pearson 殘差）的平均數為 0，標準差為 1。
- **離差**。以模型離差為基礎的殘差。

匯出模型資訊至 XML 檔。將參數估計與（選擇性）其共變異數以 XML (PMML) 格式匯出到指定檔案中。您可以使用這個模型檔案，將模式資訊套用到其他資料檔中以進行評分工作。

羅吉斯迴歸分析選項

您可以指定羅吉斯迴歸分析的選項：

統計量與圖形。可讓您要求統計量及圖形。可用的選項有分類圖、Hosmer-Lemeshow 適合度、依觀察值順序列出殘差、估計值相關性、疊代歷程、以及 $\exp(B)$ 的 CI。在「顯示」群組中選取一選項，在每一步驟或是在最終模型的最後步驟中顯示統計量及圖形。

- *Hosmer-Lemeshow* 適合度統計量。此適合度統計量較 Logistic 迴歸所使用的傳統適合度統計量更穩健，尤其適用於擁有連續共變量的模型及樣本大小較小的研究。其基礎運作方式，是以風險十分位數將觀察值予以分組，並比較每個十分位數中的觀察機率與預期機率。

逐步之機率。 讓您可以控制從等式輸入或移除變數的條件。您可以指定變數進入或移除的條件。

- **逐步之機率。** 如果變數的分數統計量之機率小於「選入」值的話，系統會將變數輸入模型中；如果機率大於「移除」值的話，系統會將變數移除。若要覆寫預設值，「選入」和「移除」請輸入正值。「選入」必須小於「移除」。

分類分割值。 讓您可以決定分類觀察值時的分割點。超過分類分割的預測觀察值會分類為正向，而小於分割的預測觀察值會分類為負向。若要改變預設值，請輸入介於 0.01 和 0.99 之間的數值。

最大疊代。 讓您可以改變模式在終止之前的最多疊代次數。

常數項納入模式中。 讓您可以指出模式是否應該要包括常數項。如果不可以，常數項應該等於 0。

LOGISTIC REGRESSION 指令的其他功能

指令語法語言也可以讓您：

- 藉由變數的值或變數標籤來確認觀察值輸出。
- 控制疊代的報告間隔。不一定要在每個疊代之後列印出參數估計，您可以要求在每 n 個疊代之後列印參數估計。
- 變更終止疊代和檢查餘數的條件。
- 指定觀察值清單的變數清單。
- 執行過程中，將每個分割檔組別的資料保留在外部暫存檔中，以節省記憶體。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

第 3 章 多項式羅吉斯迴歸

當您想要依據預測變數集合值來分類觀察值時，多項式羅吉斯迴歸就很有用。這種迴歸方法與羅吉斯迴歸相似，但因其因變數不限於兩種類別，所以用途更為廣泛。

範例。 為了更有效地銷售影片，電影製片廠想要預測常看電影的人很可能看哪種類型影片。藉由執行「多項式羅吉斯迴歸」，製片廠可以確定年齡、性別及約會狀態對其偏好影片類型的影響強度，然後製片廠就可將特定電影的廣告活動偏向可能去看的那一群人。

統計量。 計有疊代歷程、參數的係數、漸進線共變異數與相關性矩陣、模式與淨效果的概似比檢定、- 2 對數概似。Pearson 和共變異卡方適配度，Cox 和 Snell、Nagelkerke 和 McFadden R^2 。分類：依回應類別所做的觀察對預測的次數分配。交叉表格：依共變異數型樣及回應類別，所做的觀察對預測的次數分配（含殘差）及比例。

方法。 多項式對數成敗比模型適合用在完全因素模型或使用者指定的模型。參數估計係經由疊代的最大概似演算法來進行的。

多項式羅吉斯迴歸分析的資料考量

資料。 因變數必須是類別的。而自變數則可以是因素或共變量。一般而言，因素應該是類別變數，而共變量應該是連續變數。

假設。 假設任何兩個類別的勝算比是與其他所有回應類別獨立的。舉例來說，如果一項新產品引進市場，這項假設即在說明，其他所有產品的市場佔有率係等比例地受影響。同樣地，如果有共變異數型樣，回應值即假設為獨立的多項式變數。

取得多項式羅吉斯迴歸分析

1. 在功能表上，選擇：
 分析 > 迴歸 > 多名義邏輯...
2. 選取一個因變數。
3. 因素是可隨意選取的，而且可以是數值或類別變數。
4. 共變量也是可隨意選取的，但若經選取，就必須是數值變數。

多項式羅吉斯迴歸

根據預設，「多項式羅吉斯迴歸」程序會產生包含因素、共變量主效應的模式，但是您可以使用此對話框，指定自訂模型，或是要求逐步模式選取方法。

指定模式。 主效應模型會包括共變量及因素的主效應，但不包括交互作用效應項。完全因素模型包含所有的主效應，以及所有因素對因素的交互作用。該模式卻不包含共變量的交互作用。您可以建立自訂模型，以指定因素交互作用、或共變量交互作用的次集，或是要求逐步選取模式項。

因素 & 共變量。 因素與共變量均會列出。

強制選入項目。 模式中會永遠包含加入到「強制選入項目」清單的項目。

逐步模式項目。 根據下列使用者選定逐步迴歸分析法之一，會在模式中包含加入逐步清單中的項目：

- **向前選入。** 這個方法在模式一開始時不包含逐步項目。在每一個步驟中，最顯著的項目會逐步被加入到模式中，直到剩餘的逐步項目就算被加入，在統計上也不具有顯著貢獻。
- **向後消去法。** 這個方法一開始就會將逐步清單中所指定的所有項目，輸入到模式中。在每一個步驟中，該模式中的最不顯著逐步項目會逐步被移除，直到所有剩餘的逐步項目在統計上都對該模式具有顯著貢獻。
- **向前逐步。** 這個方法一開始就使用向前選入方法所選定的模式。使用此方法時，演算法會在模式中逐步項目的向後消去法、以及脫離模式之項目的向前選入之間進行切換。這個步驟會不斷持續，直到沒有任何項目符合選入或是移除條件。
- **向後逐步。** 這個方法一開始就使用向後消除方法所選定的模式。使用此方法時，演算法會在脫離模式之項目的向前選入、以及模式中逐步項目的往後消去法之間進行切換。這個步驟會不斷持續，直到沒有任何項目符合選入或是移除條件。

模式中包括截距。 讓您為模式包含或排除截距項目。

建置項目

對所選的因素和共變量而言：

交互作用。 建立所有選定變數的最高階交互作用項。

主效應。 為每個所選的變數，建立主效應項目。

完全雙向。 為所選的變數，建立所有可能的雙向交互作用。

完全三向。 為所選的變數，建立所有可能的三因素交互作用。

完全四向。 為所選的變數，建立所有可能的四向交互作用。

完全五向。 為所選的變數，建立所有可能的五向交互作用。

多項式羅吉斯迴歸參考類別

根據預設，「多項式羅吉斯迴歸」程序會將最後一個類別當作參考類別。這個對話框可讓您控制參考類別，以及類別的排列方式。

參考類別。 指定第一個、最後一個或自訂的類別。

類別次序。 使用遞增順序時，最低值是定義第一個類別，而最高值是定義最後一個類別。使用遞減順序時，最高值是定義第一個類別，而最低值是定義最後一個類別。

多項式羅吉斯迴歸統計量

您可對「多項式羅吉斯迴歸」指定下列統計量：

觀察值處理摘要。 本表格包含所指定類別變數的相關資訊。

模式。 整體模型的統計量。

- **虛擬迴歸係數 (Pseudo R-square)。** 列印出 Cox 與 Snell、Nagelkerke 及 McFadden R^2 統計量。
- **逐步摘要。** 本表格摘要在使用逐步方法時，每個步驟所輸入或移除的效果項。這個表格只有在模型對話框中指定時，才可建立。
- **模式適合度資訊。** 本表格會針對適合的、僅含截距，或虛無模型進行比較。
- **資訊準則。** 這個表格可列印「Akaike 資訊準則」(AIC) 及「Schwarz Bayesian 資訊準則」(BIC)。

- **格機率**。 列印出依共變異數型樣及回應類別，所做的觀察與所期望次數分配（含殘差）及比例的表格。
- **分類表**。 列印出觀察對所預測回應值的表格。
- **適合度卡方統計量**。 列印出 Pearson 與概似比卡方統計量。這些統計量是就共變異數型樣計算出來的，而此樣式是依所有因素及共變量，或依使用者定義的因素及共變量次集而定的。
- **Monotonicity 量數**。 顯示包含和諧成對、不和諧成對和等值成對數量等相關資訊的表格。Somers' D、Goodman 與 Kruskal's Gamma、Kendall's tau-a 以及和諧指數 C 也會呈現在本表。

參數。與模型參數相關的統計量。

- **估計值**。 列印出包含使用者指定信賴區間的模型參數估計值。
- **概似比檢定**。 列印出對模式淨效果的概似比檢定，而對整體模型的檢定會自動印出。
- **漸近相關**。 列印出參數估計值相關性矩陣。
- **漸近共變異數**。 列印出參數估計值共變異數矩陣。

定義次母體 可讓您選取因素及共變量的次集，以便定義格機率及適合度檢定所用的共變異數型樣。

多項式羅吉斯迴歸條件

您可對「多項式羅吉斯迴歸」指定下列條件：

疊代。可讓您指定想要演算法循環的最大次數、半階 (step-halving) 中的最大步驟數目、對數概似與參數中變更的收斂容差度、多久列印出一次疊代演算法的進度，以及程序要在哪次疊代時開始檢查完成、或似是完成的資料分組。

- **對數概似收斂**。 如果對數概似函數中的絕對變化小於指定值，即假定收斂。如果數值是 0，就不會使用這個準則。請指定一個非負數的數值。
- **參數收斂條件**。如果參數估計值中的絕對變化小於這個數值，即假定收斂。如果值為 0 便不會使用這個條件。

Delta。 可讓您指定小於 1 的非負數值，此數值會加入依共變異數型樣所做的回應類別交叉表中的各個空格。此將有助於穩定演算法，並防止估計值中的偏誤。

奇異性容忍值。 可讓您指定用於檢查單一性的容忍值。

多項式羅吉斯迴歸選項

您可以為「多項式羅吉斯迴歸」指定下列選項：

分散尺度。可讓您指定分散尺度數值，以用於改正參數共變異數矩陣的估計值。**離差**會使用離差函數（概似比卡方）統計值，來估計尺度數值。**Pearson**則使用皮爾森 (Pearson) 卡方統計值，來估計尺度數值。您也可以指定自己的尺度數值，但其必須為正數值。

逐步選項。 這些選項可讓您在逐步方法來建置模型時，用來控制統計條件。除非這個表格在模型對話框中指定，否則將被忽略。

- **選入機率**。 這是指變數選入的概似比統計量的機率。指定的機率越大，就越容易在模式中輸入變數。這個條件只會在選定向前選入、向前逐步或向後逐步等方法時出現。
- **選入檢定**。 這個方法是以逐步迴歸分析法輸入項目。在概似比檢定及計分檢定中二選一。這個條件只會在選定向前選入、向前逐步或向後逐步等方法時出現。
- **移除機率**。 這是指變數移除的概似比統計量的機率。指定的機率越大，就越容易在模式留下變數。這個條件只會在選定向後消除、向前逐步或向後逐步等方法時出現。

- **移除檢定**。這個方法是以逐步迴歸分析法移除項目。在概似比檢定及 Wald 檢定中二選一。這個條件只會在選定向後消除、向前逐步或向後逐步等方法時出現。
- **模式中的最小逐步效應**。當使用向後消除或向後逐步方法時，這個選項可指定該模式要包含最少數目的項目。截距不當作模式項來計算。
- **模式中的最大逐步效應**。當使用向前選入或向前逐步方法時，這個選項可指定該模式要包含最多數目的項目。截距不當作模式項來計算。
- **階層限制選入與移除項**。這個選項可讓您選擇是否要在包含模式項時設定限制。階層在包含任何項目時都會要求這個選項，所有將進行包含、且較低次序的項目部分，都必須先存在於模式中。例如，如果階層要求成立，模式中必須先同時存在有婚姻狀況和性別兩個因素，才可以加入婚姻狀況*性別交互作用。下面有三個圓形按鈕選項，可用來在判斷階層時決定共變量的角色。

多項式羅吉斯迴歸的儲存

這個「儲存」對話框可讓您將變數儲存到工作檔案，並將模式資訊匯出到外部檔案。

已儲存變數。可以儲存下列變數：

- **估計各類別的回應機率**。此為將因素/共變異數型樣分類至回應類別的模式估計機率。回應變數中有多少個類別，就會有多少個估計機率；最多可以儲存 25 個。
- **預測應答組類**。此為因素/共變異數型樣最大期望機率的回應類別。
- **已預測類別機率**。這是最大的估計回應機率。
- **實際應答機率**。這是將因素/共變異數型樣分類至所觀察類別的估計機率。

匯出模型資訊至 XML 檔。將參數估計與（選擇性）其共變異數以 XML (PMML) 格式匯出到指定檔案中。您可以使用這個模型檔案，將模式資訊套用到其他資料檔中以進行評分工作。

NOMREG 指令的其他功能

指令語法語言也可以讓您：

- 指定因變數的參考類別。
- 加入含有使用者遺漏值的觀察值。
- 藉著將虛無假設指定為參數線性組合，來自訂假設檢定。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

第 4 章 機率值分析

本程序可測量刺激強度及表示對該刺激某項回應的觀察值比例之間的關係。當出現認為是由一些自變數層級所影響或引起的二分輸出時，它就會很好用，而且特別適合實驗性資料。本程序將可讓您估計引起某種回應比例（如有效劑量中位數）所需的刺激強度。

範例。 新殺蟲劑殺螞蟻的效果如何，以及合適的使用濃度為何？您可能會想執行一項實驗，讓螞蟻暴露在不同濃度的殺蟲劑之下，然後記錄殺死的及被暴露的螞蟻數量。將機率值分析應用在這些資料上，您就可以確定濃度及殺死螞蟻之間的關係強度；而如果您想要確保殺死 95% 被暴露的螞蟻的話，還可以確定合適的濃度為何。

統計量。 迴歸係數與標準誤、截距與標準誤、Pearson 適合度卡方、觀察與期望次數分配表，以及自變數有效層級的信賴區間。圖形：轉換後的回應值圖形。

這個程序係使用 NPSOL[®]（由 Gill、Murray、Saunders 及 Wright 提出）所提議並執行的演算法，來估計模型參數。

機率值分析的資料考量

資料。 對自變數的每個數值（或多個自變數的每個數值組合）而言，您的回應值應為觀察值（含有顯示興趣回應的那些數值）個數的計算，而且觀察值變數總數應為含有自變數數值的觀察值總個數。因素變數應該是類別的，而且編碼成整數。

假設。 觀察值應該是獨立的。如果您有許多與觀察值個數有關的自變數數值（就像您在觀察的研究中可能取得的數值一樣），那卡方及適合度統計量就可能無法使用。

相關程序。 機率值分析與羅吉斯迴歸密切相關，事實上，如果您選擇 logit 轉換，那本程序實質上就會計算羅吉斯迴歸。一般來說，機率值分析很適合設計過的實驗，但羅吉斯迴歸就較適合觀察的研究。輸出上的差別會反映出這些不同的強調之處。機率值分析程序會描述有效值在不同回應率（包含有效劑量中位數）下的估計值。而羅吉斯迴歸程序會描述自變數勝算比的估計值。

取得機率值分析

1. 在功能表上，選擇：

分析 > 迴歸 > 機率...

2. 選取回應次數變數。此變數會指出對試驗刺激有回應的觀察值個數，此變數之數值不能為負數。
3. 選取觀察值總數變數。此變數會指出實施刺激的觀察值個數。此變數之數值不能為負數，而且不能小於每個觀察值回應次數變數的數值。

您可隨意選取因素變數。如果要執行這項操作，請按一下**定義範圍**來定義組別。

4. 選取一或多個共變量。此變數會包含應用在每個觀察值的刺激層級。如果您想要轉換共變量，請從「轉換」下拉清單中選取一種轉換方法。如果未套用轉換，而且有控制組別，則該控制組別就會包含在分析中。
5. 選取**機率值**或**對數**模型。
 - **機率值模型 (Probit Model)**. 將機率值轉換（逆累積標準常態分配函數）套用至回應比例。
 - **Logit 模型 (Logit Model)**. 其會將 Logit（對數勝算）轉換套用至回應比例。

機率值分析的定義範圍

此方框可讓您指定將要分析的因素變數層級。該因素層級必須編碼成連續的整數，而且您指定範圍中的所有層級都將會進行分析。

機率值分析的選項

您可以指定機率值分析的選項：

統計量。可讓您要求下列選擇性的統計量：次數分配表、相對中位數、平行檢定及 Fiducial 信賴區間。

- **相對中位數 (Relative Median Potency).** 顯示每對因素層次的中位數比例。同時也會顯示每個相對中位數的 95% 信賴區間限制。如果您沒有因素變數，或者如果您有多個共變量，則無法使用相對中位數。
- **平行檢定 (Parallelism Test).** 所有因素層次具有共同斜度的檢定假設。
- **Fiducial 信賴區間 (Fiducial Confidence Intervals).** 產生特定回應機率所需藥劑的劑量信賴區間。

如果您選取了一個以上的共變量，就無法使用 Fiducial 信賴區間及相對中位數。而相對中位數及平行檢定只有在您選取了因素變數後才能用。

自然回應率。 讓您在在缺少刺激時，仍可指出自然回應率。可用的選項共有「無」、「由資料中計算」或「數值」。

- **從資料計算 (Calculate from Data).** 從樣本資料估計自然回應率。您的資料應包含代表控制層次的觀察值，其共變量的值為 0。機率值會將控制層次的回應比例用作起始值，以便估計自然回應率。
- **值.** 在模型中設定自然回應率（如果您事先知道自然回應率，請選取此項目）。請輸入自然回應比率（比率必須小於 1）。例如，如果刺激為 0 時，回應發生的機率為 10%，請輸入 0.10。

條件。 可讓您控制疊代的參數估計演算法中的參數。您也可以撤銷最大疊代、步驟限制及最佳容差度的預設值。

PROBIT 指令的其他功能

指令語法語言也可以讓您：

- 要求機率值及對數成敗比模型的分析。
- 控制處理遺漏值的方式。
- 依基底而非基底 10 或自然對數，來轉換共變量。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

第 5 章 非線性迴歸

非線性迴歸是一種尋找因變數及一組自變數之間關係非線性模型的方法。不像僅限於估計線性模型的傳統線性迴歸，非線性迴歸尚可用來估計因變數及自變數之間的任意關係模式。這是使用疊代的估計演算法來完成的。請注意，對 $Y = A + BX^2$ 形式的簡單多項式模式而言，這項程序並不是必要的。因為我們可藉由定義 $W = X^2$ ，並使用傳統方法（如線性迴歸程序）就可估計得到的簡單線性模型 $Y = A + BW$ 模式。

範例。 母群體可依時間預測嗎？散佈圖顯示母群體及時間之間似乎有極大關係，但因關係是非線性的，所以它須要求「非線性迴歸」程序的特殊估計方法。藉由建立合適的方程式（如 Logistic 母群體成長模式），我們可得到模型的良好估計，以便針對未實際測量的時間進行母群體的預測。

統計量。 對於每一個反覆運算：統計量有參數估計值及殘差平方和。對於每一個模型：統計量則有迴歸、殘差的平方和、未修正與修正的總和，以及參數估計值、漸進線標準誤及參數估計值的漸進線相關性矩陣。

附註：受限制的非線性迴歸係使用 NPSOL[®]（由 Gill、Murray、Saunders 及 Wright 提出）所提議並執行的演算法，來估計模型參數。

非線性迴歸的資料考量

資料。 因變數和自變數應該都是數值變數。例如宗教、主修課程、或居住地區等類別變數，都必須重新編碼成二元（虛擬）變數，或是其他類型的對比變數。

假設。 只有在您指定了精確說明因變數及自變數之間關係的函數時，結果才是有效的。此外，選擇適合的開始值也是很重要的。即使您已指定正確的模式函數形式，如果您使用了不適合的開始值，您的模式可能會無法收斂，或者，您可能會得到局部最佳解，而非絕對最佳解。

相關程序。 許多起初為非線性的模式，可在轉換成線性模型後，使用「線性迴歸」程序來進行分析。如果您不確定適當模式為何，請使用「曲線估計」程序來協助確認資料中的有用函數關係。

取得非線性迴歸分析

1. 在功能表上，選擇：

分析 > 迴歸 > 非線性...

2. 從您的作用中資料集的變數清單中選擇數值因變數。

3. 若要建置模型表示式，請在「模型」欄位中輸入表示式，或是將元件（變數、參數或函數）貼入該欄位。

4. 按一下**參數**以識別模式中的參數。

分段的模式（在其網域不同部分採取不同形式的模式）必須使用單一模式陳述式中的條件邏輯來指定。

條件邏輯（非線性迴歸）

您可以使用條件邏輯來指定分段的模型。若要在模式表示式或損失函數中使用條件邏輯，請您排成一列項目級數，每一項對應一個條件。而每一項均由邏輯性表示式（在括號內）乘以當邏輯性表示式為真時產生的表示式所組成。

例如，考慮一個當 $X \leq 0$ 時等於 0、當 $0 < X < 1$ 時等於 X 及當 $X \geq 1$ 時等於 1 的分段模式。對此的表示式就是：

$(X \leq 0) * 0 + (X > 0 \ \& \ X < 1) * X + (X \geq 1) * 1$ 。

括弧中的邏輯表示式全部評估成 1 (true) 或 0 (false)。因此：

當 $X \leq 0$ 時，上述表示式會產生 $1 * 0 + 0 * X + 0 * 1 = 0$ 。

當 $0 < X < 1$ 時，表示式會產生 $0 * 0 + 1 * X + 0 * 1 = X$ 。

當 $X \geq 1$ 時，表示式會降為 $0 * 0 + 0 * X + 1 * 1 = 1$ 。

亦可藉置換不同的邏輯表示式及結果表示式，輕鬆地建立更多的複雜範例。請注意，雙不等式（如 $0 < X < 1$ 必須寫成複合表示式（如 $X > 0 \ \& \ X < 1$ ）。

可在邏輯表示式內使用字串變數：

$(\text{city} = \text{'New York'}) * \text{costliv} + (\text{city} = \text{'Des Moines'}) * 0.59 * \text{costliv}$

這樣會為 New York 市民產生一個表示式（變數 *costliv* 的數值），並為 Des Moines 居民產生另一個表示式（該項數值的 59%）。字串常數必須包含於雙引號 (") 或單引號 (') 內，如此處所示。

非線性迴歸的參數

「參數」是「非線性迴歸」程序所要估計模型的一部分。「參數」可以是可加性的常數、可乘的係數、指數或用於求函數值的數值。您所定義的所有參數均將出現（含他們的原始值）在主要對話框的「參數」清單中。

名稱。 您必須為每個參數指定名稱，這個名稱必須是有效的變數名稱，而且必須是主要對話框中模式表示式所用的名稱。

開始值。 可讓您為參數指定起始值，以儘可能更接近期望的最終解。不適當的開始值可能導致無法收斂，或是導致收斂至局部解（非絕對解）或完全不可能的解。

使用前一次分析結果的開始值。 如果您已從此對話框執行過非線性迴歸，就可以選擇此選項，從他們前一次執行時的數值來取得原始值。這樣可讓您在演算法收斂遲滯時仍繼續搜尋（原始的開始值仍將出現在主要對話框的「參數」清單中）。

附註：這個選擇會在您的剩餘作業過程中，存留於對話框內如果您變更了模式，請確定取消選取這個選項。

非線性迴歸的常見模型

下表提供了許多已發表的非線性迴歸模型的模式語法範例。隨機選取的模式不太可能完美適配您的資料。合適的參數開始值是必要的，而且有些模式會為收斂而要求限制式。

表 1. 模式語法範例

名稱(N)	模式表示式
漸進線迴歸	$b1 + b2 * \exp(b3 * x)$
漸進線迴歸	$b1 - (b2 * (b3 ** x))$
密度(D)	$(b1 + b2 * x) ** (-1 / b3)$
Gauss	$b1 * (1 - b3 * \exp(-b2 * x ** 2))$
Gompertz	$b1 * \exp(-b2 * \exp(-b3 * x))$
Johnson-Schumacher	$b1 * \exp(-b2 / (x + b3))$
對數修改	$(b1 + b3 * x) ** b2$

表 1. 模式語法範例 (繼續)

名稱(N)	模式表示式
對數 Logistic	$b1 - \ln(1 + b2 * \exp(-b3 * x))$
逐漸縮小返回的 Metcherlich 法	$b1 + b2 * \exp(-b3 * x)$
Michaelis Menten	$b1 * x / (x + b2)$
Morgan-Mercer-Florin	$(b1 * b2 + b3 * x ** b4) / (b2 + x ** b4)$
Peal-Reed	$b1 / (1 + b2 * \exp(-(b3 * x + b4 * x **2 + b5 * x **3)))$
三次方比	$(b1 + b2 * x + b3 * x **2 + b4 * x **3) / (b5 * x **3)$
二次方比	$(b1 + b2 * x + b3 * x **2) / (b4 * x **2)$
Richards	$b1 / ((1 + b3 * \exp(-b2 * x)) ** (1 / b4))$
Verhulst	$b1 / (1 + b3 * \exp(-b2 * x))$
Von Bertalanffy	$(b1 ** (1 - b4) - b2 * \exp(-b3 * x)) ** (1 / (1 - b4))$
Weibull	$b1 - b2 * \exp(-b3 * x ** b4)$
產生密度	$(b1 + b2 * x + b3 * x ** 2) ** (-1)$

非線性迴歸的損失函數

非線性迴歸中的「損失函數」是依演算法進行最小化的函數。請選取**殘差平方和**來對殘差平方和進行最小化，或選取**使用者定義的損失函數**來對不同的函數進行最小化。

如果您選取了**使用者定義的損失函數**，就必須定義損失函數，而其和（遍及所有觀察值）須依參數值之選擇來進行最小化。

- 大多數損失函數均涉及代表殘差的 *RESID_* 特殊變數（也可以明確地將預設的「殘差平方和」損失函數輸入成 *RESID_**2*。）如果您需要在損失函數中使用預測值，那就等於將因變數減去殘差。
- 有可能以條件邏輯來指定條件損失函數。

您可以選擇在「使用者定義的損失函數」欄位中輸入表示式，或將表示式元件貼入該欄位。字串常數必須包含於雙引號 (") 或單引號 (') 內，而數值常數必須輸入成美國格式，以小圓點 (.) 代表小數點符號。

非線性迴歸的參數限制

限制是在疊代搜尋解答期間，可允許參數值的限制。因為線性表示式係於採取步驟前即先進行求值，所以您可使用線性限制來防止可能導致溢出的步驟。但非線性表示式則係於採取步驟後才進行求值。

每一個方程式或不等式均要求下列元素：

- 在模式中至少包含一個參數的表示式。直接輸入表示式，或是使用可讓您將數字、運算元或括號貼入表示式的小鍵盤。您可以選擇與表示式其餘部分一起輸入所要求的參數，或是從左邊的「參數」清單中選貼過來。但您不能在限制中使用二元變數。
- 三項邏輯運算元 \leq 、 $=$ 或 \geq 的其中之一。
- 以邏輯運算元來與表示式作比較的數值常數。輸入常數，但數值常數必須輸入成美國格式，以小圓點 (.) 代表小數點符號。

非線性迴歸的儲存新變數

您可以將一些新變數儲存到作用中的資料檔，可用的選項包括「預測值」、「殘差」、「導數」及「損失函數值」。這些變數可用於隨後之分析，以檢定模式之適配與否，或確認問題觀察值。

- 殘差. 儲存變數名稱 resid 的殘差。
- 預測值. 儲存變數名稱 pred_ 的預測值。
- 導數(D). 對於每個模型參數，將會各儲存一個導數。建立導數名稱時，會在參數名稱的前六個字元之前加上 'd.'。
- 損失函數值 (Loss Function Values). 若您指定專屬的損失函數，則可使用此選項。名稱爲 loss_ is 的變數會指派給損失函數值。

非線性迴歸選項

「選項」可讓您控制非線性迴歸分析的不同方面：

重複取樣估計 (Bootstrap Estimates). 使用原始資料集中的重複樣本，來估計統計量標準誤的一種方法。其作法是採取樣後 (放回) 的方式，來取得大小與原始資料集相同的眾多樣本。並會針對這些所有樣本，來估計非線性方程式。接著，重複取樣估計的標準差會計算每個參數估計值的標準誤。而來源資料中的參數值會當作成各個重複取樣樣本的起始值使用。這會要求使用循序二次程式設計演算法。

估計方法。 如有可能，可讓您選取估計方法。（在此項或其他對話框中的某些選擇須用到循序二次程式設計演算法）。可用的選項包含「循序二次程式設計」及「Levenberg-Marquardt」。

- 循序二次程式設計. 此方法適用於受限制及無條件限制模式。若您指定受限制模式、使用者定義的損失函數或重複取樣，會自動使用循序二次程式設計。您可以為最大反覆運算及步驟限制輸入新值，同時可以在下拉清單中變更最佳容差度、函數精準度和無限步驟大小的選取內容。
- Levenberg-Marquardt 方法. 此為無條件限制模型的預設演算法。若您指定受限制模型、使用者定義的損失函數或重複取樣，則無法使用 Levenberg-Marquardt 方法。您可輸入「最大」反覆運算的新值，並可變更「平方和」收斂及「參數」收斂下拉清單中的選擇。

說明非線性迴歸的結果

非線性迴歸問題通常會面臨下列計算上的困難：

- 影響收斂的參數開始值之選擇。試著選擇合理的開始值，而且如有可能的話，儘量接近期望的最終解。
- 有時在特定問題上，某演算法執行結果比其他演算法更佳。在「選項」對話框中，請選取其他可用的演算法（如果您指定了損失函數或某些限制類型，那就不能使用拉凡格氏 (Levenberg-Marquardt) 演算法）。
- 當疊代僅因最大疊代次數到達而停止時，「最終」模式或許不是最好的解答。請在「參數」對話框中選取**使用前一次分析結果的開始值**，以繼續疊代；或是用更好的作法，選擇不同的開始值。
- 要求很大資料值乘冪的模式可能會導致溢出或潛流（數字太大或太小而使電腦無法表示）。有時您可藉由選擇合適的開始值或輸入參數的限制，來避免這些問題。

NLR 指令的其他功能

指令語法語言也可以讓您：

- 命名檔案，以便從其讀取參數估計值的原始值。
- 指定一個以上的模型陳述式和損失函數，這可使指定分段的模型更為容易。
- 提供您自己的導數，而非使用程式計算出來的。
- 指定要產生的重複取樣樣本數。

- 指定其他的疊代條件，包含設定導數檢查的關鍵值，以及定義殘差與微分之間相關的收斂條件。

CNLR（受限制的非線性迴歸）指令的其他條件可讓您：

- 指定主要疊代中可容許的次要疊代最大次數。
- 設定導數檢查的關鍵值。
- 設定步驟限制。
- 指定失敗的容差度，以決定原始值是否落在指定的界限內。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

第 6 章 加權估計

在標準線性迴歸模型中，假設研究之母群體中的變異數為常數。當這項假設不成立時（例如，屬性高的觀察值，會比屬性低的觀察值，出現更多變化），使用普通最小平方方法 (OLS) 的線性迴歸，就無法提供您最佳化模式估計。如果可以從另一個變數預測變化差異的話，「權重估計」程序就可以使用加權最小平方方法 (WLS)，來計算線性迴歸模型的係數。這樣在決定迴歸係數時，比較精確的觀察值（也就是變化較少的觀察值），其加權值就會變得比較大。接下來，「權重估計」程序會檢定加權轉換的範圍，並指出資料最適合使用哪一種。

範例。 通貨膨脹和失業對股票價格的變動有何影響？由於股價較高的股票通常比股價低的股票更有變化性，普通最小平方方法無法提供您最佳化模式估計。而「權重估計」能讓您在計算線性模型時，考慮到股份價值，對價格變化的影響。

統計量。 它可以檢定加權來源變數之各冪次的對數概似值、複相關係數 R 、 R -平方、調整過的 R -平方、WLS 模型的 ANOVA 摘要表、未標準化和標準化的參數估計值，以及 WLS 模型的對數概似值。

權重估計法的資料考量

資料。 因變數和自變數應該都是數值變數。例如宗教、主修課程、或居住地區等類別變數，都必須重新編碼成二元（虛擬）變數，或是其他類型的對比變數。而加權變數也應該是數值變數，並且與因變數中的變化有關係。

假設。 對自變數的每個值而言，因變數的分配必須是常態的。因變數和每個自變數之間的關係，應該是線性的，而且所有觀察值應該互不相關。因變數的變異性，可以橫跨數個自變數階層，但是，其差異性必須能夠透過加權變數，加以估計。

相關程序。 「探索」程序，可以用來檢視您的資料。「探索」可提供變異數的常態性和同質性檢定，以及用圖形顯示資料。如果因變數在不同的自變數階層中，它的變異性都差不多時，就可以使用「線性迴歸」程序。如果您的資料違反某項假設（如常態性）的話，請先試著轉換資料。如果您的資料並不是呈線性相關，而且轉換無效的話，請使用「曲線估計」程序中的其他模式。如果您的因變數為二分變數（例如，是否完成某項交易，或者某個項目是否有效），那麼請使用「羅吉斯迴歸」程序。如果您的因變數是需要設限的（例如手術之後的存活時間），請使用「進階統計量」選項中的「生命表」、「Kaplan-Meier」或「Cox 迴歸」。如果您的資料並不是獨立的（例如，您是在數種不同情況下，觀察同一個人），請使用「進階統計量」選項中的「重複測量」程序。

若要取得權重估計分析

1. 在功能表上，選擇：

分析 > 迴歸 > 加權估計...

2. 選取一個因變數。

3. 選取一個或多個自變數。

4. 選取一個 heteroscedasticity 來源的變數，來當作加權變數。

• **加權變數 (Weight Variable).** 資料是使用含有次方的變數倒數進行加權。迴歸方程式的計算可用於每個指定的冪次值範圍，也表示最大化對數概似函數的冪次。

- **幂次範圍 (Power Range)**. 它與加權變數一起使用來計算加權。在幂次範圍內，將有幾個適合的迴歸方程式，每個數值都有一個適合的方程式。在幂次範圍檢定方框和整個文字框中輸入的數值必須介於 -6.5 和 7.5 (含) 之間。幂次值範圍涵蓋由低至高的數值，增量程度由指定數值決定。幂次範圍內的值總數最多限制為 150 個。

權重估計的選項

在此對話框中，您可以指定權重估計分析的選項：

將最佳權數存成新變數。把加權變數加入到作用中的檔案。這個變數稱為 *WGT_n*，其中 *n* 是一個選用數字，用來指定變數的專屬名稱。

顯示 ANOVA 與估計。讓您控制統計量在輸出上的顯示方式。您可以使用的選項包括：「最佳幂次」和「各幂次值」。

WLS 指令的其他功能

指令語法語言也可以讓您：

- 提供幂次的單一值。
- 指定幂次值的清單，或者將某個範圍裡面的數值，跟幂次值的清單合併起來。

如需完整的語法資訊，請參閱《指令語法參考手冊》。

第 7 章 二階最小平方迴歸

標準的線性迴歸模型會假設因變數的錯誤和自變數沒有關係。若不是這種情況（例如，當變數之間的關係是朝兩個方向作用的），使用一般的最小平方法 (OLS) 的線性迴歸將不再提供最佳模式預估。二階最小平方迴歸使用和誤差項無關的工具變數來計算有問題的預測數的估計值（第一階段），然後使用這些計算後的值來預估因變數的線性迴歸模型（第二階段）。因為計算後的值是根據和誤差無關的變數，所以二階模型的結果會是最佳的。

範例。 商品需求性與其價格和消費者的收入有關嗎？此模式中的難題在於價格和需求會彼此互相影響。也就是價格會影響需求，而需求也會影響價格。二階最小平方迴歸模型可能會使用消費者的收入和落差的價格來計算和需求的測量誤差無關的價格取代值。這個取代值是替代之後會被測量、且為最初指定模式中的價格。

統計量。 對於每一個模型：標準和非標準的迴歸係數、多重 R 、 R^2 、調整的 R^2 、預估的標準誤、變數分析表、預測值、以及殘差。同時還有每一個迴歸係數 95% 的信賴區間，以及參數預估的相關和共變異數矩陣。

二階最小平方迴歸資料考量

資料。 因變數和自變數應該都是數值變數。例如宗教、主修課程、或居住地區等類別變數，都必須重新編碼成二元（虛擬）變數，或是其他類型的對比變數。內生說明變數應該是可量化的（不是可分類的）。

假設。 對自變數的每個值而言，因變數的分配必須是常態的。對所有自變數數值而言，因變數分配的變異性，應該都是常數。因變數和每一個自變數的關係應該是線性的。

相關程序。 如果您相信您的預測變數沒有一個與因變數的錯誤有關，那麼您就可以使用線性迴歸程序。假如資料跟假設之一（例如常態性或常數變異數）是矛盾的話，請試著轉換資料。如果您的資料並不是呈線性相關，而且轉換無效的話，請使用「曲線估計」程序中的其他模式。如果你的因變數可以分成兩個，例如是否某一特定的銷售已完成，則可以使用羅吉斯迴歸程序。如果您的資料並不是獨立的（例如，您是在數種不同情況下，觀察同一個人），請使用「進階統計量」選項中的「重複測量」程序。

取得二階最小平方迴歸分析

1. 在功能表上，選擇：

分析 > 迴歸 > 2 階段最小平方數...

2. 選取一個因變數。

3. 選取一個或多個說明（預測值）變數。

4. 選取一個或多個工具變數。

- **工具 (Instrumental).** 在二階最小平方分析的第一階段中，這些變數會用於計算內生變數的預測值。在「說明」與「工具」清單框中可能會顯示相同的變數。工具變數個數至少必須等於說明變數個數。如果列出的所有說明變數與工具變數皆相同，則其結果與「線性迴歸」程序結果相同。

未指定為工具變數的說明變數將視為內生變數。一般說來，「說明」清單中的所有 Exogenous 變數也都指定為工具變數。

二階最小平方迴歸選項

您可以選擇下列選項進行分析：

儲存新變數。 讓您增加新變數到作用中的檔案。可用的選項有「預測值」和「殘差」。

顯示參數的共變異數。 讓您列印參數估計值的共變異數矩陣。

2SLS 指令的其他功能

指令語法語言也讓您可以同時估計多重等式。如需完整的語法資訊，請參閱《指令語法參考手冊》。

第 8 章 類別變數編碼方法

在許多的程序中，您可以要求用一組對比變數自動取代類別自變數，然後再當作區塊輸入方程式或從中移除。您可以指定該組對比變數的編碼方式，通常是在 `CONTRAST` 次指令中指定。本篇附錄說明並列舉 `CONTRAST` 中不同的對比類型要求實際上如何運作。

偏差

與總平均數之離差。 以矩陣的觀點來看，這些對比具有下述形式：

```
mean ( 1/k  1/k  ...  1/k  1/k)
df(1) (1-1/k -1/k  ... -1/k -1/k)
df(2) (-1/k  1-1/k ... -1/k -1/k)
      .
      .
df(k-1) (-1/k -1/k  ... 1-1/k -1/k)
```

其中 k 為自變數的類別個數，且預設情況下會略過最後一個類別。例如，具有三個類別的自變數其離差對比如下所示：

```
( 1/3  1/3  1/3)
( 2/3 -1/3 -1/3)
(-1/3  2/3 -1/3)
```

若要略過最後一個以外的類別，請在 `DEVIATION` 關鍵字後面，用括弧指定要略過的類別編號。例如，以下的次指令取得第一個和第三個類別的離差並略過第二個類別：

```
/CONTRAST(FACTOR)=DEVIATION(2)
```

假設 *factor* 有三個類別。最後結果的對比矩陣將會是

```
( 1/3  1/3  1/3)
( 2/3 -1/3 -1/3)
(-1/3 -1/3  2/3)
```

簡單

簡單對比。 將因素的每個層級與最後一個做比較。一般的矩陣形式為

```
mean (1/k  1/k  ...  1/k  1/k)
df(1) ( 1  0  ...  0  -1)
df(2) ( 0  1  ...  0  -1)
      .
      .
df(k-1) ( 0  0  ...  1  -1)
```

其中 k 為自變數的類別個數。例如，具有四個類別的自變數其簡單對比如下所示：

```
(1/4  1/4  1/4  1/4)
( 1  0  0  -1)
( 0  1  0  -1)
( 0  0  1  -1)
```

若要使用另一個類別而非最後一個來作為參考類別，請在 `SIMPLE` 關鍵字後面用括弧指定參考類別的順序編號，但未必是和該類別相關的。例如，下述 `CONTRAST` 次指令取得已略過第二個類別的對比矩陣：

```
/CONTRAST(FACTOR) = SIMPLE(2)
```

假設 *factor* 有四個類別。最後結果的對比矩陣將會是

```
(1/4  1/4  1/4  1/4)
( 1  -1  0  0)
( 0  -1  1  0)
( 0  -1  0  1)
```

Helmert

赫爾莫特 (Helmert) 對比。 將自變數的類別與後續類別的平均數做比較。一般的矩陣形式為

```
mean (1/k      1/k      ...      1/k      1/k      1/k)
df(1) ( 1 -1/(k-1) ... -1/(k-1) -1/(k-1) -1/(k-1))
df(2) ( 0      1      ... -1/(k-2) -1/(k-2) -1/(k-2))
      .
      .
df(k-2) ( 0      0      ...      1      -1/2      -1/2)
df(k-1) ( 0      0      ...      0      1      -1)
```

其中 k 為自變數的類別個數。例如，具有四個類別的自變數其赫爾莫特 (Helmert) 對比矩陣的形式如下：

```
(1/4  1/4  1/4  1/4)
( 1 -1/3 -1/3 -1/3)
( 0      1 -1/2 -1/2)
( 0      0      1 -1)
```

差異

差分或反赫爾莫特 (Helmert) 對比。 將自變數的類別與變數之先前類別的平均數做比較。一般的矩陣形式為

```
mean (      1/k      1/k      1/k ... 1/k)
df(1) (      -1      1      0 ... 0)
df(2) (      -1/2     -1/2      1 ... 0)
      .
      .
df(k-1) (-1/(k-1) -1/(k-1) -1/(k-1) ... 1)
```

其中 k 為自變數的類別個數。例如，具有四個類別的自變數其差異對比如下所示：

```
( 1/4  1/4  1/4  1/4)
( -1      1      0      0)
(-1/2 -1/2  1      0)
(-1/3 -1/3 -1/3  1)
```

多項式

正交多項式對比。 第一自由度包含所有類別的線性效應項；第二自由度包含二次效應項；第三自由度包含三次效應項；依此類推到更高階的效應項。

您可以指定由給定類別變數所測得的處理方式層級之間間距。間距相等 (如果您略過矩陣則此為預設值) 可以指定為從 1 到 k 的連續整數，其中 k 為類別個數。如果變數 *drug* 有三個類別，則次指令

```
/CONTRAST(DRUG)=POLYNOMIAL
```

相當於

```
/CONTRAST(DRUG)=POLYNOMIAL(1,2,3)
```

然而間距相等並非永遠必要。例如，假設 *drug* 代表給予三個組別的不同藥量。如果給予第二組別的管制藥量為第一組別的兩倍，而第三組別為第一組別的四倍，則處理方式類別為間距相等，且這種狀況下的合適矩陣由連續整數構成：

```
/CONTRAST(DRUG)=POLYNOMIAL(1,2,3)
```

但是，如果給予第二組別的管制藥量為第一組別的四倍，而第三組別為第一組別的七倍，則合適的矩陣為

```
/CONTRAST(DRUG)=POLYNOMIAL(1,4,7)
```

在上述各種情況下，對比規格的結果為，*drug* 的第一自由度包含藥量層級的線性效應項，且第二自由度包含二次效應項。

多項式對比在測試趨勢及研究回應值表面的本質等方面非常有用。您也可以使用多項式對比來進行非線性曲線填入，例如曲線線性迴歸。

重複

比較自變數的相鄰層級。一般的矩陣形式為

```
mean (1/k 1/k 1/k ... 1/k 1/k)
df(1) ( 1 -1 0 ... 0 0)
df(2) ( 0 1 -1 ... 0 0)
      :
      :
df(k-1) ( 0 0 0 ... 1 -1)
```

其中 k 為自變數的類別個數。例如，具有四個類別的自變數其重複對比如下所示：

```
(1/4 1/4 1/4 1/4)
( 1 -1 0 0)
( 0 1 -1 0)
( 0 0 1 -1)
```

這些對比在輪廓分析方面及需要差異分數時特別有用。

特殊

使用者定義的對比。 允許特殊對比的項目以平方矩陣的形式來呈現，而矩陣的列和行個數與給定自變數的類別個數相同。對於 MANOVA 與 LOGLINEAR 而言，輸入的第一列永遠為平均數或對比效應項並代表一組加權值，這些加權值則指示如何取得其他自變數對於給定變數的平均值 (如果有的話)。一般而言，這種對比為向量。

包含特殊對比之矩陣的其餘各列表示變數的類別之間的比較。正交對比通常最為有用。正交對比在統計上互相獨立且並非多餘。若符合下列則為正交對比：

- 對於每列而言，對比係數總和為 0。
- 非聯合列的所有成對其對應係數的乘積總和亦為 0。

例如，假設該處理有四個層級，且您想要讓處理方式的各層級彼此互相比較。合適的特殊對比為

```
(1 1 1 1) 平均值計算的加權值
(3 -1 -1 -1) 將第一個與第二個至第四個比較
(0 2 -1 -1) 將第二個與第三個和第四個比較
(0 0 1 -1) 將第三個與第四個比較
```

您藉由 MANOVA、LOGISTIC REGRESSION 和 COXREG 的下列 CONTRAST 次指令所指定的項目為：

```
/CONTRAST(TREATMNT)=SPECIAL( 1 1 1 1 3 -1 -1 -1 0 2 -1 -1 0 0 1 -1 )
```

對於 LOGLINEAR 而言，您必須指定：

```
/CONTRAST(TREATMNT)=BASIS SPECIAL( 1 1 1 1 3 -1 -1 -1 0 2 -1 -1 0 0 1 -1 )
```

除了平均值列外其餘各列總和皆為 0。非聯合列的每一成對其乘積總和亦為 0：

```
列 2 和 3: (3)(0) + (-1)(2) + (-1)(-1) + (-1)(-1) = 0
列 2 和 4: (3)(0) + (-1)(0) + (-1)(1) + (-1)(-1) = 0
列 3 和 4: (0)(0) + (2)(0) + (-1)(1) + (-1)(-1) = 0
```

特殊對比不需為正交。但彼此之間不得為線性組合。如果是的話，程序會報告線性相依並停止處理。Helmert、差異、以及多項式對比都是正交對比。

指示燈

指標變數編碼。 亦稱為虛擬編碼，且無法用於 LOGLINEAR 或 MANOVA。新變數的數目編碼為 $k - 1$ 。對於所有 $k-1$ 變數，其參考類別中的觀察值編碼為 0。所有指標變數第 i 個類別中的觀察值編碼為 0，但第 i 個類別例外，此類別中的觀察值編碼為 1。

注意事項

本資訊適用於美國提供的產品與服務。

IBM 可能並未在其他國家提供在本文件中討論到的產品、服務或功能。有關目前在 貴地區可供使用的產品與服務相關資訊，請洽您當地的 IBM 服務代表。對於 IBM 產品、程式或服務的任何參考，目的並不是要陳述或暗示只能使用 IBM 產品、程式或服務。任何功能相等且未侵犯 IBM 智慧財產權的產品、程式或服務皆可使用。但是，評估及確認任何非 IBM 產品、程式或服務的操作之責任應由使用者承擔。

IBM 可能有一些擁有專利或專利申請中的項目包含本文件所描述的內容。本文件的提供並不表示授與您對於這些專利的權利。您可以將書面的授權查詢寄至：

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

對於與雙位元組 (DBCS) 資訊相關的授權查詢，請與貴國的 IBM 智慧財產部門聯絡，或將查詢郵寄至：

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

下列條款不適用於英國，或其他任何當地法律規定與下列條款不一致的國家：INTERNATIONAL BUSINESS MACHINES CORPORATION 聲明係以「現狀」提供，沒有任何保固；不作任何明示或默示的保證，包括但不限於不侵權、適銷性或適合某一特定用途之保證。某些州不允許特定交易中明示或默示的保固聲明，因此，此聲明或許對您不適用。

此資訊內容可能包含技術失準或排版印刷錯誤。此處資訊會定期變更，這些變更將會納入新版的聲明中。IBM 可能會隨時改善和 / 或變更此聲明中所述的產品和 / 或程式，恕不另行通知。

本資訊中任何對非 IBM 網站的敘述僅供參考，IBM 對該網站並不提供任何保證。該「網站」的內容並非此 IBM 產品的部分內容，使用該「網站」需自行承擔風險。

IBM 可能會以任何其認為適當的方式使用或散佈您提供的任何資訊，無需對您負責。

意欲針對達成以下目的而擁有本程式相關資訊之程式被授權人：(i) 在獨立建立的程式與其他程式 (包括本程式) 之間交換資訊及 (ii) 共用已交換的資訊，應聯絡：

IBM Software Group
ATTN: Licensing
200 W. Madison St.
Chicago, IL; 60606
U.S.A.

在適當條款與條件之下，包括某些情況下 (支付費用)，或可使用此類資訊。

在本文件中描述的授權程式及其適用之所有授權材料皆由 IBM 在與我方簽訂之 IBM 客戶合約、IBM 國際程式授權合約或任何相等等效力合約中提供。

此處所包含的任何效能資料皆於受控制的環境中取得。因此，其他作業環境中所取得的結果可能有顯著的不同。某些測量可能是於開發階段的系統中執行，因此無法保證在一般所用的系統中進行測量結果會相同。此外，某些測量可能已經透過外插法進行預測。實際測量結果可能不同。此文件的使用者應確認資料是否適用於特定環境中。

有關非 IBM 產品的資訊來自於那些產品的供應商、其公佈內容或其他可公開取得的來源。IBM 尚未測試那些產品，且無法確認效能準確度、相容性或任何其他與非 IBM 產品相關的理賠。有關非 IBM 產品的性能問題應直接洽詢該產品供應商。

關於 IBM 未來方針或意圖的所有聲明僅代表目標或目的，得依規定未另行通知即變更或撤銷。

此資訊包含用於日常企業運作的資料和報表範例。為了儘可能提供完整說明，範例中包含了人名、公司名稱、品牌名稱和產品名稱。所有的名稱皆為虛構，使用的名稱或地址和實際的企業如有雷同純屬巧合。

這些範例程式或任何衍生成果的每份複本或任何部分，都必須依照下列方式併入著作權聲明：

此資訊包含用於日常企業運作的資料和報表範例。為了儘可能提供完整說明，範例中包含了人名、公司名稱、品牌名稱和產品名稱。所有的名稱皆為虛構，使用的名稱或地址和實際的企業如有雷同純屬巧合。

這些範例程式或任何衍生成果的每份複本或任何部分，都必須依照下列方式併入著作權聲明：

© (貴客戶公司名稱) (年度) 。本程式之若干部分係衍生自 IBM 公司的範例程式。

© Copyright IBM Corp. _輸入一個以上年度_。All rights reserved.

商標

IBM、IBM 標誌及 ibm.com 是 International Business Machines Corp. 在世界許多管轄區註冊的商標或註冊商標。其他產品及服務名稱可能是 IBM 或其他公司的商標。IBM 商標的最新清單可在 Web 的『Copyright and trademark information』中找到，網址為 www.ibm.com/legal/copytrade.shtml。

Adobe、Adobe 標誌、PostScript 以及 PostScript 標誌為 Adobe Systems Incorporated 於美國和 / 或其他國家的註冊商標或商標。

Intel、Intel 標誌、Intel Inside、Intel Inside 標誌、Intel Centrino、Intel Centrino 標誌、Celeron、Intel Xeon、Intel SpeedStep、Itanium 和 Pentium 為 Intel Corporation 或其分公司於美國和其他國家的商標或註冊商標。

Linux 為 Linus Torvalds 於美國和 / 或其他國家的註冊商標。

Microsoft、Windows、Windows NT 和 Windows 標誌為 Microsoft Corporation 於美國和 / 或其他國家的商標。

UNIX 為 The Open Group 於美國和其他國家的註冊商標。

Java 和所有以 Java 為基礎的商標及標誌是 Oracle 及 (或) 其子公司的商標或註冊商標。

索引

索引順序以中文字，英文字，及特殊符號之次序排列。

〔二劃〕

- 二元羅吉斯迴歸分析 1
- 二次方比模型
 - 有限制的迴歸 14
- 二階最小平方迴歸 21
 - 工具變數 21
 - 指令的其他功能 22
 - 參數的共變異數 22
 - 統計量 21
 - 範例 21
 - 儲存新變數 22

〔三劃〕

- 三次方比模型
 - 有限制的迴歸 14

〔四劃〕

- 分組
 - 在多項式羅吉斯迴歸中 9
- 分散尺度數值
 - 在多項式羅吉斯迴歸中 9
- 分類
 - 在多項式羅吉斯迴歸中 7
- 分類表
 - 在多項式羅吉斯迴歸中 8

〔五劃〕

- 主效應模型
 - 在多項式羅吉斯迴歸中 7
- 加權估計 19
 - 指令的其他功能 20
 - 將最佳權數存成新變數。 20
 - 統計量 19
 - 對數概似 19
 - 範例 19
 - 疊代歷程 20
 - 顯示 ANOVA 與估計 20
- 半階
 - 在多項式羅吉斯迴歸中 9
- 平行檢定
 - 機率值分析中 12
- 皮爾森 (Pearson) 卡方
 - 估計分散尺度數值 9

- 皮爾森 (Pearson) 卡方 (繼續)
 - 適合度 8

〔六劃〕

- 共變異數矩陣
 - 在多項式羅吉斯迴歸中 8
- 共變量
 - 在羅吉斯迴歸中 4
- 向前選取法
 - 在羅吉斯迴歸中 4
- 向後消去法
 - 在羅吉斯迴歸中 4
- 多項式羅吉斯迴歸 7, 8
 - 指令的其他功能 10
 - 參考類別 8
 - 條件 9
 - 統計量 8
 - 匯出模型資訊 10
 - 模型 7
 - 儲存 10
- 字串共變量
 - 在羅吉斯迴歸中 4
- 收斂準則
 - 在多項式羅吉斯迴歸中 9
- 有限制的迴歸
 - 有限制的迴歸 15
- 自訂模型
 - 在多項式羅吉斯迴歸中 7

〔七劃〕

- 含有觀察值 0 的儲存格
 - 在多項式羅吉斯迴歸中 9
- 完全因素模型
 - 在多項式羅吉斯迴歸中 7

〔八劃〕

- 非線性迴歸 13
 - 分段的模型 13
 - 估計方法 16
 - 拉凡格氏 (Levenberg-Marquardt) 演算法 16
 - 指令的其他功能 16
 - 重複取樣估計 16
 - 參數 14
 - 參數限制 15
 - 常見的非線性模型 14
 - 條件邏輯 13

- 非線性迴歸 (繼續)
 - 統計量 13
 - 循序二次程式設計 16
 - 殘差 16
 - 結果說明 16
 - 開始值 14
 - 損失函數 15
 - 預測的值 16
 - 範例 13
 - 導數 16
 - 儲存新的變數 16
- 非線性模型
 - 有限制的迴歸 14
- 信賴區間
 - 在多項式羅吉斯迴歸中 8

〔九劃〕

- 相對中位數
 - 機率值分析中 12
- 相關性矩陣
 - 在多項式羅吉斯迴歸中 8

〔十劃〕

- 格機率表
 - 在多項式羅吉斯迴歸中 8

〔十一劃〕

- 參考類別
 - 在多項式羅吉斯迴歸中 8
- 參數估計值
 - 在多項式羅吉斯迴歸中 8
- 參數限制
 - 有限制的迴歸 15
- 密度模型
 - 有限制的迴歸 14
- 常數項
 - 常數項 5
- 產生密度模型
 - 有限制的迴歸 14
- 逐步選取
 - 在多項式羅吉斯迴歸中 7
 - 在羅吉斯迴歸中 4
- 逐漸縮小返回的 Metcherlich 法
 - 有限制的迴歸 14

〔十二劃〕

單一性
在多項式羅吉斯迴歸中 9

〔十三劃〕

概似比
估計分散尺度數值 9
適合度 8

〔十四劃〕

對比
在羅吉斯迴歸中 4
對數修改模型
有限制的迴歸 14
對數概似
在多項式羅吉斯迴歸中 8
在權重估計中 19
截距
包含或排除 7
槓桿值
在羅吉斯迴歸中 5
漸進迴歸
有限制的迴歸 14

〔十五劃〕

線性迴歸
二階最小平方迴歸 21
權重估計 19
適合度
在多項式羅吉斯迴歸中 8

〔十六劃〕

機率值分析
平行檢定 12
自然回應率 12
定義範圍 12
指令的其他功能 12
相對中位數 12
條件 12
統計量 11, 12
範例 11
疊代 12
Fiducial 信賴區間 12

〔十八劃〕

離差函數
估計分散尺度數值 9

〔十九劃〕

羅吉斯迴歸 3
二元 1
分類分割值 5
字串共變量 4
定義選擇規則 4
係數 3
指令的其他功能 6
常數項 5
統計量 3
統計與圖形 5
設定規則 4
逐步之機率。 5
殘差 5
預測的值 5
對比 4
槓桿值 5
範例 3
儲存新變數 5
類別共變量 4
疊代 5
變數選取方法 4
顯示選項 5
Hosmer-Lemeshow 適合度統計量 5
類別共變量 4

〔二十二劃〕

疊代
在多項式羅吉斯迴歸中 9
在羅吉斯迴歸中 5
機率值分析中 12
疊代歷程
在多項式羅吉斯迴歸中 9

C

Cook's D
在羅吉斯迴歸中 5
Cox 與 Snell R 平方
在多項式羅吉斯迴歸中 8

D

Delta
對含有觀察值 0 的儲存格之改正 9
DfBeta
在羅吉斯迴歸中 5

F

Fiducial 信賴區間
機率值分析中 12

G

Gauss 模型
有限制的迴歸 14
Gompertz 模型
有限制的迴歸 14

H

Hosmer-Lemeshow 適合度統計量
在羅吉斯迴歸中 5

J

Johnson-Schumacher 模型
有限制的迴歸 14

M

McFadden R 平方值
在多項式羅吉斯迴歸中 8
Michaelis Menten 模型
有限制的迴歸 14
Morgan-Mercer-Florin 模型
有限制的迴歸 14

N

Nagelkerke R 平方
在多項式羅吉斯迴歸中 8

P

Peal-Reed 模型
有限制的迴歸 14

R

Richards 模型
有限制的迴歸 14

V

Verhulst 模型
有限制的迴歸 14
Von Bertalanffy 模型
有限制的迴歸 14

W

Weibull 模型
有限制的迴歸 14



Printed in Taiwan