

如何利用變異數成份估計一致性相關係數

蔡秒玉 彰化師範大學統計資訊所/副教授

一致性相關係數的定義

在生物醫學研究中，若有一個新的測量方法要與黃金標準或是早期慣用的測量方法進行比較時，我們希望新的方法與黃金測量或是早期慣用的方法兩者之間具有相當程度的一致性(agreement)，這樣才能證明新的測量法具有一定的準確性。若針對類別型資料(categorical data)，最廣為大家接受的是 kappa 統計量，其中 Cohen's kappa 及 weighted kappa 分別是使用在二元(binary data)及有序的判讀資料(ordinal data)上(Cohen, 1960; 1968)。當判讀結果為連續型資料時，組內相關係數(Intraclass correlation coefficient; ICC)則是常用來評估家庭或集群判讀資料組內的相關性大小，亦即用來反映連續資料上的一致性(Bartko, 1966; Shrout and Fleiss, 1979)；另一個針對連續型資料，可用來估計兩種判讀方法或儀器的一致性指標為一致性相關係數(concordance correlation coefficient; CCC)，其為測量任兩對判讀資料相距於經原點的 45 度線的變異(variation of the linear relationship between each pair of data from the 45° line through the origin) (Lin, 1989)。此一致性相關係數有兩方面的優點，其第一項優點為可以評估每一筆判讀資料離所配適的迴歸線多遠，此即代表精確度(precision)，另一項優點為評估所配適的迴歸線離經原點的 45 度線多遠，此即代表準確度(accuracy) (King and Chinchilli, 2001)。本文章主要是針對一致性相關係數的估計進行探討。

根據 Lin (1989)提出針對兩個連續型變數 Y_1 和 Y_2 的一致性相關係數定義如下：

$$\rho_c = 1 - \frac{E[(Y_1 - Y_2)^2]}{E[(Y_1 - Y_2)^2 | Y_1 \text{ and } Y_2 \text{ are uncorrelated}]} = \frac{2\sigma_{12}}{\sigma_1^2 + \sigma_2^2 + (\mu_1 - \mu_2)^2},$$

其中 $\mu_1 = E(Y_1)$ ， $\mu_2 = E(Y_2)$ ， $\sigma_1^2 = Var(Y_1)$ ， $\sigma_2^2 = Var(Y_2)$ ， $\sigma_{12} = Cov(Y_1, Y_2)$ 。由

此定義可知， ρ_c 的值域落在 -1 與 1 之間，若為 -1 則表示兩儀器之間完全不一致 (perfect disagreement)；為 0 則表示彼此間互相獨立；為 1 則表示兩者之間完全一致 (perfect agreement)。當 $\mu_1 = \mu_2$ 與 $\sigma_1^2 = \sigma_2^2$ 時， ρ_c 則相當於皮爾森相關係數 (Pearson correlation coefficient)；而一致性相關係數的 95% 的信賴區間可用

Fisher's Z 變數變換得到 (Fisher's Z transformation)，即 $\hat{Z}_c = \frac{1}{2} \ln \left(\frac{1 + \hat{\rho}_c}{1 - \hat{\rho}_c} \right)$ 會近似常

態分配，其中 $\hat{\rho}_c$ 代表 ρ_c 的估計式，因此一致性相關係數的 95% 的信賴區間可經由 $\hat{Z}_c \pm 1.96 \sqrt{\text{Var}(\hat{Z}_c)}$ 轉換得到，其中 $\text{Var}(\hat{Z}_c) = \frac{\text{Var}(\hat{\rho}_c)}{(1 - \hat{\rho}_c^2)^2}$ (Lin, 1989)。

利用變異數成份進行一致性相關係數的估計

Carrasco & Jover (2003) 提出了在線性混合模式 (linear mixed model; LMM) 下，利用變異數成份 (variance components; VC) 估計一致性相關係數，其模式假設為

$$Y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \quad (1)$$

其中 Y_{ij} 代表第 i 個個體 (subject) 被第 j 個儀器 (observer) 所測量的觀測值； μ 為總體平均； α_i 是第 i 個個體的隨機效應 (subject random effect)，令其分配為 $\alpha_i \sim N(0, \sigma_\alpha^2)$ ； β_j 是第 j 個儀器的固定效應 (observer fixed effect)； e_{ij} 是第 i 個個體被第 j 個儀器所測量觀測值的隨機誤差 (random error)，令其分配為 $e_{ij} \sim N(0, \sigma_e^2)$ ， $i = 1, \dots, n$ ， $j = 1, \dots, J$ 。令隨機效應與隨機誤差項彼此之間皆互相獨立，因此在這些前提假設下，Carrasco & Jover (2003) 證明了組內相關係數與一致性相關係數兩種一致性測度估計式相同，則此一致性相關係數可經由變異數成份表示成

$$\rho_c = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2}, \quad (2)$$

$$\text{其中 } \sigma_{\beta}^2 = \frac{1}{J-1} \sum_{j=1}^J \beta_j^2 \text{。}$$

利用變異數成份進行重複測量加權一致性相關係數的估計

針對長期追蹤重複測量資料，Carrasco et al. (2009)則將第(1)式的線性混合模型中加入時間屬性的固定效應及其與個體及判讀方法效應的交互作用，如此可將第(1)式的線性混合模型改寫為

$$Y_{ijt} = \mu + \alpha_i + \beta_j + \gamma_t + \alpha\beta + \alpha\gamma + \beta\gamma_t + e_{ijt} \text{ ,} \quad (3)$$

其中 Y_{ijt} 代表第 i 個個體被第 j 個儀器所測量的第 t 次判讀觀測值； γ_t 是第 t 次時間的固定效應 ($t = 1, 2, \dots, p$)； $\alpha\beta_{ij}$ 是第 i 個個體和第 j 個儀器間交互作用的隨機效應 (random subject-observer interaction effect)，令其分配為 $\alpha\beta_{ij} \sim N(0, \sigma_{\alpha\beta}^2)$ ； $\alpha\gamma_{it}$ 是第 i 個個體和第 t 次測量時間交互作用的隨機效應 (random subject-time interaction effect)，令其分配為 $\alpha\gamma_{it} \sim N(0, \sigma_{\alpha\gamma}^2)$ ； $\beta\gamma_{jt}$ 是第 j 個儀器和第 t 次測量時間交互作用的固定效應 (fixed observer-time effect)； e_{ijt} 是第 i 個個體第 j 個儀器第 t 次測量的隨機誤差項 (random error)。同時假設所有隨機效應 α_i 、 $\alpha\beta_{ij}$ 、 $\alpha\gamma_{it}$ 與隨機誤差 e_{ijt} 皆互相獨立 (mutually independent)。

除此之外，為了考慮在不同時間點觀測值之間的相關程度，Carrasco et al. (2009)依循 King et al. (2007)所提出的加權方法，在任一個時間點測量的觀測值給予權重，其方法即為將一致性相關係數的估計式中加入一個權重矩陣 \mathbf{D} 在重複的觀測值上，若 \mathbf{D} 為一個單位矩陣 (identity matrix)，則組內相關係數與一致性相關係數的估計式相同，即在第(3)式的線性混合模型下的一致性相關係數估計式則可由個體屬性、個體和儀器間交互作用與個體和測量時間交互作用的隨機效應、儀器與測量時間交互作用固定效應以及隨機誤差項的變異數表示。如此，第(2)式的一致性相關係數估計式則可改寫為

$$\rho_{crm} = \frac{\sigma_{\alpha}^2 + \sigma_{\alpha\gamma}^2}{\sigma_{\alpha}^2 + \sigma_{\alpha\gamma}^2 + \sigma_{\alpha\beta}^2 + \sigma_{\beta\gamma}^2 + \sigma_e^2},$$

其中 $\sigma_{\beta\gamma}^2 = \frac{1}{p} \frac{1}{(J-1)} \sum_{t=1}^p \sum_{j=1}^J (\mu_{jt} - \mu_t)^2$ ， μ_{jt} 為第 j 個儀器在第 t 次時間點測量觀測值的總平均， μ_t 為第 t 次時間點測量觀測值的總平均。若 \mathbf{D} 為一個對角矩陣(diagonal matrix)，則此長期追蹤資料之加權重複測量一致性相關係數則可表示成

$$\rho_{crm} = \frac{(\sigma_{\alpha}^2 + \sigma_{\alpha\gamma}^2) \sum_{t=1}^p d_{tt}}{(\sigma_{\alpha}^2 + \sigma_{\alpha\gamma}^2 + \sigma_{\alpha\beta}^2 + \sigma_e^2) \sum_{t=1}^p d_{tt} + \frac{1}{p} \frac{1}{(J-1)} \sum_{t=1}^p \sum_{j=1}^J d_{tt} (\mu_{jt} - \mu_t)^2},$$

其中 d_{ij} 表示為加權矩陣 \mathbf{D} 中第 i 列第 j 行的元素。

利用 R 程式執行一致性相關係數的估計

目前已有 R 軟體的 cccrm 套件可以進行一致性相關係數的估計(Carrasco et al., 2013)，以下皆用 R 軟體內設的重複測量資料作為 R 程式估計一致性相關係數的範例（在附錄中亦有 SAS 程式估計一致性相關係數的介紹）。在此將以 R 軟體提供血壓重複測量資料(bpres)作為分析資料，此資料共有 384 位個案（以 ID 變數標示），每位個案皆被兩種儀器(METHOD)進行舒張壓(DIA)的測量，每種儀器各測量兩次，並記錄每位個案的年齡(AGE)與性別(SEX)兩變數，此資料型態如下表：

ID	AGE	SEX	DIA	METHOD
1	70	2	84	1
1	70	2	92	1
1	70	2	83	2
1	70	2	98	2
2	29	1	70	1
2	29	1	70	1
2	29	1	76	2
2	29	1	76	2
...

針對第(1)式的線性混合模型，可利用 R 軟體的 `ccvc` 函數進行一致性相關係數的估計，而此函數中需放入的參數如下：

1. `dataset`：資料名稱
2. `ry`：反應變數
3. `rind`：個案身份變數
4. `rmet`：儀器方法變數
5. `covar`：個案解釋變數

【ccvc 程式碼】

```
ccvc(dataset=bpres, ry="DIA", rind="ID", rmet="METHOD")
```

【結果】

CCC	LL CI95%	UL CI95%	SE CCC	Z	SE Z
0.818829	0.791561	0.842843	0.013056	1.100336	0.039426
Variance Components:					
Subject	Observer	Random Error			
77.8245	0.0995	17.1197			

此資料的分析結果為兩種血壓測量儀器的一致性程度為 0.8188，95%的信賴區間為(0.7916, 0.8428)，其估計結果顯示兩種血壓測量儀器有高度的一致性，其中個體隨機效應變異數估計值高達 77.8245，代表個人異質性程度高，儀器效應變異數估計值為 0.0995，隨機誤差變異數估計值為 17.1197。

此 `ccrm` 函數亦可以將個案相關的解釋變數加入模式中進行調整，將 `bpres` 資料中的年齡及性別兩變數加入模式後，其程式及分析結果如下：

【ccvc 程式碼】

```
ccvc(dataset=bpres, ry="DIA", rind="ID", rmet="METHOD",  
covar=c("AGE","SEX"))
```

【結果】

CCC	LL CI95%	UL CI95%	SE CCC	Z	SE Z
0.800620	0.771111	0.826696	0.014154	1.153254	0.039622
Variance Components:					
Subject	Observer	Random Error			
69.1441	0.0995	17.1197			

在調整了年齡及性別兩變數後，兩種血壓測量儀器的一致性程度仍高達 0.8006，而個體隨機效應變異數估計值降低為 69.1441。

利用 R 統計軟體套件 cccrm 中的 ccclonw 函數可以進行長期追蹤資料之重複測量加權一致性相關係數的估計。在此仍以 R 軟體內設的 bfat 資料來進行 ρ_{crm} 的估計，bfat 資料為有 82 位個案(SUBJECT)，每位個案皆被兩種儀器(METHOD)分別進行了 3 次的重複測量(VISITNO)，所測得的反應變數為體脂肪(bf)，其中設定的 3×3 加權矩陣 **D** 的對角線元素為(2,1,1)，同時假設同一位個案的觀測值之間的相關性結構為 compound symmetry，此資料型態如下表：

SUBJECT	VISITNO	bf	METHOD
101	2	21.67627	1
101	3	23.18935	1
101	4	26.31943	1
101	2	17.46278	2
101	3	18.79371	2
101	4	19.91578	2
102	2	21.67627	1
102	3	20.84363	1
102	4	23.88159	1
102	2	15.60228	2
102	3	16.76827	2
102	4	18.25625	2
...

針對第(3)式的線性混合模型，利用 R 軟體的 ccclonw 函數進行長期追蹤資料之重複測量加權一致性相關係數的估計，此函數中需放入的參數如下：

1. dataset：資料名稱
2. ry：反應變數
3. rind：個案身份變數
4. rtime：重複測量時間變數
5. rmet：儀器方法變數
6. vecD：加權矩陣對角線向量
7. covar：個案解釋變數
8. rho：同一位個案觀測值之間的相關性結構；rho=0 代表 compound symmetry，rho=1 代表 autoregressive order one

【cccvc 程式碼】

```
ccclonw(dataset=bfat,ry="bf", rind="SUBJECT",rtime="VISITNO",
rmet="METHOD",vecD=c(2,1,1),rho=0)
```

【結果】

CCC	LL CI95%	UL CI95%	SE CCC	Z	SE Z
0.565884	0.459258	0.656435	0.050317	0.641447	0.074020
Variance Components:					
Subjects	Subjects-Method	Subjects-Time	Method-Time	Error	
8.5920	2.1082	0.9204	17.6777	0.7698	

此資料的分析結果為兩種體脂肪測量儀器的一致性程度為 0.5659，95%的信賴區間為(0.4593, 0.6564)，表示兩種體脂肪測量儀器有中等程度的一致性，其中個體隨機效應變異數估計值為 8.5920，隨機誤差變異數估計值為 0.7698。此外，個案與儀器及個案與測量時間交互作用隨機效應的變異數估計值分別為 2.1082 及 0.9204，儀器與測量時間交互作用效應變異數估計值為 17.6777。

附錄：

在線性混合模式下，利用變異數成份估計一致性相關係數亦可使用 SAS macro *rm_ccc* 來進行估計，而此 SAS macro *rm_ccc* 中需放入的參數如下：

1. DATA：資料名稱
2. Y：反應變數
3. METHOD：儀器方法變數
4. TIME：重複測量時間變數
5. SUBJECT：個案身份變數
6. NL：重複測量的資料型態；NL=0 代表長期追蹤資料(原始設定)，NL=1 代表非長期追蹤資料
7. TYPE：同一位個案觀測值之間的相關性結構；TYPE=0 代表沒有此相關性結構(原始設定)，TYPE=1 代表 autoregressive order one，TYPE=2 代表 compound symmetry
8. DIAG：加權矩陣對角線向量

針對第(1)式的線性混合模型，可利用 SAS macro *rm_ccc* 進行一致性相關係數估計的程式碼如下：

【SAS macro *rm_ccc* 程式碼】

```
%rm_ccc(DATA= bpres, Y=DIA, METHOD=METHOD, TIME= VISITNO,
SUBJECT=ID, NL=1)
```

針對第(3)式的線性混合模型，可利用 SAS macro *rm_ccc* 進行長期追蹤資料之重複測量加權一致性相關係數估計的程式碼如下：

【SAS macro *rm_ccc* 程式碼】

```
%rm_ccc(DATA= bfat, Y=bf, METHOD=METHOD, TIME= VISITNO,
SUBJECT=SUBJECT, NL=0, TYPE=2, DIAG=(2,1,1))
```

參考文獻：

Bartko, J.J., 1966. The intraclass correlation coefficient as a measure of reliability. Psychological Reports 19, 3–11.

- Carrasco, J.L., Jover, L., 2003. Estimating the generalized concordance correlation coefficient through variance components. *Biometrics* 59, 849–858.
- Carrasco, J.L., King, T.S., Chinchilli, V.M., 2009. The concordance coefficient for repeated measures estimated by variance components. *Journal of Biopharmaceutical Statistics* 19, 90–105.
- Carrasco, J.L., Phillips, B.R., Puig-Martinez, J., King, T.S., Chinchilli, V.M., 2013. Estimation of the concordance correlation coefficient for repeated measures using SAS and R. *Computer Methods and Programs in Biomedicine* 109, 293–304.
- Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measure* 20, 37–46.
- Cohen, J., 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70, 213–220.
- King, T.S., Chinchilli, V.M., 2001. A generalized concordance correlation coefficient for continuous and categorical data. *Statistics in Medicine* 20, 2131–2147.
- King, T.S., Chinchilli, V.M., Wang, K.L., Carrasco, J.L., 2007. A class of repeated measures concordance correlation coefficients. *Journal of Biopharmaceutical Statistics* 17, 653–672.
- Lin, L.I., 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45, 255–268.
- Shrout, P.E., Fleiss J.L., 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin* 86, 420–428.