

*IBM SPSS Decision Trees 29*



## 附註

使用此資訊和支援的產品之前，請先閱讀第 17 頁的『注意事項』中的資訊。

## 產品資訊

除非新版中另有指示，否則此版本適用於 IBM® SPSS Statistics 29 版次 0，修訂 1 版及所有後續版次與修訂。

© Copyright International Business Machines Corporation .

---

# 目錄

<b>第 1 章 決策樹.....</b>	<b>1</b>
建立決策樹.....	1
選取類別.....	3
驗證.....	4
樹狀結構成長條件.....	4
選項.....	7
儲存模型資訊.....	10
輸出.....	10
樹狀編輯器.....	13
使用大型樹狀結構.....	14
控制樹狀結構中顯示的資訊.....	15
變更樹狀結構的顏色和字型.....	15
觀察值選擇和評分規則.....	15
<b>注意事項.....</b>	<b>17</b>
商標.....	18
<b>索引.....</b>	<b>19</b>



# 第 1 章 決策樹

下列「決策樹」功能包含在 SPSS Statistics Professional Edition 或「決策樹」選項中。

## 建立決策樹

「決策樹」程序會建立樹狀結構的分類模式。它會根據自變數（預測值）的值，將觀察值分組，或應變數（目標）的預測值。這個程序會提供用於解釋與確認分類分析的驗證工具。

這個程序可以用於：

**分段。** 識別有可能是特殊群組成員的人員。

**分層。** 將觀察值指定給其中一個類別，例如高風險、中風險與低風險群組。

**預測。** 建立規則並用這些規則來預測未來的事件，例如某人可能會借貸，或是汽車或房屋的潛在重新銷售值。

**資料縮減與變數篩檢。** 從一個大型的變數集中選取一個有用的預測值子集，用於建立一個正式的參數模式。

**交互作用識別。** 識別只與特定子群組有關的關係並在正式的參數模式中指定這些項目。

**類別合併與離散化連續的變數。** 使用遺失最少資訊的方式將組別預測值類別與連續變數重新編碼。

**範例。** 假設某家銀行打算根據信用申請人是否有合理的信用風險，來將這些申請人加以分類。根據各種因素，包括過去客戶的已知信用評等，您就可以建立一個模式來預測未來的客戶是否有可能進行借貸。

樹狀結構分析會提供一些引人注意的功能：

- 它可以讓您識別具有高風險或低風險的同質組別。
- 它可以更容易建構有關個別觀察值進行預測的規則。

資料考量

**資料。** 應變數和自變數可以是：

- 名義。當變數值代表沒有等級化的類別（例如，員工工作所在公司的部門）時，可以將變數視為名義。名義變數的範例包括區域、郵遞區號及宗教團體。
- 序數。當變數值代表具有等級化的類別時（例如，服務滿意度從非常不滿意到非常滿意的層級），則可將此變數視為序數。序數變數的範例包括代表滿意度或信賴程度的態度分數以及偏好等級分數。
- 尺度。當變數值代表具有有意義度量的已排序類別時，可以將變數視為尺度（連續），因此值之間的距離比較適當。尺度變數的範例包括以年份表示的年齡及以千元為單位的收入。

**頻次加權** 如果加權生效的話，則分數加權就會捨入為最接近的整數，所以，加權值少於 0.5 的觀察值就會被指定一個 0 的加權，進而從分析中被排除在外。

**假設。** 這個程序會假設已經將適當的測量層級指定給所有分析變數，而且某些功能會假設包含在分析中的應變數的所有值都已經定義數值標籤。

- **測量層級。** 測量層級會影響樹狀結構計算作業，因此所有變數都應該指定適當的測量層級。根據預設，數值變數是假設為尺度變數而字串變數則假設為名義變數，它們可能無法精確反映真正測量的層級。變數清單中各變數旁圖示會指明變數類型。

表 1: 測量層級圖示




圖示	測量層次
	縮放

表 1: 測量層級圖示 (繼續)

圖示	測量層次
	名義(M)
	序數(O)

您可以藉由在來源變數清單按一下滑鼠右鍵，從蹦現功能表選取測量層級，暫時變更變數的測量層級。

- **值標籤。** 此程序的對話框介面假設類別（名義、序數）應變數的「所有」非遺漏值均已定義數值標籤，或者「沒有一個」非遺漏值已定義。有些功能至少需要類別應變數的兩個遺漏值具有數值標籤，否則無法使用。如果至少兩個非遺漏值已定義數值標籤，當有任何觀察值具有其他無數值標籤的數值時，該觀察值會從分析中排除。

若要取得決策樹

1. 從功能表中選擇：

**分析 > 分類 > 樹狀結構...**

2. 選取應變數。
3. 選取一個或多個自變數。
4. 選取一個成長方法。

視需要而定，您可以：

- 變更來源清單中任何變數的測量層級。
- 強制自變數清單中的第一個變數進入模式中當作第一個分割變數。
- 選取定義觀察值影響樹狀結構成長過程之程度的影響變數。觀察值的影響變數數值較低，則影響力較小；反之則影響力較大。影響變數值必須為正數。
- 驗證樹狀結構。
- 自訂樹狀結構成長條件。
- 將終端節點數、預測值以及預測機率另存成變數。
- 以 XML (PMML) 格式儲存模式。

具有未知測量層級的欄位

若在資料集中出現一或多個未知的變數 (欄位) 測量層級，就會顯示「測量層級」警示。由於測量層級會影響此程序的結果計算，因此所有變數皆必須具有已定義的測量層級。

**掃描資料。** 讀取作用中資料集的資料，並且針對目前具有未知測量層級的任何欄位指派預設的測量層級。若為大型資料集，則讀取時可能需要一些時間。

**手動指派。** 開啟對話框，以列出具有未知測量層級所有欄位。您可以使用此對話框以指派測量層級給這些欄位。您可以在「資料編輯器」的「變數視圖」中指派測量層級。

由於測量層級是此程序的重要項目，因此您在所有欄位皆擁有已定義的測量層級之前，無法存取對話框來執行此程序。

變更測量層級

1. 在來源清單的變數上按一下滑鼠右鍵。
2. 從蹦現功能表選取測量層級。

這會暫時變更測量層級，以供在「決策樹」程序中使用。

成長方法。

可用的成長方法包括：

**CHAID.** 卡方自動互動偵測。CHAID 會在每個步驟中，選擇與應變數具有最強交互作用的自 (預測) 變數。如果每個預測變數的類別相對於應變數沒有明顯不同，則會合併這些類別。

耗盡的 **CHAID.** 對 CHAID 的修改，它會檢查每個預測值所有可能的分割。

**CRT.** 分類和迴歸樹狀結構。CRT 會盡量將資料分割成與應變數相關的同質資料區段。所有觀察值皆具有相同應變數值的終端節點，是「純」同質節點。

**QUEST.** 快速、公正、有效的統計樹狀結構。此方法不但計算快速，而且能避免如其他方法偏好有許多類別的預測變數。只有在名義應變數才能指定 QUEST。

每一個方法都有其優點與限制，包括：

功能	CHAID*	CRT	QUEST
以卡方分佈為基礎**	X		
代理自 (預測值) 變數		X	X
樹狀結構修正		X	X
多因素節點分割	X		
二元節點分割		X	X
影響變數	X	X	
事前機率		X	X
錯誤分類成本	X	X	X
快速計算	X		X

\*包括 Exhaustive CHAID。

\*\*QUEST 也會將卡方量數用於名義自變數。

## 選取類別

如果是類別 (名義、序數) 因變數，您可以：

- 控制要包含在分析中的類別。
- 識別相關的目標種類。

包含/排除類別

您可以將分析限制為因變數的特定類別。

- 在「排除」清單中之應變數值的觀察值不會包含在分析中。
- 如果是名義應變數，您也可以包含使用者遺漏的類別 (依預設，使用者遺漏的類別會顯示在「排除」清單中)。

目標類別

已選取的 (已核取的) 種類會被視為分析中主要相關的種類。例如，如果您主要是要識別最有可能借貸的個人，您可以選取信用評等種類「差」來當作目標種類。

- 沒有預設的目標類別。如果沒有選取任何類別，某些分類規則選項與獲利相關選項就無法使用。
- 如果選取多個類別，就會為每一個目標類別產生個別的獲利表與圖表。
- 將一個或多個類別指定為目標類別，對於樹狀模式、風險估計或錯誤分類結果並不會有任何影響。

類別及數值標籤

這個對話框需要應變數的已定義數值標籤。除非類別應變數至少有兩個值已經定義值標籤，否則無法使用這個對話框。主題，以取得更多資訊。

若要包含/排除類別並選取目標類別

1. 在主要的「決策樹」對話框中，選取具有兩個或多個數值標籤的類別（名義、序數）應變數。
2. 按一下「類別」。

## 驗證

驗證可以讓您評估樹狀結構對大型母群概化的程度。有兩種驗證方法可供使用：交叉驗證和分割樣本驗證。

### 交叉驗證

交叉驗證會將樣本分成子樣本數或**折疊**。接著會產生樹狀結構模型，然後從每個子樣本中排除資料。第一個樹狀結構是以第一個樣本折疊中以外的所有觀察值為基礎，第二個樹狀結構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。對於每一個樹狀結構，都會藉由將樹狀結構套用至在產生時被排除的子樣本來評估錯誤分類風險。

- 您最多可以指定 25 個樣本折疊。如果值越高，則從每一個樹狀結構模型中排除的觀察值數量也就越少。
- 交叉驗證會產生一個單一、最終的樹狀結構模型。最終樹狀結構的交叉驗證風險評估是計算所有樹狀結構風險的平均值。

### 分割樣本驗證

透過分割樣本驗證，就可以使用訓練樣本來產生模式，以及在保留樣本上測試模式。

- 您可以指定一個訓練樣本大小（以總樣本大小的百分比表示），以及指定一個會將樣本分割成訓練與測試樣本的變數。
- 如果您使用變數來定義訓練與測試樣本，則具有變數值 1 的觀察值就會被指定給訓練樣本，而所有其他觀察值則會被指定給測試樣本。變數不可以是應變數、加權變數、影響變數或強制自變數。
- 您可以同時顯示訓練樣本與測試樣本，或是僅顯示測試樣本。
- 在小型資料檔（含有少量觀察值的資料檔）上使用分割樣本驗證時，必須小心。小型的訓練樣本大小可能會產生品質不佳的模式，因為這些樣本大小在某些類別中可能沒有足夠的觀察值，所以無法適當地讓樹狀結構成長。

若要驗證決策樹

1. 在主「決策樹」對話框中，按一下**驗證**。
2. 選取「**交叉驗證**」或「**分割樣本驗證**」。

附註：兩種驗證方法都會隨意將觀察值指定給樣本組別。如果想要能夠在後續的分析中重製完全一樣的結果，則在第一次執行分析之前，應該要設定亂數種子（「轉換」功能表、「亂數產生器」），然後再為後續的分析將種子重設為該值。主題，以取得更多資訊。

## 樹狀結構成長條件

可用的成長條件是根據成長方法、應變數的測量層次或兩者的組合而定。

### 成長限制

「成長限制」標籤可以讓您限制樹狀結構中的層級數量，以及控制個上層節點與子節點的最小觀察值數量。

**最大樹狀結構深度。** 控制根節點底下成長的最大層次數量。「**自動**」設定會將樹狀結構限制在 CHAID 與 Exhaustive CHAID 方法之根節點底下的三個層級數量，以及 CRT 與 QUEST 方法的五個層級數量。

**最小觀察值個數。** 控制節點的觀察值個數下限。不符合這些準則的節點不會被分割。

- 增加下限值會產生具有較少節點的樹狀結構。
- 降低下限值會產生具有較多節點的樹狀結構。

對於觀察值較少的資料檔，母節點的預設值為 100 個觀察值而子節點的預設值為 50 個觀察值，這種情況有時候可能會造成樹狀結構中根節點底下沒有節點；在這種情況下，降低下限值可能會產生更有用的結果。

若要指定成長限制

1. 在主「決策樹狀結構」對話框中，按一下**準則**。



2. 按一下「**成長限制**」標籤。

## CHAID 準則

對於 CHAID 與 Exhaustive CHAID 方法，您可以控制：

**顯著性水準。** 您可以控制用於分割節點與合併種類的顯著值。對這兩個準則而言，預設的顯著性層次都是 0.05。

- 對於分割節點，該值必須大於 0 且小於 1。較低的值會產生具有較少節點的樹狀結構。
- 若為合併種類，值必須大於 0 且小於或等於 1。若要防止種類合併，請指定值 1。若為尺度自變數，這表示最終樹狀結構中變數的種類數目是指定的間隔數（預設值是 10）。如需相關資訊，請參閱主題第 5 頁的『CHAID 分析的尺度區間』。

**卡方統計量。** 如果是序數應變數，則用於決定節點分割與類別合併的卡方，就是使用概似比方法計算的。如果是序數應變數，您可以選擇方法：

- **Pearson。** 這個方法會提供更快速的計算，但是用在小型樣本時則必須小心。此為預設的方法。
- **概似比。** 這個方法比皮爾遜方法更為穩健，但是在計算時必須花費較多的時間。對於小型的樣本，這是較佳的方法。

**模式估計。** 對於名義與序數應變數，您可以指定：

- **最大疊代次數。** 預設值為 100。如果樹狀結構因為已經到達最大疊代次數而停止成長，您可以增加最大值的數字，或是變更一個或兩個控制樹狀結構成長的條件。
- **資料格期望頻次中的最少變更。** 該值必須大於 0 小於 1。預設值為 0.05。較低的值會產生具有較少節點的樹狀結構。

**使用 Bonferroni 方法調整顯著值。** 對於多重比較而言，用於合併與分割準則的顯著性值是透過使用 Bonferroni 方法調整的。此為預設值。

**允許在節點中重新分割已合併的類別。** 除非您明確地防止類別進行合併，否則程序會嘗試一起合併自 (預測值) 變數類別以產生會描述模型之最簡單的樹狀結構。這個選項允許該程序重新分割已經合併的類別（如果這樣做可以提供更佳解答的話）。

若要指定 CHAID 條件

1. 在主「決策樹狀結構」對話框中，選取 **CHAID** 或 **消耗 CHAID** 作為增長方法。
2. 按一下「**條件**」。
3. 按一下「**CHAID**」標籤。

## CHAID 分析的尺度區間

在 CHAID 分析中，在進行分析之前，尺度自 (預測值) 變數一定會先被分成離散的組別（例如 0-10、11-20、21-30 等）。您可以控制群組的初始/最大數（即使程序可能會在初始分割之後合併連續的組別）：

- **固定數。** 所有的尺度自變數一開始都會被分成相同的組別個數。預設值為 10。
- **自訂。** 每一個尺度自變數一開始都會被分為該變數指定的組別數。

若要指定尺度自變數的區間

1. 在主要的「決策樹」對話框中，選取一個或多個尺度自變數。
2. 如果是成長方法，請選取「**CHAID**」或「**Exhaustive CHAID**」。
3. 按一下「**條件**」。
4. 按一下「**區間**」標籤。

在 CRT 與 QUEST 分析中，所有的分割都是二元分割，且尺度與序數自變數都是以相同方式處理，所以，您無法為尺度自變數指定區間數。

## CRT 條件

CRT 成長方法會嘗試最大化節點內的同質性。如果某個節點無法表示觀察值的同質子集，就表示有雜質。例如，如果某個終端節點中的所有觀察值都具有應變數的相同值，由於該觀察值為純同質節點，因此不需要進一步分割。

您可以選取用來測量雜質的方法，以及進行分割節點時所需的最小雜質減少量

**雜質測量。** 如果是尺度應變數，會使用雜質的最小平方差 (LSD) 測量，它會計算節點內變異數，並為任何頻次加權或影響值調整。

如果是類別（名義、序數）應變數，您可以選取雜質測量：

- **Gini。** 會找到與應變數值相關之子節點同質性最大化的分割。Gini 是根據應變數之每個類別成員的平方機率為基礎。當節點中的所有觀察值都落在單一個類別中時，它就會達到最小值（零）。此為預設的測量。
- **Twoing。** 應變數的類別會被分成兩個子類別群組。會找到能夠以最佳方式分開兩個組別的分割。
- **Ordered Twoing。** 與 Twoing 類似，差別在於只有相鄰類別才可以加分組。這個測量只能用於序數應變數。

**改善中的最小變更。** 這是分割節點時所需的最小雜質減少量。預設值是 0.0001。較高的值會產生具有較少節點的樹狀結構。

若要指定 CRT 條件

1. 請選取「**CRT**」作為成長方法。
2. 按一下「**條件**」。
3. 按一下「**CRT**」標籤。

## QUEST 條件

如果是 QUEST 方法，您可以指定用於分割節點的顯著性水準。除非顯著性層次小於或等於指定的層次，否則您無法使用自變數來分割節點。該值必須大於 0 小於 1。預設值為 0.05。較小的值會從最終模型中排除更多的自變數。

若要指定 QUEST 條件

1. 在主「決策樹狀結構」對話框中，選取名義因變數。
2. 如果是成長方法，請選取「**QUEST**」。
3. 按一下「**條件**」。
4. 按一下「**QUEST**」標籤。

## 刪改樹狀結構

您可以使用 CRT 和 QUEST 方法避免透過**修剪**樹狀結構，過度適配模型：在達到停止準則時，樹狀結構會就會停止成長，然後樹狀結構會根據風險中所指定的最大差異自動修正為最小的子樹狀結構。風險值是以標準誤來表示。預設值為 1。該值必須是非負數。若要取得具有最低風險的子樹狀結構，請指定 0。

若要修正樹狀結構

1. 在主「決策樹」對話框中，如果是成長方法，請選取「**CRT**」或「**QUEST**」。
2. 按一下「**條件**」。
3. 按一下「**修正**」標籤。

刪改與隱藏節點

當您建立已經刪改的樹狀結構時，從樹狀結構中刪改的任何節點都無法用在最終樹狀結構中。您可以使用互動方式來隱藏或顯示在最終樹狀結構中所選擇的子節點，但是您無法顯示在樹狀結構建立過程中所刪改的節點。如需相關資訊，請參閱主題第 13 頁的『[樹狀編輯器](#)』。

## 代理

CRT 與 QUEST 可以將代理用於自變數（預測值）。對於該變數值已遺漏的觀察值而言，會使用其他具有與原始變數高度關聯的自變數來進行分類。這些替代的預測值稱為代理。您可以指定要用在模式中的最大代理數。

- 根據預設，最大的代理數是自變數的數量減去 1 的數字。換句話說，對於每一個自變數，所有其他的自變數都可以當作代理來使用。
- 如果您不希望模式使用代理，請為代理數指定 0。

若要指定代理

1. 在主「決策樹」對話框中，如果是成長方法，請選取「**CRT**」或「**QUEST**」。
2. 按一下「**條件**」。
3. 按一下「**代理**」標籤。

## 選項

可以使用的選項會因為成長方法、應變數的測量層次，及/或是否有應變數之值的已定義數值標籤等而有所不同。

### 錯誤分類成本

如果是類別（名義、序數）因變數，則錯誤分類成本可以讓您包含與不正確分類相關的相對懲罰資訊。例如：

- 拒絕信用良好之客戶所花費的成本，可能不同於擴展日後會借貸之客戶的信用所花費的成本。
- 將具有高度心臟疾病風險的人員錯誤分類為具有低度心臟疾病風險人員所付出的成本，可能會高於將具有低度心臟疾病風險的人員錯誤分類為具有高度心臟疾病風險人員所付出的成本。
- 將大量郵件傳送給不太可能回應的人，成本可能比較低，但是如果將大量郵件傳送給可能會回應的人，成本相對上可能會比較高（以損失的收益而言）。

錯誤分類成本與數值標籤

除非類別因變數至少有兩個值已經定義數值標籤，否則無法使用這個對話框。主題，以取得更多資訊。

若要指定錯誤分類成本

1. 在主要的「決策樹」對話框中，選取具有兩個或多個數值標籤的類別（名義、序數）應變數。
2. 按一下「**選項**」。
3. 按一下「**錯誤分類成本**」標籤。
4. 按一下**自訂**。
5. 在網格中輸入一個或多個錯誤分類成本。輸入的值必須是非負數。（正確的分類會顯示在對角線上，而且一律為 0）。

**填滿矩陣。** 在許多情況下，您可能想要成本對稱——也就是，將 A 分類為 B 的成本與將 B 分類為 A 的成本相同。下列控制項可讓您更容易指定對稱成本矩陣：

- **複製下三角形。** 將下半三角形矩陣（對角線底下）的值複製到對應的上半三角形資料格中。
- **複製上三角形。** 將矩陣下半三角形（對角線底下）內的值複製到對應的上半三角形資料格中。
- **使用平均資料格值。** 對於每一半矩陣的每一個資料格而言，兩個值（上與下三角形）是相加之後的平均值，而這個平均值會取代原來的兩個值。例如，如果將 A 錯誤分類為 B 所付出的成本為 1，而將 B 錯誤分類為 A 所付出的成本為 3，則這個控制會使用平均值 2 ( $(1+3)/2 = 2$ ) 來取代原來的兩個值。

## 利潤

如果是種類應變數，您可以指定應變數層次的收益與支出值。

- 利潤是以收益減去支出的方式來計算。
- 利潤值會影響獲利表中利潤與 ROI（投資報酬率）的值，但是不會影響基本樹狀結構的模式結構。

- 收益與支出值都必須是數值，而且也都必須指定給網格中所顯示之應變數的所有類別。

### 利潤與數值標籤

這個對話框需要應變數的已定義數值標籤。除非類別應變數至少有兩個值已經定義值標籤，否則無法使用這個對話框。主題，以取得更多資訊。

若要指定利潤

1. 在主要的「決策樹」對話框中，選取具有兩個或多個數值標籤的類別（名義、序數）應變數。
2. 按一下「選項」。
3. 按一下「利潤」標籤。
4. 按一下自訂。
5. 為網格中所列出的所有應變數類別輸入收益與支出值。

## 事前機率

如果是具有類別應變數的 CRT 與 QUEST 樹狀結構，您可以指定組別成員的事前機率。**事前機率**就是在瞭解自變數（預測值）之前，對應變數之每一個類別總體相對次數的估計。使用事前機率可以協助更正由非整體母群之取樣中的資料所造成的任何樹狀結構成長情況。

**從訓練範例 (經驗先驗) 取得。** 如果資料檔中的應變數值分佈是表示母群分佈，請使用這個設定。如果您是使用分割樣本驗證，就會使用訓練樣本中的觀察值分佈。

附註：由於觀察值是隨機指定給分割樣本驗證中的訓練樣本，因此無法事先知道訓練樣本中實際的觀察值分佈。如需相關資訊，請參閱主題第 4 頁的『驗證』。

**在所有類別保持相等。** 如果應變數的類別在母群中都是顯示為相等，請使用這個設定。例如，如果一共有四個類別，則每各類別中都會有大約 25% 的觀察值。

**自訂。** 為網格中所列的每一個應變數類別輸入一個非負數值。值可以是比例、百分比、次數個數，或是在所有類別中表示數值分布的其他值。

**使用錯誤分類成本調整先驗。** 如果您定義自訂錯誤分類成本，就可以根據這些成本來調整事前機率。如需相關資訊，請參閱主題第 7 頁的『錯誤分類成本』。

### 利潤與數值標籤

這個對話框需要應變數的已定義數值標籤。除非類別應變數的至少兩個值已經定義數值標籤，否則無法使用這個對話框。主題，以取得更多資訊。

若要指定事前機率

1. 在主要的「決策樹」對話框中，選取具有兩個或多個數值標籤的類別（名義、序數）應變數。
2. 如果是成長方法，請選取「**CRT**」或「**QUEST**」。
3. 按一下「選項」。
4. 按一下「事前機率」標籤。

## 評分

如果是具有序數應變數的 CHAID 與 Exhaustive CHAID，您可以自訂應變數之每一個種類的分數。分數會定義應變數各類別之間的順序與距離。您可以使用評分來增加或減少序數值之間的相對距離，或是變更值的順序。

- **為每個類別使用序數等級。** 應變數的最低類別會被指定一個 1 的分數，下一個較高的類別會被指定一個 2 的分數，依此類推。此為預設值。
- **自訂。** 為網格中所列的每一個因變數類別輸入一個數值分數。

範例

表 3: 自訂評分值		
數值標籤	原始值	評分
非技術人員	1	1
技術人員	2	4
事務人員	3	4.5
Professional	4	7
管理	5	6

- 分數會增加非技術人員與技術人員之間的相對距離，而且會減少技術人員與事務人員之間的相對距離。
- 分數會將管理人員與專業人員的順序反轉。

#### 分數與數值標籤

這個對話框需要應變數的已定義數值標籤。除非類別應變數至少有兩個值已經定義值標籤，否則無法使用這個對話框。主題，以取得更多資訊。

若要指定分數

1. 在主要的「決策樹狀結構」對話框中，選取具有兩個或多個已定義之數值標籤的序數因變數。
2. 如果是成長方法，請選取「**CHAID**」或「**Exhaustive CHAID**」。
3. 按一下「**選項**」。
4. 按一下「**分數**」標籤。

## 遺漏值

「遺漏值」標籤會控制名義值、使用者遺漏值與自（預測值）變數值的處理方式。

- 序數與尺度使用者遺漏的自變數值的處理方式會因為成長方法而有所不同。
- 名義應變數的處理方式是在「類別」對話框中所指定。如需相關資訊，請參閱主題第 3 頁的『[選取類別](#)』。
- 如果是序數與尺度應變數，一定會排除具有系統遺漏或使用者遺漏的應變數值的觀察值。

**視為遺漏值處理。** 使用者遺漏值會被視為系統遺漏值來處理。系統遺漏值的處理方式會因為成長方法而有所不同。

**視為有效值處理。** 名義自變數的使用者遺漏值會被視為樹狀結構成長與分類中的普通值來處理。

方法相依規則

如果某些（非全部）自變數值是系統遺漏值或使用者遺漏值：

- 如果是 CHAID 與 Exhaustive CHAID，系統遺漏與使用者遺漏的自變數值會以單一、組合的類別包含在分析中。如果是尺度與序數自變數，則演算法會先使用有效的值來產生類別，然後決定是否要將遺失的類別和其最相似的（有效的）類別加以合併，或是將其維持為一個個別的類別。
- 如果是 CRT 與 QUEST，則具有遺失自變數值的觀察值會從樹狀結構成長過程中被排除，但是會使用代理來加以分類（如果方法中含有代理的話）。如果名義使用者遺漏值是被視為遺漏值來處理的話，也會使用這個方法來處理這些值。如需相關資訊，請參閱主題第 7 頁的『[代理](#)』。

若要指定名義、自變數使用者遺失處理

1. 在主要的「決策樹」對話框中，選取至少一個名義自變數。
2. 按一下「**選項**」。
3. 按一下「**遺漏值**」標籤。

## 儲存模型資訊

您可以將模型的資訊儲存為工作資料檔中的變數，也可以將整個模型以 XML (PMML) 格式儲存至某個外部檔案。

已儲存變數

**終端節點數。** 每個觀察值獲指定的終端節點。值就是樹狀結構節點數。

**預測值。** 由模式所預測之應變數的類別（群組）或值。

**預測的機率。** 與模型的預測相關聯的機率。系統會為應變數的每一個類別儲存一個變數。不適用於尺度因變數。

**樣本指定（訓練/測試）。** 如果是分割樣本驗證，這個變數會指出訓練或檢定樣本中是否有使用觀察值。訓練樣本的值為 1，檢定樣本的值則為 0。除非您已經選取分割樣本驗證，否則無法使用。請參閱第 4 頁的『[驗證](#)』主題，以取得更多資訊。

以 XML 格式匯出樹狀結構模型

您可以將整個樹狀結構模型儲存為 XML (PMML) 格式。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。主題，以取得更多資訊。

**訓練樣本。** 將模式寫入至指定的檔案。如果是分割樣本驗證樹狀結構，這是用於訓練樣本的模型。

**測試樣本。** 將檢定樣本的模型寫入至指定的檔案。除非您已經選取分割樣本驗證，否則無法使用。

## 輸出

可用的輸出選項依成長方法、應變數的測量層級與其他設定而定。

### 樹狀結構顯示

您可以控制樹狀結構的初始外觀，或完全隱藏樹狀結構顯示。

**樹狀結構。** 依預設，樹狀結構表是包含在「檢視器」中所顯示的輸出中。取消選取這個選項，就可以從輸出中排除樹狀結構圖。

**顯示。** 這些選項會控制「檢視器」中樹狀結構圖的初始外觀。所有這些屬性都可以藉由編輯產生的樹狀結構來加以修改。

- **方向。** 樹狀結構可以顯示為根節點在頂端，可以從上到下展開、或根節點在左右兩側，可以從左到右或從右到左展開。
- **節點內容。** 節點可以顯示表格、圖表，或同時顯示兩者。如果是類別應變數，表格會顯示次數個數與百分比，而圖表則為長條圖。如果是尺度應變數，表格會顯示平均數、標準差、觀察值數與預測的值，而圖表則為直方圖。
- **尺度。** 依預設，大型的樹狀結構都會自動調整，嘗試讓樹狀結構能夠符合頁面的大小。您可以指定高達 200% 的自訂尺度百分比。
- **自變數統計量。** 如果是 CHAID 與 Exhaustive CHAID，統計量包括  $F$  值（用於尺度應變數）或卡方值（用於種類應變數，以及顯著值和自由度。如果是 CRT，則會顯示改善值。如果是 QUEST，則會為尺度與序數自變數顯示  $F$ 、顯著值與自由度；如果是名義自變數，則會顯示卡方值、顯著值與自由度。
- **節點定義。** 節點定義會顯示每個節點分割使用之自變數的值。

**表格格式中的樹狀結構。** 樹狀結構中每個節點的摘要資訊，包括上層節點數、自變數統計量、節點的自變數值、尺度應變數的平均數與標準差，或是類別應變數的個數與百分比。

若要指定初始的樹狀結構顯示

1. 在主「決策樹狀結構」對話框中，按一下**輸出**。
2. 按一下「**樹狀結構**」標籤。

### 統計資料

可用的統計量表格根據應變數的測量層級、成長方法與其他設定而定。

## 模型

**摘要。**摘要包括使用的方法、模式中所包括的變數，以及模式中所指定但未包括的變數。

**風險。**風險估計與其標準誤。樹狀結構預測準確性的測量。

- 如果是類別應變數，風險估計就是在事前機率和錯誤分類成本調整之後，不正確分類之觀察值的比例。
- 如果是尺度應變數，風險估計是在節點變異數的範圍中。

**分類表。**如果是類別（名義、序數）應變數，這個表格就會顯示為每個應變數類別正確分類與不正確分類之的觀察值數。不適用於尺度應變數。

**成本、事前機率、分數與利潤值。**如果是類別應變數，這個表格會顯示分析中所使用的成本、事前機率、分數與利潤值。不適用於尺度應變數。

## 自變數(V)

**模型的重要性。**如果是 CRT 成長方法，則會根據其對模型的重要性來將每個自（預測值）變數分等。不適用於 QUEST 或 CHAID 方法。

**根據分割來代理。**對於 CRT 與 QUEST 成長方法，如果模式包括代理，則會列出樹狀結構中每個分割的代理。不適用於 CHAID 方法。如需相關資訊，請參閱主題 [第 7 頁的『代理』](#)。

## 節點效能

**摘要。**如果是尺度應變數，這個表格會包括節點數、觀察值數，以及應變數的平均值。如果是具有已定義之利潤的類別應變數，則表格會包括節點數、平均利潤以及 ROI（投資報酬率）值。不適用於沒有已定義之利潤的類別應變數。如需相關資訊，請參閱主題 [第 7 頁的『利潤』](#)。

**依目標分類。**如果是具有已定義之目標類別的類別應變數，則表格會包括百分比獲利、回應百分比以及根據節點或百分位數組別所區分的索引百分比（提升）。每個目標類別都會產生個別的表格。不適用於沒有已定義之目標類別的尺度應變數或類別應變數。如需相關資訊，請參閱主題 [第 3 頁的『選取類別』](#)。

**列。**節點效能表可以根據終端節點、百分位數或兩者來顯示結果。如果您選取同時使用兩者，每個目標類別就會產生兩個表格。百分位數表會根據排序順序，為每個百分位數顯示累積值。

**百分位數增量。**如果是百分位數表，您可以選取百分位數增量：1、2、5、10、20 或 25。

**顯示累積統計量。**如果是終端節點表，會在每個表格顯示更多的欄位，來顯示累積結果。

若要選取統計量輸出。

1. 在主「決策樹狀結構」對話框中，按一下**輸出**。
2. 按一下「**統計量**」標籤。

## 圖表

可用的圖表是根據應變數的測量層級、成長方法與其他設定而定。

**自變數對模型的重要性。**根據自變數（測量值）之模式重要性長條圖。只適合與 CRT 成長方法搭配使用。

## 節點效能

**獲利。**「獲利」是指每個節點之目標類別的總觀察值的百分比，計算方式為： $(\text{節點目標 } n / \text{目標總計 } n) \times 100$ 。獲利圖表是獲利累積百分位的折線圖，計算方式為： $(\text{累積百分位目標 } n / \text{目標總計 } n) \times 100$ 。每個目標類別都會產生分隔線圖表。只適用於有定義之目標類別的類別應變數。如需相關資訊，請參閱主題 [第 3 頁的『選取類別』](#)。

獲利圖表會繪製與您在百分位數表之獲利的「獲利百分比」行中所見相同的值，這個百分位數表也會報表累積值。

**指標(N).**指標為目標種類之節點回應值百分比與整個樣本之整體目標種類回應值百分比相較之下，所得出的比例。索引圖表是累積百分位數索引值的折線圖。僅適用於類別應變數。累積百分位索引的計算方式為： $(\text{累積百分位回應百分比} / \text{回應百分比總計}) \times 100$ 。每個目標類別都會產生獨立圖表，且必須定義目標類別。

指數圖表會繪製與您在百分位數表之獲利的「指數」行中所見相同的值。

回應(E). 指定的目標種類中，節點內的觀察值百分比。回應圖表是累積百分位數回應的折線圖，計算方式為： $(\text{累積百分位數目標 } n / \text{累積百分位總計 } n) \times 100$ 。只適用於有定義之目標類別的類別應變數。

回應圖表會繪製與您在百分位數表之獲利的「回應」行中所見相同的值。

**平均數。** 應變數的累積百分位數平均值線性圖。僅適用於尺度應變數。

**平均利潤。** 累積平均利潤的線性圖。僅適用於具有已定義之利潤的類別應變數。如需相關資訊，請參閱主題第 7 頁的『利潤』。

平均利潤圖表會繪製與您在百分位數表之獲利摘要的「利潤」行中所見相同的值。

**投資報酬率 (ROI)。** 累積的 ROI (投資報酬) 的線性圖。ROI 是以利潤對支出的比率來計算。僅適用於具有已定義之利潤的類別應變數。

ROI 圖表會繪製與您在百分位數表之獲利摘要的「ROI」行中所見相同的值。

**百分位數增量。** 對於所有的百分位數圖表，這個設定會控制圖表上所顯示的百分位數增量：1、2、5、10、20 或 25。

若要選取圖表輸出。

1. 在主「決策樹狀結構」對話框中，按一下**輸出**。
2. 按一下「**圖形**」標籤。

## 選擇和評分規則

「規則」標籤會提供以指令語法、SQL 或範例 (純英文) 文字等形式來產生選項或分類/預測規則的功能。您可以在「檢視器」中顯示這些規則及/或將這些規則儲存至某個外部檔案。

**語法。** 控制在「檢視器」中顯示之輸出以及儲存為外部檔案兩者的選擇規則。

- **IBM SPSS Statistics.** 指令語法語言。規則是以定義用於選取觀察值子集之過濾條件的一組指令來表示，或以用於為觀察值評分的 COMPUTE 陳述式來表示。
- **SQL。** 標準的 SQL 規則是用來從資料庫中選取或擷取記錄，或是將值指定給這些記錄。產生的 SQL 規則不包含任何表格名稱或其他資料來源資訊。
- **簡單文字。** 純英文虛擬程式碼。規則表示為一組邏輯 "if...then" 陳述式，這一組陳述式可以描述模型的分類或每一個節點的預測。這種形式的規則可以用來定義變數和值標籤或變數名稱和資料值。

**類型。** 若是 IBM SPSS Statistics 和 SQL 規則，可控制所產生規則的類型：選擇或評分規則。

- **指定值給觀察值。** 此規則可用來指定模型的預測給符合節點成員資格條件的觀察值。另外會為符合節點成員資格條件的各節點產生不同的規則。
- **選取觀察值。** 此規則可用來選取符合節點成員資格條件的觀察值。有關 IBM SPSS Statistics 或 SQL 規則，會產生單一規則，以選取符合選擇條件的所有觀察值。

**IBM SPSS Statistics** 和 SQL 規則中包含代理。您可以在 CRT 和 QUEST 中，包含規則中模型的代理預測值。包含代理的規則可能會相當複雜。一般來說，如果只要推導有關樹狀結構的概念資訊，請排除代理。如果有些觀察值有不完整的自變數 (預測值) 資料，而您要模擬樹狀結構的規則，請包含代理。如需相關資訊，請參閱主題 第 7 頁的『代理』。

**節點。** 控制所產生規則的範圍。將產生一個單獨的規則用於範圍中納入的每個節點。

- **所有的終端節點。** 為每個終端節點產生規則。
- **最佳終端節點。** 根據指數值，為前  $n$  個終端節點產生規則。如果數量超過樹狀結構中終端節點的數量，就會為所有的終端節點產生規則 (請參閱以下注意事項)。
- **最佳終端節點會往上移至指定的觀察值百分比。** 根據指數值，為前  $n$  百分比觀察值的終端節點產生規則 (請參閱以下注意事項)。
- **其索引值符合或超過分割值的終端節點。** 為其指數值大於或等於指定值的所有終端節點產生規則。大於 100 的指數值表示該節點中目標類別內的觀察值百分比已經超過根節點中的百分比。(請參閱以下注意事項)。
- **所有節點。** 為所有節點產生規則。



附註 1: 以索引值為根據的節點選項功能僅適用於有已定義之目標類別的類別因變數。如果您已經指定多個目標類別，就會為每個目標類別產生一組個別的規則。

Note 2: For IBM SPSS Statistics and SQL rules for selecting cases (not rules for assigning values), **All nodes** and **All terminal nodes** will effectively generate a rule that selects all cases used in the analysis.

**將規則匯出至檔案。** 將規則儲存在外部文字檔中。

您也可以根據最終樹狀結構模型中選取的節點，以互動方式產生並儲存選擇或評分規則。如需相關資訊，請參閱主題 [第 15 頁的『觀察值選擇和評分規則』](#)。

附註: 如果您將指令語法格式的規則套用到另一個資料檔，則該資料檔必須包含與最終模型中之自變數相同名稱、使用相同之單位測量，並且有使用者定義遺漏值（如果有的話）的變數。

若要指定選項或分數規則

1. 在主「決策樹狀結構」對話框中，按一下**輸出**。
2. 按一下「規則」標籤。

## 樹狀編輯器

使用「樹狀模型編輯器」時，您可以：

- 隱藏和顯示選擇的樹狀結構分支。
- 控制節點內容、分割節點的統計量，以及其他資訊的顯示。
- 變更節點、背景、框線、圖表和字型顏色。
- 變更字型樣式和大小。
- 變更樹狀結構對齊方式。
- 依據選擇的節點，選擇要進一步分析的觀察值子集。
- 依據選擇的節點，建立和儲存選擇或評分觀察值的規則。

若要編輯樹狀結構模型：

1. 在「檢視器」視窗中，連按兩下樹狀結構模型。  
或
2. 從「編輯」功能表選擇，或用滑鼠右鍵按一下蹦現功能表選擇：

### 編輯

隱藏和顯示節點

若要隱藏（收合）上層節點下分支中的所有子節點：

1. 在上層節點的右下角，按一下小方框中的減號 (-)。

在該分支上層節點下所有的子節點將會隱藏。

若要顯示（展開）上層節點下分支中的所有子節點：

2. 在上層節點的右下角，按一下小方框中的加號 (+)。

附註：隱藏分支中的子節點與修正樹狀結構是不一樣的。如果您要的是已修正的樹狀結構，您必須在建立樹狀結構之前要求修正，而且已修正的分支不會包含在最後的樹狀結構之中。如需相關資訊，請參閱主題 [第 6 頁的『刪改樹狀結構』](#)。

選擇多個節點

您可以依據目前選擇的節點，選擇觀察值、產生評分和選擇規則，以及執行其他動作。若要選擇多個節點：

1. 按一下您要選擇的節點。
2. 按住 Ctrl 鍵不放，然後再按您要選擇的節點。

您可以選擇一個分支中多個相鄰的節點和/或上層節點，以及其他分支中的子節點。但是，您不能選擇同一個節點分支中的上層節點和子節點/其下節點。

## 使用大型樹狀結構

有時候，樹狀結構模型包含有太多節點和分支，很難或甚至不可能檢視整個完整的樹狀結構。在使用大型樹狀結構時，有一些實用的功能：

- **樹狀結構圖。** 您可以使用樹狀結構圖（尺寸較小，是樹狀的簡化版）瀏覽樹狀結構和選擇節點。如需相關資訊，請參閱主題 [第 14 頁的『樹狀圖』](#)。
- **縮放比例。** 您可以變更縮放比例，縮小或放大樹狀結構顯示。如需相關資訊，請參閱主題 [第 14 頁的『縮放樹狀結構顯示』](#)。
- **節點和分支顯示。** 您可以利用只顯示節點中的表格或圖表，和（或）隱藏節點標記或自變數的顯示資訊，使樹狀結構看起來更精簡。如需相關資訊，請參閱主題 [第 15 頁的『控制樹狀結構中顯示的資訊』](#)。

### 樹狀圖

樹狀結構圖提供精簡、簡化的樹狀結構檢視，讓您可以瀏覽樹狀結構和選擇節點。

若要使用樹狀結構圖視窗：

1. 從「樹狀模型編輯器」功能表選擇：

#### 檢視 > 樹狀對映

- 目前選擇的節點會在「樹狀結構模型編輯器」和樹狀結構圖視窗中反白顯示。
- 樹狀結構圖中的紅色方框表示目前正在「樹狀結構模型編輯器」中檢視的區域。按一下滑鼠右鍵並拖曳方框可變更檢視區域中顯示的樹狀結構區段。
- 如果您在樹狀結構圖中選擇了目前不在「樹狀模型編輯器」視圖區域中的節點，則檢視區域會移至包含該選取節點的區域以供檢視。
- 在樹狀結構圖和「樹狀模型編輯器」中選擇多個節點的方式相同：按住 Ctrl 鍵並選取多個節點。您不能選擇同一個節點分支中的上層節點和子節點/其下節點。

### 縮放樹狀結構顯示

依預設，樹狀結構會自動縮放至符合「檢視器」視窗的大小，因此部分樹狀結構在剛開始時較不容易讀取。您可以選擇預設的縮放比例設定值，或是輸入您自訂的縮放比例值，範圍從 5% 到 200%。

若要變更樹狀結構的縮放比例值：

1. 在工具列的下拉清單中，選擇縮放比例，或是輸入自訂的比例值。

或

2. 從「樹狀模型編輯器」功能表選擇：

#### 檢視 > 比例...

您可以在建立樹狀結構模型前指定縮放比例值。如需相關資訊，請參閱主題 [第 10 頁的『輸出』](#)。

### 節點摘要視窗

節點摘要視窗提供已選擇節點的放大檢視。您也可以依據選擇節點，使用摘要視窗來檢視、套用，或是儲存選項或評分規則。

- 在節點摘要視窗中，使用「檢視」功能表切換檢視摘要表格、圖表或規則。
- 在節點摘要視窗中，使用「規則」功能表選擇您要查看的規則類型。如需相關資訊，請參閱主題 [第 15 頁的『觀察值選擇和評分規則』](#)。
- 所有節點摘要視窗中的檢視會反映所有已選擇節點的組合摘要。

若要使用節點摘要視窗：

1. 在「樹狀模型編輯器」中選擇節點。若要選取多個節點，可以按住 Ctrl 鍵並按一下來選取。

2. 從功能表中選擇：

#### 檢視 > 摘要

## 控制樹狀結構中顯示的資訊

「樹狀模型編輯器」中的「選項」功能表可讓您控制顯示節點內容、自變數（預測變數）名稱和統計量、節點定義和其他設定值。其中許多設定值也可以從工具列進行控制。

## 變更樹狀結構的顏色和字型

您可以在樹狀結構中變更如下的顏色：

- 節點框線、背景和文字顏色
- 分支顏色和分支文字顏色
- 樹狀結構背景顏色
- 預測類別反白顯示的顏色（類別應變數）
- 節點圖表顏色

您可以變更樹狀結構中所有的字型、樣式和大小。

附註：您無法變更個別節點或分支的顏色或字型屬性。顏色變更會套用至所有相同類型的元素，以及字型變更（不同於顏色）會套用至所有圖表的元素。

若要變更顏色和字型屬性：

1. 使用工具列變更整個樹狀結構的字型屬性，或是不同樹狀結構元素的顏色（當您將滑鼠游標移至工具列的控制項上方，「工具提示」會顯示說明資訊）。  
或
2. 在「樹狀模型編輯器」的任意處連按兩下開啟「性質」視窗，或是在功能表中選擇：  
**檢視 > 內容**
3. 有關框線、分支、節點背景、預測類別，和樹狀結構背景，按一下「**顏色**」標籤。
4. 有關字型顏色和屬性，按一下「**文字**」標籤。
5. 有關節點圖表顏色，按一下「**節點圖表**」標籤。

## 觀察值選擇和評分規則

您可以利用「樹狀模型編輯器」，執行下列動作：

- 依據選擇的節點，選擇觀察值子集。如需相關資訊，請參閱主題第 15 頁的『過濾觀察值』。
- 產生 IBM SPSS Statistics 指令語法或 SQL 格式的觀察值選擇或評分規則。如需相關資訊，請參閱主題第 16 頁的『儲存選擇和評分規則』。

當您執行「決策樹」程序來建立樹狀結構模式時，您也可以依據多個準則自動儲存規則。如需相關資訊，請參閱主題第 12 頁的『選擇和評分規則』。

## 過濾觀察值

如果您想要進一步瞭解特定節點或節點群組中的觀察值，您可以依據選擇的節點來選取要進一步分析的觀察值子集。

1. 在「樹狀模型編輯器」中選擇節點。若要選取多個節點，可以按住 Ctrl 鍵並按一下來選取。
2. 從功能表中選擇：

**規則 > 過濾觀察值...**

3. 輸入過濾變數名稱。選擇節點中的個案將收到變數值 1。所有其他觀察值將會收到數值 0，並將在接下來的分析中被執行，直到您變更過濾狀態為止。
4. 按一下「**確定**」。

## 儲存選擇和評分規則

您可以將觀察值選擇或評分規則儲存在外部檔案，然後套用那些規則至不同的資料來源。這些規則是依據「樹狀模型編輯器」中選擇的節點。

**語法。** 控制在「檢視器」中顯示之輸出以及儲存為外部檔案兩者的選擇規則。

- **IBM SPSS Statistics.** 指令語法語言。規則是以定義用於選取觀察值子集之過濾條件的一組指令來表示，或以用於為觀察值評分的 COMPUTE 陳述式來表示。
- **SQL。** 標準 SQL 規則是用來從資料庫中選擇/擷取記錄，或指定值給這些記錄。產生的 SQL 規則不包含任何表格名稱或其他資料來源資訊。

**類型。** 您可以建立選擇或評分規則。

- **選取觀察值。** 此規則可用來選取符合節點成員資格條件的觀察值。有關 IBM SPSS Statistics 或 SQL 規則，會產生單一規則，以選取符合選擇條件的所有觀察值。
- **指定值給觀察值。** 此規則可用來指定模型的預測給符合節點成員資格條件的觀察值。另外會為符合節點成員資格條件的各節點產生不同的規則。

**包括代理。** 您可以在 CRT 和 QUEST 中，包含規則中模型的代理預測值。包含代理的規則可能會相當複雜。一般來說，如果只要推導有關樹狀結構的概念資訊，請排除代理。如果有些觀察值有不完整的自變數（預測值）資料，而您要模擬樹狀結構的規則，請包含代理。如需相關資訊，請參閱主題 [第 7 頁的『代理』](#)。

若要儲存觀察值選擇或評分規則：

1. 在「樹狀模型編輯器」中選擇節點。若要選取多個節點，可以按住 Ctrl 鍵並按一下來選取。

2. 從功能表中選擇：

**規則 > 匯出...**

3. 選擇您需要的規則類型，然後輸入檔名。

附註：如果您將指令語法格式的規則套用到另一個資料檔，則該資料檔必須包含與最終模型中之自變數相同名稱、使用相同之單位測量，並且有使用者定義遺漏值（如果有的話）的變數。

## 注意事項

---

本資訊係針對 IBM 在美國所提供之產品與服務所開發。IBM 可能會提供此資料的其他語言版本。然而，貴客戶可能需要擁有該語言的產品或產品版本副本，才能進行存取。

IBM 可能不會在其他國家或地區提供本文件所討論的產品、服務或特性。請洽詢當地的 IBM 業務代表，以取得當地目前提供的產品和服務之相關資訊。本文件在提及 IBM 的產品、程式或服務時，不表示或暗示只能使用 IBM 的產品、程式或服務。只要未侵犯 IBM 之智慧財產權，任何功能相當之產品、程式或服務皆可取代 IBM 之產品、程式或服務。不過，任何非 IBM 之產品、程式或服務，使用者必須自行負責作業之評估和驗證責任。

本文件所說明之主題內容，IBM 可能擁有其專利或專利申請案。提供本文件不代表授與這些專利的授權。您可以用書面方式來查詢授權，來函請寄到：

*IBM Director of Licensing*

*IBM Corporation*

*North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
US*

若要查詢有關雙位元組 (DBCS) 資訊的授權事宜，請洽詢所在國家或地區的 IBM 智慧財產部門，或書面提出授權查詢，來函請寄到：

*Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan*

International Business Machines Corporation 只依「現況」提供本出版品，不提供任何明示或默示之保證，其中包括且不限於不侵權、可商用性或特定目的之適用性的隱含保證。有些地區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。

本參考資訊中可能會有技術上或排版印刷上的訛誤。因此，IBM 會定期修訂；並將修訂後的內容納入新版中。IBM 隨時會改進及/或變更本出版品所提及的產品及/或程式，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。

如果本程式之獲授權人為了 (i) 在個別建立的程式和其他程式（包括本程式）之間交換資訊，以及 (ii) 相互使用所交換的資訊，因而需要相關的資訊，請洽詢：

*IBM Director of Licensing*

*IBM Corporation*

*North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
US*

這些資訊可能可以使用，但必須遵循適當的條款，在某些情況中需要付費。

IBM 基於雙方之「IBM 客戶合約」、「IBM 國際程式授權合約（或任何同等合約）條款，提供本文件所提及的授權程式與其所有適用的授權資料。

本文件中引用的效能資料及用戶範例僅供敘述之目的。實際效能結果可能會依據特定配置和作業條件而有所不同。

本文件所提及之非 IBM 產品資訊，係取自產品供應商，或其發佈的聲明或其他公開管道。IBM 並未測試過這些產品，也無法確認這些非 IBM 產品的執行效能、相容性或任何對產品的其他主張是否完全無誤。有關非 IBM 產品功能之問題，應直接洽詢產品供應商。

所有關於 IBM 未來方針或目的之聲明，隨時可能更改或撤銷，不必另行通知，且僅代表目標與主旨。

本資訊含有日常企業運作所用之資料和報告範例。為了盡可能詳盡說明，這些範例都包括個人、公司、品牌及產品的名稱。所有這些名稱全為虛構，任何與實際人員或商業企業類似之處，純屬巧合。

著作權授權：

本資訊含有原始語言之範例應用程式，用以說明各作業平台中的程式設計技術。貴客戶可以為了研發、使用、銷售或散布符合範例應用程式所適用的作業平台之應用程式介面的應用程式，以任何形式複製、修改及散布這些範例程式，不必向 IBM 付費。這些範例並未在所有情況下完整測試。因此，IBM 不保證或暗示這些程式的可靠性、服務性或功能。這些程式範例以「現狀」提供，且無任何保證。IBM 對因使用這些程式範例而產生的任何損害概不負責。

這些範例程式或任何衍產生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：

© Copyright IBM Corp. 2021. 此程式碼部分衍生自 IBM 公司 程式範例。

© Copyright IBM Corp. 1989 - 2021. All rights reserved.

## 商標

---

IBM、IBM 標誌及 [ibm.com](http://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](http://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 及 Pentium 是 Intel Corporation 在美國及（或）其他國家或地區商標或註冊商標。

Linux 是 Linus Torvalds 在美國及/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 及所有 Java 型商標及標誌是 Oracle 及/或附屬公司的商標或註冊商標。

# 索引

## Special Characters

- 分割樣本驗證
  - 樹狀結構 [4](#)
- 加權觀察值
  - 決策樹中的分數加權 [1](#)
- 用於分割節點的顯著性層級 [6](#)
- 交叉驗證
  - 樹狀結構 [4](#)
- 成本
  - 錯誤分類 [7](#)
- 收合樹狀結構分支 [13](#)
- 利潤
  - 事前機率 [8](#)
  - 樹狀結構 [7, 10](#)
- 刪改決策樹
  - 與隱藏節點 [6](#)
- 決策樹
  - 強制第一個變數進入模式中 [1](#)
  - 測量層次 [1](#)
  - CHAID 方法 [1](#)
  - CRT 方法 [1](#)
  - Exhaustive CHAID 方法 [1](#)
  - QUEST 方法 [1](#)
- 決策樹狀結構
  - QUEST 方法 [6](#)
- 指令語法
  - 建立決策樹狀結構的選項與評分語法 [12, 15](#)
- 風險估計
  - 樹狀結構 [10](#)
- 索引值
  - 樹狀結構 [10](#)
- 規則
  - 建立決策樹狀結構的選項與評分語法 [12, 15](#)
- 測量層次
  - 決策樹 [1](#)
- 評分
  - 樹狀結構 [8](#)
- 亂數種子
  - 決策樹狀結構驗證 [4](#)
- 節點
  - 選擇多個樹狀結構節點 [13](#)
- 節點數
  - 從決策樹儲存為變數 [10](#)
- 預測的機率
  - 從決策樹儲存為變數 [10](#)
- 預測值
  - 從決策樹儲存為變數 [10](#)
- 語法
  - 建立決策樹狀結構的選項與評分語法 [12, 15](#)
- 樹狀結構
  - 分割樣本驗證 [4](#)
  - 尺度自變數的區間 [5](#)
  - 文字屬性 [15](#)
  - 交叉驗證 [4](#)
  - 字型 [15](#)
  - 利潤 [7](#)

- 樹狀結構 (繼續)
  - 刪改 [6](#)
  - 事前機率 [8](#)
  - 使用大型樹狀結構 [14](#)
  - 表格中的樹狀結構目錄 [10](#)
  - 限制層級數量 [4](#)
  - 風險估計 [10](#)
  - 索引值 [10](#)
  - 控制節點大小 [4](#)
  - 控制樹狀結構顯示 [10, 15](#)
  - 產生規則 [12, 15](#)
  - 終端節點統計量 [10](#)
  - 評分 [8](#)
  - 節點圖表顏色 [15](#)
  - 預測值重要性 [10](#)
  - 圖表 [11](#)
  - 編輯 [13](#)
  - 樹狀結構方向 [10](#)
  - 樹狀圖 [14](#)
  - 選擇多個節點 [13](#)
  - 遺漏值 [9](#)
  - 錯誤分類成本 [7](#)
  - 錯誤分類表格 [10](#)
  - 儲存模型變數 [10](#)
  - 縮放樹狀結構顯示 [14](#)
  - 隱藏分支和節點 [13](#)
  - 顏色 [15](#)
  - 顯示與隱藏分支統計量 [10](#)
  - CHAID 成長條件 [5](#)
  - CRT 方法 [6](#)
  - 選擇多個樹狀結構節點 [13](#)
  - 遺漏值
    - 樹狀結構 [9](#)
  - 錯誤分類
    - 成本 [7](#)
    - 樹狀結構 [10](#)
  - 隱藏節點
    - 與刪改 [6](#)
  - 隱藏樹狀結構分支 [13](#)
  - 雜質
    - CRT 樹狀結構 [6](#)
  - 驗證
    - 樹狀結構 [4](#)

## C

- CHAID
  - 分割與合併準則 [5](#)
  - 反覆運算次數上限 [5](#)
  - 尺度自變數的區間 [5](#)
  - 重新分割合併的種類 [5](#)
  - Bonferroni 法調整 [5](#)
- CRT
  - 刪改 [6](#)
  - 雜質測量 [6](#)

## G

Gini [6](#)

## O

Ordered Twoing [6](#)

## Q

QUEST  
刪改 [6](#)

## S

SQL  
建立選項與評分的 SQL 語法 [12](#), [15](#)

## T

Twoing [6](#)





