

相對重要度指數的探討

Discussion of Relative Importance Index

顏 月 珠 *

Yen Yueh-chu

摘 要

本文探討一統計量——相對重要度指數 $RI_i = \frac{\sum_{j=1}^n (m - R_{ij} + 1)}{\sum_{i=1}^m \sum_{j=1}^n (m - R_{ij} + 1)}$

及其推論，此統計量適用於調查項目呈等級區分的調查。

一、前言

1930 年代末期，統計學家開始用不同的方法尋找機率分配，希望盡量減少或不修改所建立的機率模型，並用簡單方法尋求所要的正確機率或近似機率，因此逐漸發展而成爲統計推論的另一支——無母數統計學（nonparametric statistics）。

關鍵詞（key words）：

相對重要度指數（relative importance index）

無母數統計學（nonparametric statistics）

*國立台灣大學財務金融系副教授

無母數統計學發展的歷史雖短，但潛力無窮。在人文社會科學研究的範疇內，有許多現象是不能窺視全貌，即不能知道母體的型態與性狀，若貿然以有母數統計方法（parametric statistical method）推論之，常不能充分顯示母體的特徵，甚至誤入歧途而導至錯誤的結論，故當母體為變動的、型態未知的、只能以名義尺度或序列尺度表示的、或母體極端偏斜等情況下，沒有適用的有母數統計方法可作分析推論，則無母數統計學是唯一的選擇。

相對重要度指數（relative importance index） RI_i 是無母數統計學的一個統計量（statistic），1988年由美國森林生物統計學家Leuschner, W.A., Gregoire, T.G. 及 Buhyoff, G.J. 提出 [1]，可代替傳統之平均等級（mean rank） \bar{R}_i 來說明各研究項目的重要程度。只可惜 Leuschner 等將相對重要度指數 RI_i 的公式弄得太繁複 [2]，且 RI_i 與平均順序等級 \bar{R}_i 的線性轉換又錯得離譜 [3]，故筆者重新定義公式，找尋相對重要度指數與平均等級的關係，並探討可用的統計推論，推演檢定公式，且佐以實證。

二、相對重要度指數的意義

設一個可以等級 $R_{ij} = 1, 2, \dots, m$ 表示的問項，訪問 n 個受訪者；或一個集區設計（block design）有 n 個集區 m 個處理，則各項目或處理的相對

-
- [1] W.A. Leuschner, T.G. Gregoire, G.J. Buhyoff (1988), “ RI_i : A statistic for Reporting Ranked Responses”, *Journal of Leisure Research*, Vol. 20, No. 3, pp. 228-232.

- [2] Leuschner 等提出之公式為：
$$RI_j = \frac{\sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^N W_{ijk}}{\sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^N W_{ijk}} \times 100$$
，式中 W_{ijk} 是第 j 項目為第 k 個被詢問者所指派之第 i 等級的權數。

- [3] Leuschner 等提出，當 $m = n$ 時，相對重要度指數 RI_j 與平均順序等級 R_j 的線性轉換公式為：
$$RI_j = \{R_j \times 2 / [m(m+1)]\} \times 100。$$

重要度指數 RI_i 為：

$$RI_i = \frac{\sum_{j=1}^n (m - R_{ij} + 1)}{\sum_{i=1}^m \sum_{j=1}^n (m - R_{ij} + 1)} \quad (1)$$

式中 R_{ij} 表示第 i 個項目（或處理）之第 j 個受訪者（或集區）所填註的等級，未填註的不予計算， $i = 1, 2, \dots, m$ ， $j = 1, 2, \dots, n$ ， RI_i 代表第 i 個項目或處理之重要度所占的比例或百分比， $\sum_{i=1}^m RI_i = 100\% = 1$ ，故當 $RI_1 = 10\%$ ， $RI_2 = 20\%$ ，……，則表示第二個項目或處理之重要性大於第一個，且其重要程度占有所有項目或處理的 20% 。

所謂傳統的平均等級是項目或處理填註等級 R_{ij} 後，求算此 m 個項目或處理各 n 個 R_{ij} 的順序等級平均數，即平均順序等級 \bar{R}_i 為：

$$\bar{R}_i = \frac{\sum_{j=1}^n R_{ij}}{n} \quad (2)$$

由於 $R_i + R_{m-i+1} = m + 1$ ，即第 i 個項目或處理的等級 R 可為：

等 級	R						
順序 R_i	1	2	3	……	$m-2$	$m-1$	m
逆序 R_{m-i+1}	m	$m-1$	$m-2$	……	3	2	1
$R_i + R_{m-i+1}$	$m+1$	$m+1$	$m+1$	……	$m+1$	$m+1$	$m+1$

故另一種形式的平均等級（平均逆序等級） \bar{R}_{m-i+1} 為：

$$\bar{R}_{m-i+1} = (m + 1) - \bar{R}_i \quad (3)$$

平均順序等級（即傳統的平均等級 \bar{R}_i ）小者，其重要程度反而大，故平均等級的解釋常令人懷疑，尤其是填註等級有欠缺時，其意義令人不解，相對重要度指數 RI_i 沒有這些缺點。當 RI_i 大時，表示此項目或處理的重要程度大，指數 RI_i 與重要程度有一致的方向，即指數 RI_i 能對各研究項目或處理之重要程度提供合理的解釋，使人容易了解，此為指數 RI_i 勝過傳統平均等級的原因之一。

不論項目或處理的個數 m 及受訪者或集區數 n 的大小，以及 m 、 n 相等與否，只要等級 R_{ij} 沒有欠缺，則公式(1)之分子、分母可為：

$$\text{分子} = \sum_{j=1}^n (m - R_{ij} + 1) = \sum_{j=1}^n (m + 1) - \sum_{j=1}^n R_{ij} = n(m + 1) - \sum_{j=1}^n R_{ij}$$

$$\begin{aligned} \text{分母} &= \sum_{i=1}^m \sum_{j=1}^n (m - R_{ij} + 1) = \sum_{i=1}^m \sum_{j=1}^n (m + 1) - \sum_{i=1}^m \sum_{j=1}^n R_{ij} \\ &= mn(m + 1) - \frac{nm(m + 1)}{2} = \frac{nm(m + 1)}{2} \end{aligned}$$

故公式(1)可改寫為：

$$RI_i = \frac{n(m + 1) - \sum_{j=1}^n R_{ij}}{\frac{nm(m + 1)}{2}} \quad (4)$$

由此可得相對重要度指數 RI_i 、平均順序等級 \bar{R}_i 或平均逆序等級 \bar{R}_{m-i+1} 之間的線性關係為：

$$RI_i = \frac{2}{m} \left(1 - \frac{\bar{R}_i}{m + 1} \right) \quad (5)$$

$$RI_i = \frac{2\bar{R}_{m-i+1}}{m(m + 1)} \quad (6)$$

三、有關的統計推論

相對重要度指數 RI_i 是無母數統計學的一個統計量，據此可作的統計推論擇要分述如下：

(一) m 個項目或處理之重要性的檢定

欲檢定 m 個項目或處理之重要性是否有顯著差異，假設的建立可為：

$$\begin{cases} H_0 : m \text{ 個項目 (處理) 之重要性一致} \\ H_1 : m \text{ 個項目 (處理) 之重要性不一致} \end{cases}$$

檢定統計量可依 Friedman 二因子分類變異數分析 [4] 之公式推演而得：

$$\chi_r^2 = 3n (m+1) \left(m - \sum_{i=1}^m RI_i^2 - 1 \right) \quad (7)$$

當 $m \geq 3$, $n > 9$, χ_r^2 分配接近卡方分配。設 $\chi^2_{(1-\alpha, m-1)}$ 為由顯著水準 α 及自由度 $m-1$ 所決定之臨界值，則當 $\chi_r^2 > \chi^2_{(1-\alpha, m-1)}$ 時，拒絕 H_0 ，表示 m 個項目（或處理）之重要性可能不一致。自由度喪失一度，乃因 $\sum_{i=1}^m RI_i = 1$ 的緣故。

本檢定的理論基礎：當 m 個項目或處理之重要性一致，則相對重要度指數 RI_i 應相去不多，使 χ_r^2 小；故當 RI_i 相去甚多， χ_r^2 大，則應拒絕重要性一致的虛無假設。

(二) m 個項目或處理之重要性完全不同的檢定

經 m 個項目或處理之重要性是否可能一致檢定後，若所得結論為否定的，則可進一步作事後比較，即進行多重比較（multiple comparisons），以決定 $\binom{m}{2} = \frac{m(m-1)}{2}$ 組比較是否可能完全不同。此多重比較之假設建立為：

$$\begin{cases} H_0 : \text{第 } i, k \text{ (} i \neq k \text{) 個項目 (處理) 之重要性相同} \\ H_1 : \text{第 } i, k \text{ 個項目 (處理) 之重要性不同} \end{cases}$$

依多組有關樣本之多重比較檢定法 [5]，推得檢定之決策法則為：

$$|RI_i - RI_k| > Z_{(1-\frac{\alpha}{m(m-1)})} \sqrt{\frac{2}{3nm(m+1)}} \quad (8)$$

[4] W. W. Daniel (1979), "Applied Nonparametric Statistics", pp. 224-230.

[5] 同 [4], pp. 231-232.

則拒絕 H_0 ，表示第 i 及 k 個項目或處理之重要性有可能不同。式中 RI_i ， RI_k 分別為第 i ， k 個項目或處理的相對重要度指數， Z 之下標 $1 - \frac{\alpha}{m(m-1)}$ 是因考量 $\binom{m}{2}$ 組比較的緣故，即

$$1 - \left[\frac{\alpha}{2} \bigg/ \frac{m(m-1)}{2} \right] = 1 - \frac{\alpha}{m(m-1)} \quad (9)$$

(三) m 個項目或處理之重要性遞減的檢定

若欲檢定 m 個項目或處理之重要性是否可能有遞減的情況，假設的建立為：

$$\begin{cases} H_0 : m \text{ 個項目 (處理) 之重要性一樣} \\ H_1 : m \text{ 個項目 (處理) 之重要性遞減} \end{cases}$$

檢定統計量可借用 Page 檢定 [6] 的觀念，衍繹而為：

$$L = \frac{nm(m+1)}{2} \left[(m+1) - \sum_{i=1}^m iRI_i \right] \quad (10)$$

式中 $i = 1, 2, \dots, m$ 分別為 m 項相對重要度指數 RI_i 之權數。若 m 及 n 夠大，則 L 分配趨近於常態分配，期待值 $E(L)$ 及變異數 $V(L)$ 分別為：

$$E(L) = \frac{nm(m+1)^2}{4} \quad (11)$$

$$V(L) = \frac{n(m^3 - m)^2}{144(m-1)} \quad (12)$$

取 Z 統計量：

$$Z = \frac{L - E(L)}{\sqrt{V(L)}} \quad (13)$$

令顯著水準為 α ，則 $Z > Z_{(1-\alpha)}$ 時，拒絕 H_0 ，表示 m 個項目或處理的重要性可能有遞減情況。

本檢定的理論基礎：若各項目或處理的重要性一致，則各相對重要度指

[6] M. Hollander, D. A. Wolfe (1973), "Nonparametric Statistical Methods", pp. 147-150.

數相去不遠；換言之，若重要性遞減的情況存在，則 RI_1 有大於 RI_2 的趨向， RI_2 有大於 RI_3 的趨向，依此類推，再加上權數 1, 2, ...m 的加權，則使統計量 L 加大。故當 L 大時，應該拒絕重要性一致的虛無假設。

(四)相對重要度指數母數 P_i 之估計

由於相對重要度指數 RI_i 之和 $\sum_{i=1}^m RI_i = 100\% = 1$ ，表示各項目或處理之 RI_i 為一分量，可比照樣本比例 \hat{P} 依 Clopper-Pearson 法 [7] 進行任一項目或處理之相對重要度指數母數 P_i 的估計。即在樣本大小 $n > 30$ 的條件下，母數 P_i 之 $1 - \alpha$ 信賴度的近似信賴區間為：

$$RI_i - Z_{(1-\frac{\alpha}{2})} \sqrt{\frac{RI_i(1-RI_i)}{n}} \leq P_i \leq RI_i + Z_{(1-\frac{\alpha}{2})} \sqrt{\frac{RI_i(1-RI_i)}{n}}$$

式中 $Z_{(1-\frac{\alpha}{2})}$ 為 $1 - \frac{\alpha}{2}$ 機率所決定之標準常態值。

四、實證研究

一包裝決定以黃綠二色印製，但比例該如何？乃進行一包裝滿意度實驗設計，A、B、C 分別代表綠色部分所佔比例為 $\frac{1}{4}$ 、 $\frac{1}{2}$ 、 $\frac{3}{4}$ 者，經抽訪 30 位超市顧客，得下表：

顧客		1	2	3	4	5	6	7	8	9	10	11	12	13	14
等級 R_{ij}	A	1	1	1	2	1	1	1	2	1	2	1	2	1	2
	B	2	3	2	1	3	3	2	1	2	3	2	1	3	1
	C	3	2	3	3	2	2	3	3	3	1	3	3	2	3

15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	1	1	2	2	1	2	1	1	1	1	1	1	3	2	2
3	2	2	3	1	3	1	3	2	2	3	2	2	1	1	3
2	3	3	1	3	2	3	2	3	3	2	3	3	2	3	1

[7] 同 [6], pp. 23-24.

$m = 3, n = 30$, A、B、C 之等級和分別為 42、63、75 代入公式(4)，可得相對重要度指數分別為 $RI_1 = 0.4333$, $RI_2 = 0.3167$, $RI_3 = 0.2500$ 。進一步可進行以下的無母數統計推論：（令 $\alpha = 0.05$ ）

$$\textcircled{1} \begin{cases} H_0 : \text{顧客對包裝 A、B、C 的喜愛一樣} \\ H_1 : \text{顧客對包裝 A、B、C 的喜愛不一樣} \end{cases}$$

$$\chi_r^2 = 3(30)(3+1)[3(0.4333^2 + 0.3167^2 + 0.2500^2) - 1] \\ = 18.5914 > \chi^2_{(0.95,2)} = 5.99$$

差異顯著，故拒絕 H_0 ，表示顧客對包裝 A、B、C 的喜愛可能不一樣。

$$\textcircled{2} \begin{cases} H_0 : \text{顧客對 A、B (或 A、C 或 B、C) 的喜愛一樣} \\ H_1 : \text{顧客對 A、B (或 A、C 或 B、C) 的喜愛不一樣} \end{cases}$$

已知 $\alpha = 0.05$ ， $m = 3$ ， $1 - \frac{0.05}{3(3-1)} = 0.9917$ ， $Z(0.9917) = 2.395$ ，
臨界值為：

$$2.395 \sqrt{\frac{2}{3(30)(3)(3+1)}} = 0.1031$$

三組比較分別為：

$$A \text{ 與 } B \quad | \quad 0.4333 - 0.3167 \quad | \quad = 0.1166 > 0.1031$$

$$A \text{ 與 } C \quad | \quad 0.4333 - 0.2500 \quad | \quad = 0.1833 > 0.1031$$

$$B \text{ 與 } C \quad | \quad 0.3167 - 0.2500 \quad | \quad = 0.0667 < 0.1031$$

A 與 B 比較為差異顯著，A 與 C 比較為差異顯著，B 與 C 比較為差異不顯著，故知顧客對包裝 A、B、C 的喜愛可能不是完全不一樣。

$$\textcircled{3} \begin{cases} H_0 : \text{顧客對包裝 A、B、C 的喜愛一樣} \\ H_1 : \text{至少有一種包裝較被偏愛} \end{cases}$$

$$L = \frac{30(3)(3+1)}{2} \{ (3+1) - [1(0.4333) + 2(0.3167) \\ + 3(0.2500)] \} = 392.9940$$

$E(L) = 360$ ， $V(L) = 60$ ，則

$$Z = \frac{392.9940 - 360}{\sqrt{60}} = 4.2595 > Z_{(0.95)} = 1.645$$

差異顯著，故拒絕 H_0 ，表示顧客對包裝 A、B、C 的喜愛可能不是一樣的，而至少有一種較被偏愛。

④ 綜合前三種檢定可知，顧客最喜愛的包裝可能是綠色部分占 $\frac{1}{4}$ 的 A 包裝。至於所有超市顧客喜愛 A 包裝的相對比例 P 有多大？則可進行 P 之近似區間的估計為：

$$0.4333 - 1.96 \sqrt{\frac{0.4333(1-0.4333)}{30}} \leq P \leq 0.4333 + 1.96 \sqrt{\frac{0.4333(1-0.4333)}{30}}$$

$$\text{得} \quad 0.2560 \leq P \leq 0.6106$$

即所有超市顧客喜愛 A 包裝的相對比例，亦即 A 包裝的相對重要度指數母數 P 介於 25.60% 與 61.06% 之間的機率為 95%。

五、結論

相對重要度指數 RI_i 是無母數統計學一個新統計量，各項目或處理之 RI_i 與其重要程度方向一致，便於解釋且簡單易懂。

由樣本資料求得各項目或處理的相對重要度指數 RI_i 後，進行 χ^2_r 檢定以判斷各項目或處理的重要性是否一致，進行多重比較以判斷各項目或處理的重要性是否完全不同，進行 L 檢定以判斷各項目或處理的重要性是否有遞減的情況存在，最後再進行近似信賴區間的估計以決定母數 P_i ，如此，則可對研究的課題提供清晰的結論。

相對重要度指數能對各研究項目或處理的重要程度提供簡易合理的解釋，例如比較顧客對各種包裝的喜愛次序，遊客對各種解說方式的偏好，兒童對各種遊憩設施的需求程度……等，皆可使用此統計量 RI_i 表示樣本資料

中之各項目或處理的重要性。再與一些多組有關樣本的無母數統計推論方法配合使用，在某一顯著水準 α 或信賴度 $1 - \alpha$ 的條件下，可知顧客對包裝及商品等，遊客對解說及行程等，投資者對股票及房地產等之喜愛、偏好及需求的程度與順序，故能提供行銷、遊憩、教育、財務管理……等各行各業之經營管理者較有價值的訊息。

DISCUSSION OF RELATIVE IMPORTANCE INDEX

Yen Yueh-chu

ABSTRACT

This paper discussed a statistic — relative importance index $RI_i =$

$$\frac{\sum_{j=1}^n (m - R_{ij} + 1)}{\sum_{i=1}^m \sum_{j=1}^n (m - R_{ij} + 1)}$$

and its inferences, for questionnaires requesting respondents

$$\sum_{i=1}^m \sum_{j=1}^n (m - R_{ij} + 1)$$

to rank items.