積差相關在體育研究上的應用

王俊明
國立體育學院
體育研究所教授

壹、積差相關的基本概念

一、積差相關的意義

積差相關(product-moment correlation)是指兩個變項之間關係密切的程度。積差相關是由統計學家皮爾遜(K. Pearson)所發展出來的統計方法，積差相關適用於雙變項的資料，而且這兩個變項都是連續變數。由於積差相關是用來了解兩個變項之間的關係，所以是一種簡單的相關，而且其表現的型態是以直線的方式來呈現。

假如是多個變項和一個變項之間的相關，就稱爲多元相關，譬如以急行跳遠來說，一百公尺的跑速、垂直跳的高度、腰力和急行跳遠的距離之間的相關，即是多元相關，主要是用來了解這三個預測變項可以聯合解釋急行跳遠成績總變異量的百分比。假如是數個變項和數個變項之間的相關，就稱爲典型相關。譬如想了解教練領導行爲(訓練與教學的行爲、民主的行爲、專制的行爲、社會支持的行爲、正向回饋的行爲)和團隊適應(人際的適應、情緒的適應、比賽的適應、訓練的適應)之間的相關，就需要用典型相關才能求得。

二、積差相關係數的範圍

積差相關係數的範圍在-1 到+1 之間，其值越大，表示兩個變項之間的關係越密切。積差相關一般可分爲正相關、負相關及零相關。如 $r_{xy} = 0.8$，表示 X 變項和 Y 變項是正相關，而 $r_{xy} = -0.8$，表示 X 變項和 Y 變項是負相關，當 X 變項和 Y 變項是零相關時，其積差相關係數是 0。因此兩個變項是正相關時，表示 X 變項越大，Y 變項也跟著越大。如智力越高的人，其學業成就會越高。當兩個變項是負相關時，X 變項越大，Y 變項反而越小。如考試焦慮越高，考試成績就越低。而零相關時，表示 X 變項和 Y 變項沒有任何關聯。如智力和體重是沒有相關。

三、$r_{xy} = 0$ 的意義

$r_{xy} = 0$ 有兩種意義，一種是 X 變項和 Y 變項是零相關，另一種是 X 變項和 Y 變項是曲線相關。因此，當積差相關係數等於零的時候，在解釋它的意義時要特別小心。假如遽下結論說兩個變項是沒有相關，有時可能會犯下錯誤。$r_{xy} = 0$ 時，頂多只能說它沒有直線相關。

四、$r_{xy} = 0.85$ 與 $r_{xy} = -0.85$ 的異同。

正相關和負相關對於兩個變項之間而言，只是表示其方向的不同而已，其關係密切的程度是一樣的，因此不能說正相關大於負相關。以 $r_{xy} = 0.85$ 與 $r_{xy} = -0.85$ 來說，這兩個相關係數所表示的兩個變項之間的關係是相同的，只是 0.85 是表示 X 變項越大，Y 變項也越大；而-0.85 則表示 X 變項越大，Y 變項反而越小。

五、積差相關係數在統計上的意義

當求得一個相關係數時，這個相關係數是否有統計上的意義，不是由其數字直接來判斷，而是要根據三個指標來決定：自由度、顯著水準的大小、考驗的方向。假設有一個體育老師想了解體能和運動表現之間是否有相關存在，他求得這兩個變項之間的積差相關是 $r_{xy} = 0.54$，而受試者人數是 15，顯著水準定爲 $\alpha = .05$，此項問題的考驗是屬於雙側考驗。由積差相關的統計考驗表查出，其臨界值爲 0.514。因爲所計算的積差相關係數大於查表的臨界值，因此可以體能和運動表現之間有相關關係。假如，所計算的相關值未能大於臨界值，這個相關就不具有統計上的意義。

六、積差相關係數的性質

積差相關係數本質上並不屬於比率變數，譬如 $r_{xy} = 0.80$ 不是 $r_{xy} = 0.40$ 的兩倍；此外，積差相關係數也不屬於等距變數，譬如 $r_{xy} = 0.70$ 到 $r_{xy} = 0.80$ 的距離等於 $r_{xy} = 0.30$ 到 $r_{xy} = 0.40$ 的距離。由上述可知，積差相關係數應不屬於連續變數。基本上，積差相關係數應該比較偏向次序變數。

七、積差相關係數的解釋

積差相關只是表示兩個變項之間的關係，因此在解釋時只能就兩者之間的相關關係來解釋，不能將其解釋爲因果關係。譬如假如得知少年犯在攻擊行爲量表的得分和其家庭氣氛量表的得分有相關關係，不能說少年犯的攻擊行爲是由家庭不良氣氛所造成。我們只能說少年犯的攻擊行爲可能和家庭氣氛有關係。事實上，少年犯攻擊行爲的形成也許有可能是來自於同儕的影響，也可能是看了電視上暴力的節目而學得的。

若要了解兩個變項之間的因果關係，只有透過實驗法，操弄自變項，控制無關變項，然後觀察依變項，才能了解自變項和依變項之間的因果關係。相關法所得的研究結果，只能了解兩個變項之間的相關關係而已。

此外，當一個積差相關係數在統計上達顯著水準時，是否在實用上必有其意義？此點是有待探討的。譬如某一位教師想了解考試運氣與考試成績之間是否有相關關係，受試者爲 102 人，所得積差相關係數爲 0.2。假設此研究的顯著水準爲 $\alpha = .05$。經查表得知積差相關的臨界值爲 0.195。由於所計算的值大於查表的臨界值，因此可以說所獲得的積差相關係數在統計上是有意義的。假如由此結果就遽下結論說考試運氣果然會影響考試的成績，這種結論事實上是不可靠的。因爲將此積差相關係數加以平方得到 $r^2=0.04$，這個值就是所謂的決定係數(coefficient of determination)，也就是說考試運氣能解釋考試成績的總變異量只有 4%。由此得知，考試運氣對考試成績的影響只是微乎其微。可能考試前一晚的睡眠，或是當天的情緒對考試的影響都遠大對考試運氣。因此，任何積差相關若是達顯著水準，不能隨意的就做下結論，還要進一步看其是否有實用上的意義。有時積差相關係數能達顯著水準，只是因爲受試者人數多，造成其臨異值變小，因此所得的積差相關係數很容易達顯著水準，這是研究者們在做研究上應該注意的。

八、積差相關係數與團體變異的關係

小學生的智力和學業成績之間的相關會大於大學生智力和學業成績之間的相關，爲何會有此種現象？主要原因是小學生的異質性較高，而大學生因爲經過篩選的關係，其同質性較高。

以下就用一個例子來說明團體變異與積差相關的情形：

表一　國中生智力與學業表現的積差相關資料

學生	智力		學業表現		Z_xZ_y
	X	Z_x	Y	Z_y	
A	112	0.7792	76	-0.1302	-0.1015
B	120	1.2375	82	0.8462	1.0472
C	116	1.0084	86	1.4972	1.5098
D	114	0.8938	74	-0.4557	-0.4073
E	116	1.0084	88	1.8226	1.8379
F	86	-0.7104	74	-0.4557	0.3237
G	84	-0.8250	72	-0.7811	0.6444
H	78	-1.1688	72	-0.7811	0.9129
I	80	-1.0542	76	-0.1302	0.1373
J	78	-1.1688	68	-1.4321	1.6738

以上十名學生智力和學業表現的積差相關係數是 0.7578，但假如以智力和學業表現較佳的前五名學生(視為高分組)，求其智力和學業表現的積差相關，則降為 0.5306，後五名學生(視為低分組)的積差相關亦降為 0.4082。當十名學生一起求積差相關時，由於其間異質性較多，因此智力和學業表現的積差相關會較高。而當高分組和低分組個別求智力和學業表現的積差相關時，因同質性高，而使得其積差相關係數大為降低。其積差相關係數的降低並不是因為人數的減少所造成，現可將十名學生隨機選取五名求其係數，假設現隨機抽取 A、B、C、I、J 五名學生，可求得其智力與學業表現的積差相關係數為 0.7968。由此可見，積差相關係數的高低並不會受人數的多寡而影響，而是因其群體內的變異而產生變化。

由以上的說明可知，一個研究者在進行相關的研究時，在解釋其積差相關係數的高低時，應該進一步考慮受試者的特

質，而不是僅就其係數的高低做解釋。譬如以一般大學生三分鐘登階的成績和其 1600 公尺跑走的成績求積差相關，另以大學田徑代表隊同樣的兩項成績求積差相關，前者所得的積差相關係數應該比後者的積差相關係數要高。因為，前者的異質性高。

貳、積差相關係數的計算

一、以 z 分數來計算積差相關係數

積差相關的定義公式就是將兩個變項的原始分數化為 z 分數，然後將兩兩相對的 z 分數相乘，如表一的例子所示。積差相關係數的定義公式為：$r_{xy} = \frac{\sum ZxZy}{N}$

由表一的例子得知 $\sum Z_x Z_y$ 為 7.5782，由於人數共有 10 人，因此由定義公式計算得到 $r_{xy} = 0.7578$。在實際計算積差相關係數時，我們並不會用 z 分數的方法來做，因為這種方法不但繁瑣，而且牽涉到太多的小數點，在計算時較會有誤差產生。

不過此定義公式可以幫我們了解為何團體的變異變大時，其積差相關係數也會跟著變大。主要是團體的變異變大時，每個人的 z 分數的差異也會跟著變大，這將使變異較大的團體其 $Z_x Z_y$ 加大，因此團體變異較大的 $\sum Z_x Z_y$ 就會比團體變異較小的 $\sum Z_x Z_y$ 為大。由此得知，當團體變異變大時，其積差相關係數也會變大。

二、以共變數來計算積差相關係數

其次，積差相關係數也可用共變數來求得，其計算公式為：

$r_{xy}=\frac{C_{xy}}{S_xS_y}$，其中分子的 C_{xy} 就是共變數。所謂共變數就是兩個變項共同的變異量。由此公式可知，兩個變項共同的變異量除以兩個變項個別的變異量所得的值就是積差相關係數。共變數的公式可進一步改爲計算式：$C_{xy}=\frac{\sum(X-\overline{X})(Y-\overline{Y})}{N}=\frac{\sum XY-\frac{\sum X\sum Y}{N}}{N}$。

共變數的公式非常類似變異數的公式：$S_x^2=\frac{\sum X^2-\frac{(\sum X)^2}{N}}{N}$，假如將變異數的公式改爲：$S_x^2=\frac{\sum XX-\frac{\sum X\sum X}{N}}{N}$，將此公式與共變數的公式比較，可發現兩者的差異只在於共變數有兩個變項，而變異數只有一個變項。事實上，我們也不會用共變數的方法來求得積差相關係數。共變數的公式能讓我們了解當兩個變項沒有共同的變異量時，就沒有積差相關可言。而且我們可進一步將共變數的公式改爲：$C_{xy}=r_{xy}S_xS_y$。因爲 S_x 和 S_y 不會小於零，所以 S_xS_y 和 r_{xy} 的正負號會是一致的。亦即當共變數爲負值時，積差相關係數也是負的；而共變數爲正值時，積差相關係數則爲正值；當共變數爲零時，積差相關係數當然也是零。

三、以運算公式來計算積差相關係數

積差相關的運算公式是由共變數的公式所轉換而來的，其差別只是運算公式都是以原始的得分來運算，所得的值比較精確。目前一些套裝軟體有關積差相關係數的計算也都是採用下列的運算公式來求得。

$$r_{xy}=\frac{\sum XY-\frac{\sum X\sum Y}{N}}{\sqrt{\sum X^2-\frac{(\sum X)^2}{N}}\sqrt{\sum Y^2-\frac{(\sum Y)^2}{N}}}$$

參、積差相關在統計分析上的應用

一、一個積差相關係數的假設考驗

一個積差相關係數的假設考驗只是要了解所求得的積差相關係數是否達顯著水準，這種考驗必須知道自由度、顯著水準的大小、考驗的方向。譬如有某位田徑教練想了解跳遠選手的運動成就動機是否和跳遠成績有相關關係。假設該教練找到 12 位跳遠選手給予實施運動成就動機量表，並取得他們的跳遠成績，最後計算兩項成績的積差相關，得 $r_{xy}=0.61$。由於此教練只想了解兩項成績是否有相關關係，因此這個考驗是屬於雙側考驗。此項問題的自由度是人數減 2，得到 10。一般社會科學的研究，其顯著水準大多是將 α 定爲.05。有了這三個考驗的指標，經查積差相關的臨界值，得知是.576，因爲所計算的值大於查表的臨界值。因此，該名教練可說運動成就動機和跳遠的成績有相關關係。

假如有某位教練想要了解情境焦慮對運動表現是否有不好的影響，亦即情境焦慮越大，運動的成績會越差。同樣的，他選了 12 名跳遠選手，並求得其情境焦慮的得分和跳遠的成績，同時算出這兩項成績的積差相關係數爲 0.52。要考驗此教練的想法是否獲得支持，就要進行一個積差相關係數的考驗。此項研究的自由度是 10，考驗的方向性應屬於單側考驗(亦即考驗情境焦慮和跳遠成績是否有負相關)。假設此題的顯著水準定爲

α =.05。經查積差相關的考驗表之後，得知其臨界值爲.497。由於計算的積差相關係數大於查表的臨界值，因此，該名教練的想法可以獲得支持。

二、兩個積差相關係數的差異比較

假如要比較兩個積差相關係數的差異，就會牽涉到獨立樣本和相依樣本的問題。以下就分別說明這兩種的差異比較：

(一)獨立樣本

某運動生理學家認爲一般大學生的體能和運動表現的相關應大於運動選手的體能和運動表現的相關。現在他以 103 名大學生求得體能和運動表現的相關爲 0.64，以 53 名運動選手求得體能和運動表現的相關爲 0.42，問是否可以支持該運動生理學家的說法？此項研究的公式爲：$Z = \dfrac{Z_{r1} - Z_{r2}}{\sqrt{\dfrac{1}{N_1 - 3} + \dfrac{1}{N_2 - 3}}}$

由於所計算得的兩個積差相關係數分屬不同的群體，必須轉換爲 z 分數才能互相比較，因此將 0.64 轉換爲 0.758，而將 0.42 轉換爲 0.448，然後代入公式計算其 z 值。經計算後，所得 z 值爲 1.79。假設此項研究的顯著水準爲 α =.05，因此研究爲單側考驗，查常態分配表後，得知其臨界值爲 $z_{1-.05}$ =1.65。因爲計算值大於查表值,所以該運動生理學家的說法獲得支持。

(二)相依樣本

某教練想根據選手一百公尺的成績(X_2)和垂直跳的成績(X_3)和跳遠成績(X_1)求積差相關，然後比較一百公尺的成績和跳遠成績的積差相關係數($r_{x_1x_2}$)與垂直跳的成績和跳遠成績的

積差相關係數($r_{x_1x_3}$)兩者是否有差異(α =.05)。該教練共以 123 名選手來進行此項研究，分別求得各組的積差相關係數爲 $r_{x_1x_2}$=0.60；$r_{x_1x_3}$=0.45；$r_{x_2x_3}$=0.50。相依樣本的積差相關係數比較公式如下：$t=\frac{(r_{x1x2}-r_{x1x3})\sqrt{(N-3)(1+r_{x2x3})}}{\sqrt{2(1-r_{x1x2}^2-r_{x1x3}^2-r_{x2x3}^2+2r_{x1x2}r_{x1x3}r_{x2x3})}}$

由此公式計算的 t 值爲 2.104，因爲本研究的顯著水準爲.05，而且是雙側考驗，故其臨界值爲 t=1.980。由於計算值大於查表的臨界值，因此可以說一百公尺的成績和跳遠成績的積差相關係數高於垂直跳的成績和跳遠成績的積差相關係數。

三、積差相關係數在迴歸分析的應用

積差相關在迴歸分析的應用可分爲簡單直線迴歸，和多元迴歸分析。前者是一對一的關係，後者是多對一的關係。在簡單直線迴歸分析中，有一個預測變項(predictable variable)和一個效標變項(criterion variable)。迴歸分析主要是了解預測變項對效標變項的總變異量可以解釋多少的百分比。以表一的例子而言，以智力來預測學業表現，因兩者的積差相關係數是 0.7578，將積差相關係數平方後得到 r^2=0.5743。r^2 稱爲決定係數，它的意義是智力可以解釋學業表現總變異量的 57.43％。由此例子我們進一步可以得到一個簡單迴歸公式 $\hat{Y}$=50.5462 + 0.2668X。由此項迴歸公式可以從智力來預測學生的學業表現，譬如某個學生的智力是 100，代入此公式後，可得其學業表現爲 77.23。

一般而言，簡單直線迴歸在體育統計上的應用較少，而以多元迴歸分析爲多。在多元迴歸分析方面，通常可分爲多元逐步迴歸分析、多元同時迴歸分析、多元階層迴歸分析等三種。

不管迴歸分析的類型如何，大多是以積差相關爲基礎。因此若要了解高深的迴歸分析，一定要了解積差相關的理論。

肆、積差相關在體育測驗上的應用

一、積差相關在重測信度上的應用

所謂重測信度是指同一個測驗在不同的時間，重複測量相同的一群受試者兩次，根據兩次分數求得的相關，稱爲重測信度係數(郭生玉，民 77)。因爲重測信度是經過一段時間所得，此係數可以表示其穩定性，因此又稱爲穩定係數。由於同一份測驗重複測兩次，當然免不了有記憶或是練習的影響。因此，兩次測驗的時間必須有一段間隔。間隔的時間越短，重測信度就越高。反之。則會變低。至於應該間隔多長的時間，則需視測驗的種類、目的及受試者而調整。一般而言，情意測驗的重測信度至少要間隔兩週，認知測驗需要間隔長一些時間(譬如一個月)，而技能測驗則以不受疲勞影響即可。兒童在做認知測驗的重測信度時，其間隔時間不宜太長，因爲兒童的發展甚快，如果間隔半年以上，所獲得的重測信度就不太可靠。

重測信度雖然是以積差相關來計算，但其範圍卻不像積差相關係數從-1 到+1，信度的範圍是從 0 到+1。重測信度究竟要多高才能被接受？一般而言，重測信度大多可以達 0.7 以上，少數在 0.6 至 0.7 之間。若是在 0.6 以下的，其信度就算非常不理想。重測信度也會受題數的多寡影響，題數越多，信度就會越低。

二、積差相關在折半信度上的應用

折半信度不像重測信度需要做兩次，它只要做一次即可。一般是將一份測驗分成兩半(通常是分成奇數題和偶數題)，然後將兩半各自算其分數求積差相關。由於一份測驗被分成兩半，題數減少，信度就會被低估。此時可用斯布公式(Spearman-Brown formula)來調整。其公式如後：$r_{xx} = \frac{2r_{hh}}{1+r_{hh}}$。

折半信度屬於內部一致性係數的一種，它的範圍也是從 0 到+1。在計算折半信度時，假如兩半測驗的變異數不相等，會稍微高估測驗的信度。

三、積差相關在效標關聯效度上的應用

所謂效度關聯效度是以所得的測驗分數和外在效標的分數求積差相關，其相關係數越高，代表其效度越好。譬如某一研究者編製一份運動成就動機量表，此量表可由選手以自評的方式求其得分，其次則由教練以同一份量表對每一個選手給予評分，然後將兩個量表求積差相關，即可獲得效標關聯效度。

一般而言，效標關聯效度常會遭遇效標不易獲得的問題。假如所得效標不切實際的話，其效度通常會有偏低的現象。如前述以教練的評分做為效標，假如是資深的教練，應該是良好的效標。但若是才剛接任的教練，其評分就不可靠了。

伍、結　語

積差相關的應用不只是用來了解兩個變項之間關係密切的程度，而且可用來做為某個變項預測另一個變項之用，其用途甚大，實為進行科學研究的人所不得不懂的一門學問。在使用

積差相關時，必須注意幾點的是：一、積差相關只用來了解兩個變項間的直線關係，對於曲線的關係就要用趨向分析(trend analysis)才能解決。二、當受試者的人數太少時，應該要注意是否有 outlier 存在。當有獨特的受試者時，即使只有一個，也會嚴重影響其積差相關係數的大小。三、積差相關係數是否有統計上的意義，常會受樣本數的大小所影響。假如人數過少，即使有高的相關(如 0.8)可能也達不到顯著水準。反之，若有數百個受試者，即使是低的相關(如 0.2)，也會很容易達顯著水準。因此，在解釋積差相關的意義時要特別考慮受試者人數的多寡。

參考文獻

林清山(民 81)：心理與教育統計學。台北：東華書局。

黃瓊蓉譯(民 90)：心理與教育統計學。台北：學富文化事業限公司。

Darlington, R.B., & Carlson, P.M.(1987). Behavioral statistics: Logic and methods. New York: The Free Press.

附件一

積差相關係數臨界值表

附 件 二

積 差 相 關 係 數 的 z 分 數 對 照 表