

# 行政院國家科學委員會專題研究計畫 成果報告

## 評估動態內部一致性之修正 alpha 係數

計畫類別：個別型計畫

計畫編號：NSC94-2416-H-009-012-

執行期間：94年08月01日至95年07月31日

執行單位：國立交通大學經營管理研究所

計畫主持人：丁承

計畫參與人員：鄭天德

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 95 年 11 月 2 日

# 行政院國家科學委員會專題研究計畫成果報告

評估動態內部一致性之修正  $\alpha$  係數

## A Modification of Coefficient Alpha for Assessing Dynamic Internal Consistency

計畫編號：NSC 94-2416-H-009-012

執行期限：94 年 8 月 1 日至 95 年 7 月 31 日

計畫主持人：丁 承 國立交通大學經營管理研究所

計畫參與人：鄭天德 國立交通大學經營管理研究所

### 一、中文摘要

$\alpha$  係數係用於測量某一個構念之衡量問項在同一時點之內部一致性，Green (2003) 提出一再測  $\alpha$  係數，藉以量測前後二不同時間點不同問項間之一致性程度，然而該再測  $\alpha$  係數係基於單變量的觀點求取。在多變量的結構下，可採驗證型因素分析計算構念之複合信度，但跨期複合信度仍未見探討。本研究參考 Green 的概念，進一步基於多變量的觀點，提出評估某一構念各量測問項跨期內部一致性之動態複合信度指標，可更適當地反映跨期信度，蒙地卡羅模擬顯示該指標具良好性質，且計算簡易，在動態信度評估上具應用價值。

**關鍵詞：** $\alpha$  係數、複合信度、驗證型因素分析、構念、動態複合信度、內部一致性、蒙地卡羅模擬、再測  $\alpha$  係數

### Abstract

Coefficient alpha is a reliability index to assess internal consistency of the items used to measure the same construct. Green (2003) proposed a test-retest alpha to assess internal consistency across two different time points. However, the test-retest alpha is based on the univariate viewpoint. Under the multivariate data structure, composite reliability resulting from confirmatory factor analysis can be used, but needs to be further studied for the test-retest situation. In this research, a dynamic composite reliability

index is proposed for evaluating internal consistency for test-retest data. The index, based on Green's idea as well as multivariate viewpoint, can more appropriately reflect dynamic reliability. Monte-Carlo simulation results indicated that the index performs well. It is easy to compute, and is useful for empirical studies.

**Keywords:** coefficient alpha, composite reliability, confirmatory factor analysis, construct, dynamic composite reliability, internal consistency, Monte-Carlo simulation, test-retest alpha

### 二、緣由與目的

$\alpha$  係數 (coefficient alpha) (Cronbach, 1951) 係用於測量某一個構念 (construct) 之衡量問項 (items) 在同一時間點 (同期) 之內部一致性 (internal consistency) 程度。近期關於  $\alpha$  係數之探討如 Barchard & Hakstian (1997)、Bonett (2003)、Enders (2003)、Enders & Bandalos (1999)、Feldt & Charter (2003)、Green & Hershberger (2000)、Hakstian & Barchard (2000)、Komaroff (1997)、Miller (1995)、Osburn (2000)、Raykov (1997, 1998, 2001, 2004)、Raykov & Shrout (2002)、Schmitt (1996) 及 Shevlin et al. (2000) 等，均僅係針對同期問項間的內部一致性進行討論。至於跨期之信度實證上，大多以再測信度 (即 test-retest correlation) 為指標，如 Hawks et al. (2004)、Richardson et al. (2003)、Stahl et al. (1982) 及 Kinicki et al. (2002) 等。

由於  $\alpha$  係數未具評估跨期內部一致性信度之功能，故跨期內部一致性之探討已陸續出現，Becker (2000) 及 Vautier & Jmel (2003) 將受測者分成兩群，並採用跨期交叉折半信度為基礎計算信度。Green (2003) 針對前後二不同時間點 ( $t = 1$  及  $t = 2$ ) 不同問項間之一致性程度提出一量測指標，稱為再測  $\alpha$  係數 (test-retest  $\alpha$ ) (本文以  $G-\alpha$  表之)。舉例言之，如圖1所示，設某構念  $F$  有四個量測問項，在時間  $t = 1$  (test) 及  $t = 2$  (retest) 之前後因素分別以  $F_1$  及  $F_2$  表之，其量測變項則分別以  $V_{11}, V_{21}, V_{31}, V_{41}$  及  $V_{12}, V_{22}, V_{32}, V_{42}$  表之，該八變項間可計算其共變數矩陣如圖2 (Green, 2003)，該圖左上角或右下角次矩陣之主對角線係反映同一時點相同問項之變異數；非主對角線元素則反映同一時點不同問項之共變數，據此可計算各特定時點之  $\alpha$  係數。圖2左下角次矩陣之主對角線係反映相同問項於不同時間之共變數；非主對角線元素則反映不同時間不同問項之共變數，此即  $G-\alpha$  公式分子項之形成依據，然而， $G-\alpha$  係基於單變量的觀點求取。依圖1之資料結構所示，我們可基於多變量的觀點進行驗證型因素分析 (confirmatory factor analysis, CFA)，並計算  $F_1$  及  $F_2$  之複合信度 (composite reliability) (Fornell & Larcker, 1981) 以反映各期內問項間之一致性，但如何求取跨期之複合信度在文獻上卻未見探討，故本研究擬參考 Green 的概念，進一步基於多變量的觀點提出一跨期複合信度指標。

複合信度公式係基於標準化後之因素負荷量表達，相當於以標準化後之資料進行 CFA，即以量測問項之相關係數矩陣 (可由圖 2 之共變數矩陣獲得，如圖 3 所示) 投入，並求得各參數之估計值。相對於  $G-\alpha$ ，本研究所提出之再測複合信度應較能反映多變量之資料結構，評估跨期資料之內部一致性較具客觀性。

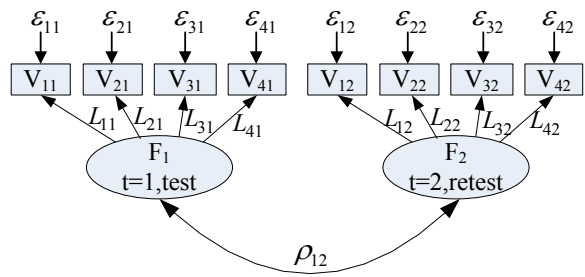


圖1. 構念  $F$  及其量測問項之 test-retest 示意圖

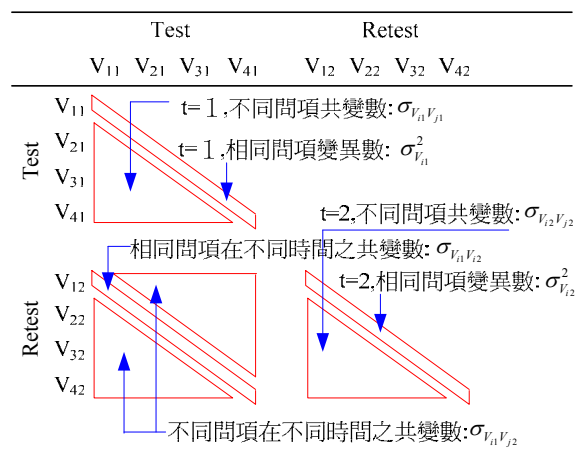


圖 2. 圖 1 結構下量測變項共變數矩陣示意圖 (Green, 2003)

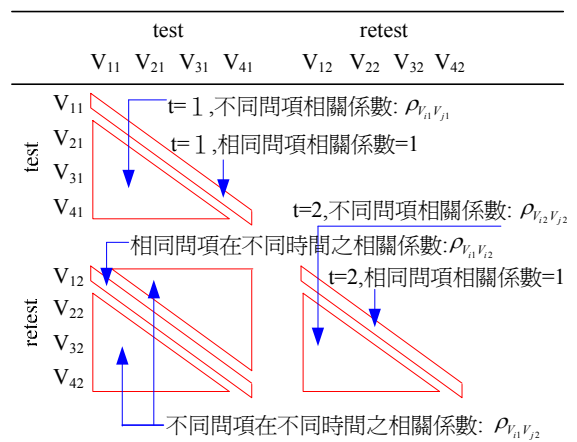


圖 3. 對應於圖 2 之相關係數矩陣示意圖

### 三、結果與討論

驗證型因素分析模式可表示如下：

$$\begin{aligned}
 V_{11} &= L_{11}F_1 + \varepsilon_{11}, \\
 V_{21} &= L_{21}F_1 + \varepsilon_{21}, \\
 &\vdots \\
 V_{K1} &= L_{K1}F_1 + \varepsilon_{K1}, \\
 V_{12} &= L_{12}F_2 + \varepsilon_{12}, \\
 V_{22} &= L_{22}F_2 + \varepsilon_{22}, \\
 &\vdots \\
 V_{K2} &= L_{K2}F_2 + \varepsilon_{K2},
 \end{aligned} \tag{1}$$

其中  $K$  表問項之個數。我們假設誤差項  $\varepsilon_{ij}$  間無關，與  $F_1, F_2$  亦無關， $\text{Var}(\varepsilon_{ij}) = \sigma_{\varepsilon_{ij}}^2$ ， $\text{Var}(F_1) = \text{Var}(F_2) = 1$ ， $F_1$  和  $F_2$  之相關係數以  $\rho_{12}$  表之，因素負荷量  $L_{it}$  表第  $t$  期之因素與其對應問項  $V_{it}$  之相關係數。當  $K=4$ ，模式(1)即與圖 1 相對應。在 CFA 各參數估計完成後，即可據以計算  $t=1$  及  $t=2$  之複合信度，分別以  $CR_1$  及  $CR_2$  表示，公式如下：

$$CR_1 = \frac{\sum_{i=1}^K \hat{L}_{i1}^2}{\left(\sum_{i=1}^K \hat{L}_{i1}\right)^2 + \sum_{i=1}^K \hat{\sigma}_{\varepsilon_{i1}}^2}.$$

當資料標準化後， $\rho_{V_{i1}V_{j1}} = \text{Cov}(L_{i1}F_1 + \varepsilon_{i1}, L_{j1}F_1 + \varepsilon_{j1}) = L_{i1}L_{j1}$ ，故

$$\begin{aligned}
 CR_1 &= \frac{\sum_{i=1}^K \hat{L}_{i1}^2 + \sum_{i \neq j} \hat{\rho}_{V_{i1}V_{j1}}}{\sum_{i=1}^K \hat{L}_{i1}^2 + \sum_{i=1}^K \hat{\sigma}_{\varepsilon_{i1}}^2 + \sum_{i \neq j} \hat{\rho}_{V_{i1}V_{j1}}} \\
 &= \frac{K \bar{\hat{L}}_1^2 + K(K-1)\bar{\hat{\rho}}_1}{K + K(K-1)\bar{\hat{\rho}}_1} \\
 &= \frac{\bar{\hat{L}}_1^2 + (K-1)\bar{\hat{\rho}}_1}{1 + (K-1)\bar{\hat{\rho}}_1}.
 \end{aligned} \tag{2}$$

同理可得

$$CR_2 = \frac{\bar{\hat{L}}_2^2 + (K-1)\bar{\hat{\rho}}_2}{1 + (K-1)\bar{\hat{\rho}}_2}. \tag{3}$$

上面  $CR_1$  及  $CR_2$  之式(2) 及式(3) 中， $\bar{\hat{L}}_t^2 = \sum_{i=1}^K \hat{L}_{it}^2 / K$ ， $\bar{\hat{\rho}}_t = \sum_{i \neq j} \hat{L}_{it} \hat{L}_{jt} / (K(K-1))$ ， $t=1, 2$ ， $\bar{\hat{\rho}}_t$  即第  $t$  期不同問項間相關係數估計值的平均數。

有關跨期間項之相關，可藉因素負荷量與  $\rho_{12}$  之乘積表達如下：

$$\rho_{V_{i1}V_{j2}} = \text{Cov}(L_{i1}F_1 + \varepsilon_{i1}, L_{j2}F_2 + \varepsilon_{j2}) = L_{i1}L_{j2}\rho_{12} \tag{4}$$

至於跨期負荷，則為

$$\begin{aligned}
 \rho_{F_1, V_{i2}} &= \text{Cov}(F_1, L_{i2}F_2 + \varepsilon_{i2}) = L_{i2}\rho_{12}, \\
 \rho_{F_2, V_{i1}} &= \text{Cov}(F_2, L_{i1}F_1 + \varepsilon_{i1}) = L_{i1}\rho_{12}
 \end{aligned} \tag{5}$$

依據式(2)、(3)、(4)及(5)，並參考 Green (2003) 的概念，可基於多變量的觀點，進一步發展出一評估跨期內部一致性之動態複合信度 (dynamic composite reliability) 指標  $DCR$  如下：

$$\begin{aligned}
 DCR &= \frac{K \bar{\hat{L}}^2(2) + K(K-1)\bar{\hat{\rho}}(2)}{\sqrt{K + K(K-1)\bar{\hat{\rho}}_1} \sqrt{K + K(K-1)\bar{\hat{\rho}}_2}} \\
 &= \frac{\bar{\hat{L}}^2(2) + (K-1)\bar{\hat{\rho}}(2)}{\sqrt{1 + (K-1)\bar{\hat{\rho}}_1} \sqrt{1 + (K-1)\bar{\hat{\rho}}_2}},
 \end{aligned} \tag{6}$$

上式分子項中之  $\bar{\hat{\rho}}(2)$  表跨 2 期(即期差 1 期)不同問項間相關係數的平均數，可依 CFA 所得因素負荷量及  $F_1, F_2$  相關係數估計值求算之：

$$\bar{\hat{\rho}}(2) = \frac{\sum_{i \neq j} \hat{\rho}_{V_{i1}V_{j2}}}{K(K-1)} = \frac{\sum_{i \neq j} \hat{L}_{i1} \hat{L}_{j2} \hat{\rho}_{12}}{K(K-1)},$$

另  $\bar{\hat{L}}^2(2) = (\sum_{t=1}^2 \sum_{i=1}^K \hat{L}_{it}^2 \hat{\rho}_{12}^2) / 2K$  表全部問項被不同期因素跨期解釋的平均能力，括弧中數字仍表跨 2 期；上式分母項取第 1 期  $CR_1$

分母項與第 2 期  $CR_2$  分母項之幾何平均係參考 Green (2003) 的作法。

本研究採用蒙地卡羅模擬 (Monte-Carlo simulation) 評估  $DCR$ ，以模式 (1) 為模擬架構。我們設  $F_1 \sim N(0, 1)$  (標準常態)，並就不同  $\rho_{12}$  之參數值 (0.8, 0.6, 0.4) 以下列方式產生  $F_2$  (管中閔, 2000, 14.3 節)：

$$F_2 = \rho_{12} \times F_1 + \sqrt{1 - \rho_{12}^2} \times Z,$$

其中  $Z$  為與  $F_1$  獨立的標準常態隨機變數，而  $F_2$  亦為標準常態。接著就不同問項數 ( $K = 3, 4, 5, 6$ )，依各不同因素負荷量  $L_{ij}$  之參數設定值 (設於 0.4 與 0.9 之間，取數種組和)，在不同樣本數 (50, 100, 200) 下分別產生  $t=1$  及  $t=2$  之各問項資料。由於各組合下之母體  $DCR$  皆可依各參數設定值算出，故可據以評估樣本  $DCR$  的性質，模擬結果顯示樣本  $DCR$  皆近似不偏，且標準誤 (隨樣本數之增加而下降) 小，文末附表例示當  $K = 4$  時各設定組合下母體  $DCR$  與樣本  $DCR$  模擬 2000 次之平均值與標準誤。

最後以 Green (2003) 之附錄資料為例計算  $DCR$ ，並與標準化再測  $\alpha$  係數進行比較。該例之  $K = 4$ ，藉由 SAS PROC CALIS 所得 CFA 相關估計結果，可知  $\bar{\hat{\rho}}_1 = 0.61528$ ,  $\bar{\hat{\rho}}_2 = 0.7737$ ,  $\hat{\rho}_{12} = 0.81355$ ,  $\bar{\hat{\rho}}(2) = 0.563$ ,  $\bar{L}^2(2) = 0.4676$ ，以上代入式 (6) 即可獲得  $DCR = 0.7015$ 。接著計算標準化再測  $\alpha$  係數  $G-\alpha$ ，其公式如下：

$$G-\alpha = \frac{K \bar{\hat{\rho}}^*(2)}{\sqrt{1 + (K-1) \bar{\hat{\rho}}_1^*} \sqrt{1 + (K-1) \bar{\hat{\rho}}_2^*}}, \quad (7)$$

式中  $\bar{\hat{\rho}}_i^*$  仍表示第  $t$  期不同問項間相關係數估計值的平均數，但估計方式與前不同，係直接採用圖 3 左上角 ( $t = 1$ ) 與右下角 ( $t = 2$ ) 次矩陣之下三角矩陣元素的平均數； $\bar{\hat{\rho}}^*(2)$  亦為跨 2 期不同問項間相關係數估計值的平均數，係逕採圖 3 左下角次矩

陣非主對角線元素之平均。所得之  $\bar{\hat{\rho}}_1^* = 0.6058$ ,  $\bar{\hat{\rho}}_2^* = 0.7726$ ,  $\bar{\hat{\rho}}^*(2) = 0.55826$ ，標準化  $G-\alpha = 0.7304$ 。比較結果顯示  $DCR$  略小於  $G-\alpha$ 。

當研究架構中有多個構念同時進行再測時，由於構念間具相關性，故更應基於多變量的觀點同時進行參數校估，此時採 CFA 尤可顯現效益，並充分提供求算  $DCR$  所需的資訊。

#### 四、計畫成果自評

本研究參考 Green (2003) 再測  $\alpha$  係數的概念，提出評估某一構念各量測問項跨期內部一致性之動態複合信度指標  $DCR$ ，因係基於多變量的觀點，故能更適當地反映跨期信度，且統計性質佳，計算簡易，在動態信度評估上具應用價值。本文預定投稿至學術期刊如 *Psychological Methods*。

#### 五、參考文獻

- 管中閔 2000. 統計學：觀念與方法，台北市：華泰文化事業公司。
- Barchard, K. A., & Hakstian, A. R. 1997. The effects of sampling model on inference with coefficient alpha. *Educational & Psychological Measurement*, 57: 893–905.
- Becker, G. 2000. Coefficient alpha: Some terminological ambiguities and related misconceptions. *Psychological Reports*, 86: 365–372.
- Becker, G. 2000. How important is transient errors in estimating reliability? Going beyond simulation studies. *Psychological Methods*, 5: 370–379.
- Bonett, D. G. 2003. Sample size requirement for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics*, 27: 335–340.
- Cronbach, L.J. 1951. Coefficient alpha and internal alpha and the internal structure of tests. *Psychometrika*, 16: 297–334.

- Enders, C. K. 2003. Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods*, 8: 322–337.
- Enders, C. K., & Bandalos, D. L. 1999. The effects of heterogeneous item distributions on reliability. *Applied Measurement in Education*, 12: 133-150.
- Feldt, L. S., & Charter, R. A. 2003. Estimating the reliability of a test split into two parts of equal or unequal length. *Psychological Methods*, 8: 102–109.
- Fornell, C., & Larcker, D. F. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18: 39-50.
- Green, S. B. 2003. A coefficient alpha for test-retest data. *Psychological Methods*, 8: 88-101.
- Green, S. B., & Hershberger, S. L. 2000. Correlated errors in true score models and their effect on coefficient alpha. *Structural Equation Modeling*, 7: 251–270.
- Hawks, S., Merrill, R. M., & Madanat, H. N. 2004. The intuitive eating scale: development and preliminary validation. *American Journal of Health Education*, 35: 90–99.
- Kinicki, A. J., McKee-Ryan, F. M., Schriesheim, C. A., & Carson, K. P. 2002. Assessing the construct validity of the job descriptive index: a review and meta-analysis. *Journal of Applied Psychology*, 87(1): 14–32.
- Komaroff, E. 1997. Effect of simultaneous violations of essential tau-equivalence and uncorrelated error on coefficient alpha. *Applied Psychological Measurement*, 21: 337–348.
- Miller, M. B. 1995. Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, 2: 255–273.
- Osburn, H. G. 2000. Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, 5: 343–355.
- Raykov, T. 1997. Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32: 329–353.
- Raykov, T. 1998. Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement*, 22: 375–385.
- Raykov, T. 2001. Bias of coefficient  $\alpha$  for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25: 69–76.
- Raykov, T. 2004. Point and interval estimation of reliability for multiple-component measuring instruments via linear constraint covariance structure modeling. *Structural Equation Modeling*, 11: 342–356.
- Raykov, T., & Shrout, P. E. 2002. Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling*, 9: 195–212.
- Richardson, B., Kitchen, G., & Livingston, G. 2004. Developing the KAMA instrument (knowledge and management of abuse). *Age and Ageing*, 32: 286-291.
- Schmitt, N. 1996. Uses and abuses of coefficient alpha. *Psychological Assessment*, 8: 350-353.
- Shevlin, M., Miles, J. N. V., Davies, M. N. O., & Walker, S. 2000. Coefficient alpha: A useful indicator of reliability? *Personality & Individual Differences*, 28: 229–237.
- Stahl, M. J. & Harrell, A. M. 1982. Evolution and validation of a behavioral decision theory measurement approach to achievement, power, and affiliation. *Journal of Applied Psychology*, 67: 744–751.
- Vautier, S., & Jmel, S. 2003. Transient error or specificity? An alternative to the staggered equivalent split-half procedure. *Psychological Methods*, 8: 225–238.

附表.  $K=4$  時各設定組合下母體  $DCR$  與樣本  $DCR$  模擬 2000 次之平均值與標準誤

Case 1: $L_{11} = 0.6, L_{21} = 0.7, L_{31} = 0.8, L_{41} = 0.9; L_{12} = 0.7, L_{22} = 0.6, L_{32} = 0.5, L_{42} = 0.4$									
$\rho_{12}$	0.8			0.6			0.4		
母體 $DCR$	0.571			0.404			0.254		
樣本數	50	100	200	50	100	200	50	100	200
樣本 $DCR$ 平均值 (標準誤)	0.566 (0.099)	0.570 (0.069)	0.569 (0.050)	0.400 (0.121)	0.407 (0.081)	0.405 (0.058)	0.264 (0.118)	0.256 (0.087)	0.252 (0.061)
Case 2: $L_{11} = 0.5, L_{21} = 0.5, L_{31} = 0.6, L_{41} = 0.6; L_{12} = 0.4, L_{22} = 0.4, L_{32} = 0.5, L_{42} = 0.5$									
$\rho_{12}$	0.8			0.6			0.4		
母體 $DCR$	0.432			0.306			0.193		
樣本數	50	100	200	50	100	200	50	100	200
樣本 $DCR$ 平均值 (標準誤)	0.405 (0.121)	0.421 (0.087)	0.433 (0.063)	0.308 (0.125)	0.308 (0.096)	0.308 (0.065)	0.220 (0.119)	0.200 (0.091)	0.195 (0.065)