

The Reversal Paradox 統計悖論

林學銘、白其卉

前言

統計學是觀察並收集、整理及分析資料的學問，然而，在推導結果的過程會發生與直觀相違背的現象，稱為統計悖論(paradox)。統計悖論不僅具有學術上的價值，對於研究結果的解釋與推論亦息息相關。本文簡介常見的三種統計悖論，以流行病學資料為例，比較連續型、類別型資料或者二者兼具的資料型態下，可能發生的悖論現象與成因。

Simpson's Paradox

辛普森悖論或稱作 Yule–Simpson's Paradox，由英國統計學家(Edward H. Simpson)於 1951 年提出，是一種常見的統計悖論，指探討二組類別型變項是否具有相關性時，可能因為加入第三個類別型變項而得到完全相反的結果。舉例來說，某研究者想知道出生體重(birth weight, BW)是否與日後得到高血壓(high blood pressure, BP)有關。由下表可觀察到：

表 1. 高血壓者(>135 mmHg)以出生體重及成年體重分組之數目與百分比

	Normal BP	High BP	Total	Percentage of subjects with high BP
Overall:				
Low birth weight	354	132	486	27.2%
High birth weight	328	186	514	36.2%
Total	682	318	1,000	31.8%
Current weight ≤ 90 Kg				
Low birth weight	329	99	428	23.1%
High birth weight	221	55	276	19.9%
Total	550	154	704	21.9%
Current weight > 90 Kg				
Low birth weight	25	33	58	56.9%
High birth weight	107	131	238	55.0%
Total	132	164	296	55.4%

在整體評估時，出生體重輕者比出生體重大者有較低的高血壓發生百分比，亦即低的出生體重具有保護性；但若考慮成年體重(current weight, CW)亦會影響高血壓的發生，以現在體重為作分組(大於 90 kg versus 小或等於 90 kg)，此時不論是高或低的成年體重，出生體重輕者在發生高血壓的風險皆較高，低的出生體重反而具危險性。

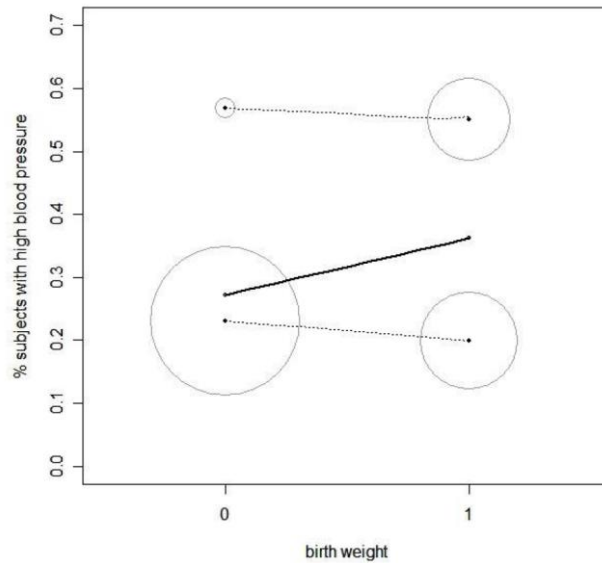


圖 1. 圖示說明 Simpson's Paradox。圖上半部二個圓表示較低的成年體重(小或等於 90 kg)，下半部二個圓為較重成年體重(大於 90 Kg)，圓面積大小則為樣本數的比例。虛線表示在成年體重分組的情況下，出生體重較輕者(birth weight: 0)比出生體重較重者(birth weight: 1)成年後發生高血壓的風險皆較高。整體來看時，出生體重較輕者有較低風險發展成高血壓(粗實線)。

悖論以純數學角度來看並無特殊之處，但在統計學解釋上卻具有重要意義。上述例子說明當考慮另一個類別變項的影響時，變項之間的相關性可能完全被扭轉(reversal)。此一相反結論的現象來自於資料本質上的差異，由於出生體重較輕者在成年後被分組的樣本數差距過大(428 versus 58)，使得在整體觀察上成年體重輕的人有較大的比重，進而影響(降低)了整體低的出生體重者的高血壓發生百分比(如圖 1 所示)。

Lord's Paradox

Lord's Paradox 是由 Frederic M. Lord 利用共變異數分析(ANCOVA)討論一非實驗性研究設計(non-experimental design)的文章後所命名。當研究者欲探究二因子的關係時，其 Y 是連續型資料而暴露 X_1 是類別型資料，此時需校正另一個連續型的變項 X_2 而發生相反結論的情形即為 Lord's Paradox。此現象常見於長期追縱的世代研究分析時，校正基線的共變量後對於結果的影響。

同樣以上述出生體重對於日後成人發生高血壓為例，成年體重與血壓為連續變項，出生體重則為二項式變項(BWb)。Two-sample t-test 顯示出生體重較重者相對於出生體重較輕者具有較高的平均血壓，然而成年體重做為共變項以 ANCOVA 進行校正後，結果呈現出生體重較重者反而較出生體重者較輕者有相對較低的平均血壓。

此悖論來自當現在體重與血壓及出生體重都具有正相關時，血壓與二組出生體重之間的關係會發生改變；當出生體重不為隨機分組，組別間會存在系統性的差異，亦即出生體重與成年體重兩者不獨立。圖 2 說明此一現象：只考慮血壓與類別型變項出生體重時，出生體重較重組有較高的血壓平均值(粗黑色

線)。若校正了現在體重的影響，則對於每個成年體重來說，平均血壓在出生體重較輕者都高於出生體重較重者(紅色水平線)。

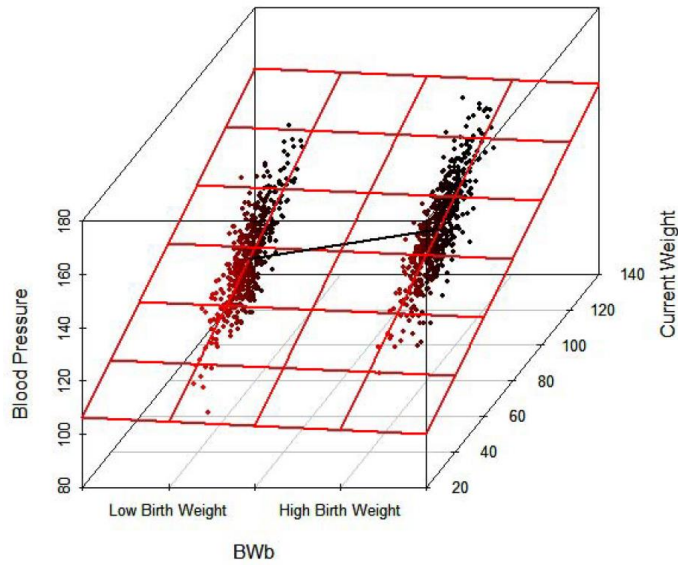
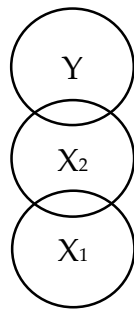


圖 2. 3D 散佈圖說明 Lord's Paradox。粗黑色線表示高的出生體重者平均血壓高於低的出生體重者。然而，水平紅色線表示成年體重的影響，平均血壓由低的出生體重傾斜向高的出生體重。

Suppression

相對於 Simpson's Paradox 與 Lord's Paradox，Suppression 是線性迴歸分析中較不常見卻重要的現象。Suppression 的定義為：當一個變項 X_1 與 Y 變項沒有相關性時，但 X_1 變項加入迴歸模型時，卻增加該模型($Y=X_2$)的解釋力(R^2 增加)，此一違反直覺的現象即為 Suppression。



考慮一迴歸方程式為：

$$Y = b_0 + b_1X_1 + b_2X_2$$

r_{y1} : X_1 與 Y 的相關係數 ($r_{y1}=0$)

r_{y2} : X_2 與 Y 的相關係數

r_{12} : X_1 與 X_2 的相關係數

由方程式可得迴歸係數 b_1 與 b_2 分別為：

$$b_1 = \frac{r_{y1} - r_{12}r_{y2}}{1 - r_{12}^2} \quad \& \quad b_2 = \frac{r_{y2} - r_{12}r_{y1}}{1 - r_{12}^2}$$

決定係數 R^2 則表示為：

$$R^2 = r_{y1} * \beta_1 + r_{y2} * \beta_2。$$

在 r_{y1} 為 0 的情況下，

$$R^2 = r_{y2} * b_2 = \frac{r_{y2}^2}{1 - r_{12}^2}，$$

由於 $1 - r_{12}^2$ 永遠小於 1，故可知 $R^2 = \frac{r_{y2}^2}{1 - r_{12}^2}$ 永遠大於 r_{y2}^2 。

雖然 X_1 與 Y 的相關性不存在($r_{y1}=0$)，但 X_2 與 Y 之間有相關，此時方程式加入 X_1 而壓抑或過濾掉(suppressed) X_2 與 Y 之間不相關的部分，因此增加了 X_2 對模型的解釋力。

同樣以上述體重與血壓研究為例，此時出生體重、血壓及成年體重皆為連續型資料， Y 為血壓， X_1 為成年體重， X_2 為出生體重。由表下可知，出生體重與成年體重與血壓都有正相關，這表示當出生體重或成年體重增加時，血壓會隨之上升。然而，同時考慮二者與血壓的關係時，由表 2 複迴歸(model 3)方程式可知：

$$Y = -3.708 X_1 + 0.465 X_2 + 104.330，$$

此方程式表示血壓會隨著出生體重增加而下降，與簡單迴歸(mode 1 & model 2)呈現相反結論，此即所謂 Suppression 現象。

表 2. 簡單與複迴歸表，說明出生體重、成年體重與血壓之間的關係。三個模型 Y 皆為血壓。

Model	Regression Coefficients (Standard Errors)	Standardised Coefficients	P-values	R ²
1	Intercept	123.258 (1.981)	(< 0.001) †	0.011
	Birth weight	1.861 (0.556)	0.001	
2	Intercept	98.173 (1.755)	(< 0.001) †	0.251
	Current weight	0.382 (0.021)	< 0.001	
3	Intercept	104.330 (1.948)	(< 0.001) †	0.283
	Birth weight	-3.708 (0.553)	< 0.001	
	Current Weight	0.465 (0.024)	0.610	

圖 3 亦顯示，雖然出生體重與血壓呈現明顯的正相關(粗黑色線)，然而同時考慮成年體重的效應下，出生體重愈高則血壓愈低(水平紅色線)，與原本結論相左。

此簡單迴歸與複迴歸結果相反的情形，何者能真實的反映出真正的因果關係？端看成年體重是否需被校正以及現在體重是否為出生體重與血壓之間的干

擾因子，此關係也有賴生物學與臨床上的知識來判斷。

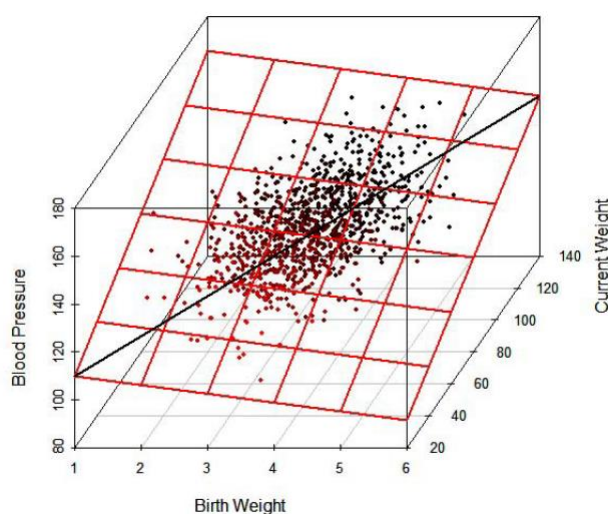


圖 3. 3D 散布圖說明 Suppression。粗黑色線表示出生體重與血壓具正相關。然而，水平紅色線則顯示考量成年體重，出生體重與血壓的關係呈相反結果。

結論

本文列舉的三個統計悖論雖然出現於不同資料類型的變項(如下表 3)，其相同之處為在於，非隨機分派的研究中，加入第三個因子做為校正時，X 變項與 Y 變項的關係可以是被減弱、增強甚至是反轉(reversal)的。了解統計悖論的概念與成因有助於洞察許多研究的爭議與矛盾之處，不正確的使用統計模型也可能會導致錯誤的結論。是故，對於非隨機分派的資料的研究，在建立統計模型時，不只需考量統計的因果推論準則，研究前的先驅知識與理論建立也扮演致關重要的角色。

表 3. Simpson's paradox、Lord's paradox 與 Suppression 比較表

Type of Reversal Paradox	Outcome (illustrated example)	Exposure (illustrated example)	Covariate/'Confounder' (illustrated example)
Simpson's Paradox	Categorical (hypertension)	Categorical (birth weight: high vs. low)	Categorical (current weight: high vs. low)
Lord's paradox	Continuous (blood pressure)	Categorical (birth weight: high vs. low)	Continuous (current weight)
Suppression	Continuous (blood pressure)	Continuous (birth weight)	Continuous (current weight)

Reference

- Tu YK, Gunnell D, Gilthorpe MS. Simpson's Paradox, Lord's Paradox, and Suppression Effects are the same phenomenon--the reversal paradox. *Emerg Themes Epidemiol* 2008, 22;5:2. doi: 10.1186/1742-7622-5-2.
- Simpson EH. The interpretation of interaction in contingency tables. *J R Stat Soc Ser B* 1951, 13:238-41.
- Yule GU. Notes on the theory of association of attributes in statistics.

- Biometrika 1903, 2:121-34.
- Hennessy E, Alberman E. Intergenerational influences affecting birth outcome. II. Preterm delivery and gestational age in the children of the 1958 British birth cohort. *Paediatr Perinat Epidemiol* 1998, 12(1):61-75.
- Website of the Department of Health, United Kingdom
[<http://www.doh.gov.uk>]
- Paik M. A graphical representation of a three-way contingency table: Simpson's paradox and correlation. *Am Stat* 1985, 39:53-54.
- Lord FM. A paradox in the interpretation of group comparisons. *Psychol Bull* 1967, 68:304-5.
- Lord FM. Statistical adjustments when comparing preexisting groups. *Psychol Bull* 1969, 72:337-8.
- Pedhazur EJ. *Multiple regression in behavioral research: Explanation and prediction*. Fort Worth: Harcourt; 1997.
- Arah OA. The role of causal reasoning in understanding Simpson's paradox, Lord's paradox, and the suppression effect: covariate selection in the analysis of observational studies. *Emerg Themes Epidemiol*. 2008 Feb 26;5:5. doi: 10.1186/1742-7622-5-5.
- Tu YK, West R, Ellison GT, Gilthorpe MS. Why evidence for the fetal origins of adult disease might be a statistical artifact: the "reversal paradox" for the relation between birth weight and blood pressure in later life. *Am J Epidemiol*. 2005 Jan 1;161(1):27-32.