

統計，讓數字說話！



Statistics

concepts and controversies

Chapter 2

實驗



實驗可對因果關係**提供好的證據**。

為什麼要做實驗呢？

這樣我們才可以確實針對我們**感興趣的處理**，
研究**效果如何**。



觀測研究與實驗

◆ 觀測研究（observational study）：

觀察一些個體，並度量我們感興趣的變數，但並不試圖影響回應。抽樣調查是一種重要的觀測研究。觀測研究之目的是描述一個團體或一種狀況。

◆ 實驗：

刻意地將一些個體做某種「處理」（treatment），以觀察他們的反應。實驗之目的，是要研究是否該處理會使回應有改變。



為什麼要實驗？

- ◆ 實驗所提供的訊息比抽樣調查多。
- ◆ 實驗可將個體做某種「處理」，以觀察他們的反應。
- ◆ 實驗是要研究是否該「處理」會使回應有改變。
- ◆ 實驗優於觀測研究，可對因果關係提供好的證據。



Ex. 阿斯匹靈和心臟病發作的關係

按時服用阿斯匹靈是不是能防止心臟病發作？

- ◆ 實驗：醫師健康研究。
- ◆ 實驗個體：共有兩萬兩千位男性醫師。
- ◆ 實驗方法：隨機抽取一半人數，每隔一天服用一顆阿斯匹靈；另一半醫師服用安慰劑（不含有效成分）。
- ◆ 實驗結果：安慰劑組中有239人心臟病發作；而阿斯匹靈組只有139人。
- ◆ 實驗差異：實驗結果差距已證明，阿斯匹靈可以減少心臟病發作。



Ex.修習外國語與增進英文能力關係

在美國，於中學修習外國語是否會增加學生的英文能力？

- ◆ 實驗個體：高三學生曾經修習外國語 vs. 不曾經修習外國語者
- ◆ 實驗方法：比較兩組學生的英文閱讀及文法的標準測驗。
- ◆ 實驗結果：修過外國語的學生之平均成績大於未修過的學生。
- ◆ 實驗差異：觀測究並沒有提供證據，證明修習另一種語言會增進英文能力。



實驗用語言

◆ 個體：

實驗加諸於其上的物體。當個體為人的時候，我們稱其為**受試對象**。

◆ 變數：

個體某一**被量的特質**。

◆ 反應變數（response variable）：

我們想研究其變化的變數，反應變數是一個「**結果**」。



實驗用語言 (Cont.)

- ◆ 解釋變數 (explanatory variable) :

可以解釋或造成反應變數變化的變數。

- ◆ 處理 :

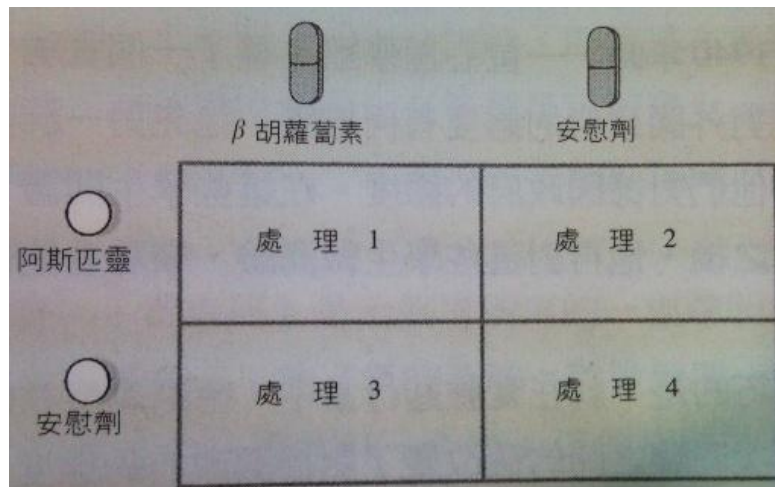
任何加諸於個體的特定實驗條件。若一個實驗當中有數個解釋變數，則處理是指每個變數都設定一個特定值後的組合。



Ex. 醫師健康研究

醫師健康研究中考慮了兩種藥的效果：阿斯匹靈及 β 胡蘿蔔素，欲研究其效應。

- ◆ 實驗個體：受試者分別進行四種不同的處理。
- ◆ 實驗方法：觀察心臟病發作、數種癌症及一些其他相關的醫學結果。
- ◆ 實驗目的：研究服用阿斯匹靈的效果及服用 β 胡蘿蔔素的效果外，及研究交互作用。





交絡

交絡 (confounding)

- ◆ 潛在變數 (**lurking variable**)：是對研究中其他變數間的關係有**重要影響**、**卻沒有被列入**研究範圍的變數。
- ◆ 當兩個變數對反應變數的影響混在一起、無法區分時，我們稱這兩個變數是**交絡**。交絡的變釋可以是**解釋變數**，**也可為潛在變數**。



Ex. 胃冷凍

- ◆ **Ex. 胃冷凍**是治療胃潰瘍的一種療法。病人先吞下一個連接著管子、放了氣的氣球；然後把一種經過冷凍的溶液打入汽球中。這種療法的想法是使胃涼下來可減少胃酸的分泌，因此減輕潰瘍症狀。
- ◆ **實驗目的：**胃冷凍→減輕痛苦？
- ◆ **實驗方法：**一組接受胃冷凍治療；另一組接受安慰劑療法。
- ◆ **實驗結果：**處理組82病人有34%改善，安慰組78病人有38改善。
- ◆ **實驗證明：**胃冷凍根本不比安慰劑療法。



隨機化比較實驗

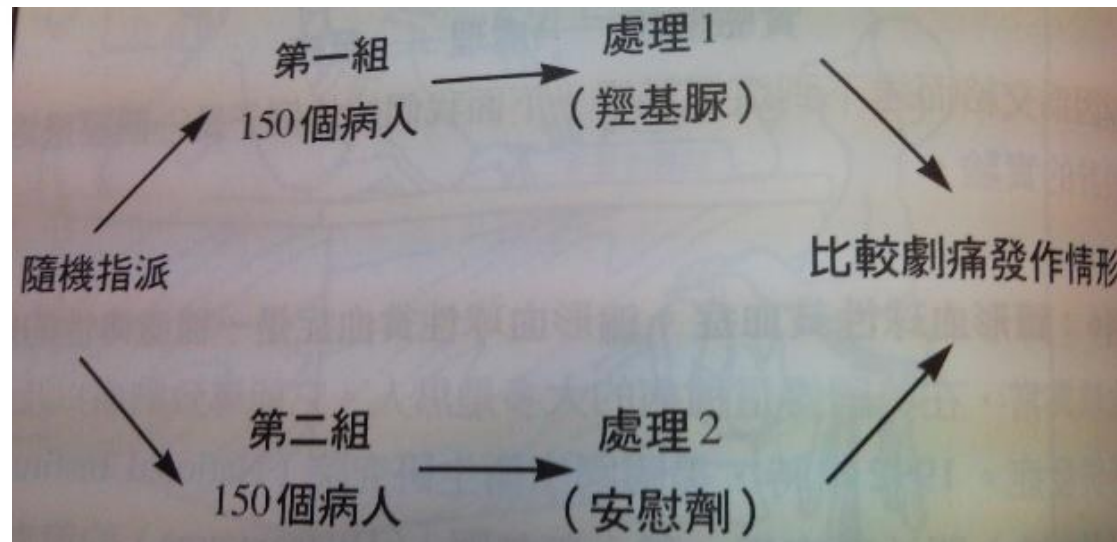
- ◆ 設計實驗的第一個目標，是要確定實驗可以顯示解釋變數對於反應變數的影響。
- ◆ 實驗個體 → 處理 → 反應
- ◆ 常因交絡而達不到這個目標。



Ex. 鎌形血球形貧血症

- ◆ Ex. 美國衛生研究探討「羥基脲」對鎌形血球形貧血症的治療效應。

- ◆ 實驗方法：



- ◆ 實驗結果：羥基脲組的劇痛發作次數只有控制組的一半。

證實羥基脲是鎌形血球性貧血症的有效療法



實驗設計的邏輯

- ◆ 隨機化比較實驗的設計是要讓我們得到「**原因和結果**」。
- ◆ 用**隨機化**的方法將實驗個體分組。
- ◆ 用**比較的設計確保**，除了正在實驗的處理外，其他所有因素都會同樣作用在所有的組上。
- ◆ 反應變數的差異必是處理的效用所致。



實驗設計的原則

- ◆ **要控制**潛在變數對反應的影響，最簡單的方法是同時比較**好幾個處理**。
- ◆ **隨機化**：用非人為的隨機方法來指派受試對象到不同的處理組。
- ◆ **重複**：對許多受試對象重複實驗，以減低結果中的機遇變異。



統計顯著性

統計顯著性（statistical significance）：

- ◆ 如觀察到的效果如果大到某種程度，光靠機遇產生這種結果的機率很小時，即稱此結果有統計顯著性。



隱性偏差

隱性偏差：

- ◆ 隨機化比較實驗的邏輯假設：除了實驗要比較的處理之外，所有受試對象受到的待遇是一樣的。如果有任何其他不相等的處理，就可能造成偏差。



雙盲實驗

雙盲（double-blind）實驗：

- ◆ 在雙盲實驗中，對於每個受試對象得到那種處理，受試對象和所有工作人員都完全矇在鼓裡。在研究結束、結果出來之前，只有該研究的統計學家確知此事。



拒絕參加、不合作者及退出者

- ◆ 人們不願當「實驗品」，而拒絕參加。
- ◆ 參加實驗卻不遵循實驗處裡的受試對象，叫不合作者。
- ◆ 實驗時間較長而中途退出者。



處理夠不夠實際？

- ◆ 避免出現太人工的實驗情境。
- ◆ 避免讓人們知道自己正被研究而改變行為事實出現「霍桑效應」。



結論能推廣到什麼程度？

- ◆ 實驗設計的邏輯非常有說服力，但必須經過不同背景多次實驗探討才能適用。
- ◆ 好的實驗必須結合統計原則及對研究專業領域的了解。



實驗設計

- ◆ 完全隨機化設計 (completely randomized design)
- ◆ 區集設計 (block design)
- ◆ 配對設計 (matched pairs design)

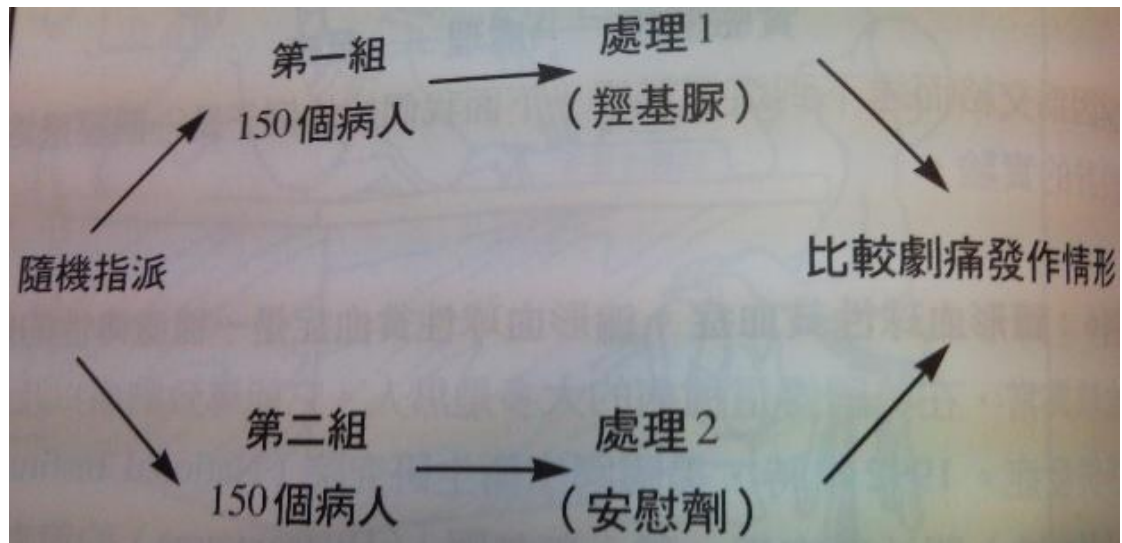


完全隨機化設計

- ◆ 在完全隨機化的實驗設計中，所有的實驗個體都是**隨機配置**給所有的處理。

Ex.

- ◆ 隨機指派
- ◆ 一個處理分配一組人
- ◆ 每組人數相近
- ◆ 每一組隨機指派處理





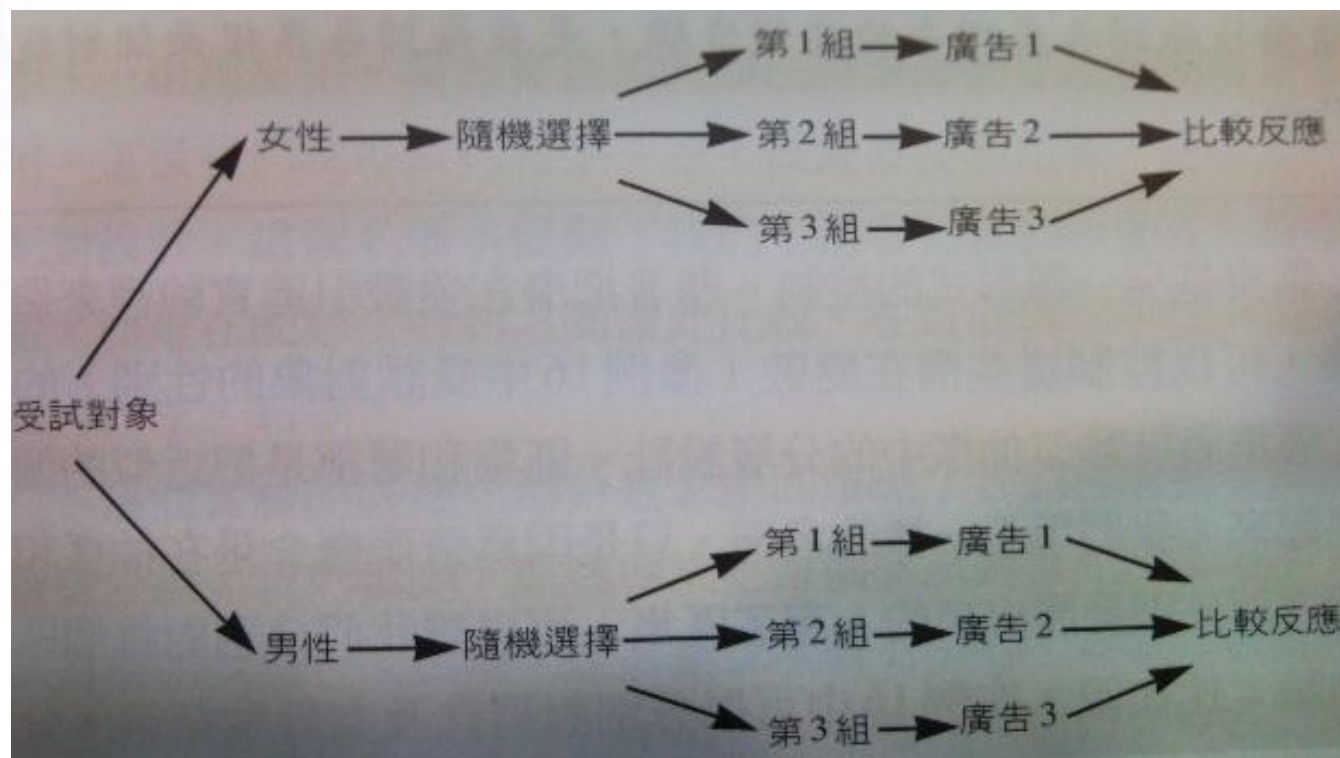
區集設計

- ◆ **區集**是一組實驗個體或受試對象，這些個體或受試對象在被認為**會影響反應的某些方面很類似**。
- ◆ **區集設計**當中，將個體隨機指派到各處理去的這個步驟，是在每個區集裡面個別執行的。



Ex. 男性、女性和廣告

- ◆ 女性和男性對廣告的反應不一樣。有個實驗要比較同一產品的三支電視廣告之效用，除了對這些廣告的整體反應外，也想要知道**男性和女性分別的反應如何**。





配對設計

- ◆ 配對設計是區集設計的一種。
- ◆ 配對設計是比較兩種處理，每個區集只包含兩個儘可能相似的個體。
- ◆ Ex. 左手對右手何者有力實驗
 - (1). 擠壓體重計，做為手力之度量
 - (2). 隨機決定受試者先用哪隻手
 - (3). 左、右手的順序隨機化
 - (4). 避免受經驗效應影響



政策實驗的缺點

- ◆ 公共政策實驗經常會面臨嚴重的實際困難，使得實驗沒法子好好執行。
- ◆ 這類實驗常常很花錢，而且花很多時間來完成。
- ◆ 實驗結果對於公共政策的主要問題影響很小，這些問題是受政治所左右的。



政策實驗的優點

- ◆ 隨機化比較實驗可以提供清楚的證據，顯示政策的改變確實可以造成結果的改變。



道德與實驗

- ◆ 醫藥方面的進展，主要是以後的病人受惠，但風險卻由現在的病人承擔。
- ◆ 給患者施予不確定療效的治療是否道德？
- ◆ 為測試新藥，用安慰劑當控制，是道德還是醫學？還是人命不值醫學？
- ◆ 法律是否代表公平？是否一定得要凌駕醫學與統計？
- ◆ 愛滋病的研究是要治標（速度）？還是要治本（徹底）？



行為及社會科學實驗

- ◆ 行為實驗常常不能讓受試者知道研究的真正目的。
- ◆ 受試者被要求參加，應該被告知什麼？什麼不應該被告知？
- ◆ 是否有自尊受損？情緒是否受到傷害？身體是否受傷？



結論

◆ 統計的技術面

- 統計的**隨機實驗**必須顧全公平及公正原則，以降低隱性偏差。
- 檢視雙盲效應，**避免霍桑效應**。
- 用足夠多的樣本及同時比較數個處理，並謹慎的運用**隨機化原則**，是統計實驗設計的基本原則。
- 透過適當的實驗設計，將可更切近**實驗目的**。

◆ 統計的社會面

- 兼顧**社會道德**
- 尊重**個人隱私**
- 重視**人性尊嚴**
- 尊重**生命**



Thanks for your attention