一、樣本代表性

「樣本代表性」是行為科學界內通行的一個抽象概念，意指抽樣調查的數據是否正確，也正由於此一語彙的抽象性，有些研究者常以研究經驗或批判洞察力來判斷其代表性，較缺乏具體衡量的準則。

「樣本代表性」應有兩類意義，第一類是「統計學、計量分析思想」的意義。

第二類，是「應用統計、抽樣實務」的意義。

二、統計學、計量分析思想的樣本代表性

一般的數學、即 Euclid-Newton 數學，若能發揮測量與預測能力，其對象必須具備「反身性、等加性」，吳統雄特命名為「第1類計量思想」產生的是「第1類知識」。

但生命現象、行為現象，並不具備「反身性、等加性」，倒是可能具備「常態分配」性質，故 Pearson, Fisher 等人開拓了「推論統計學」、原稱「生物統計」方法，吳統雄特命名為「第2類計量思想」產生的是「第2類知識」。

統計學有兩種容易混淆的目的，一種是收集到資料樣本後，測量與報告的統計值，只「限於觀察樣本的本身」，稱為「敘述統計」，沒有推論全體的基礎與價值。

第二種則是要「從樣本『推論』母群（注[1]）」，即預測「全體」的現象，而以樣本為代表，是為「推論統計學」。

如果要「推論」，就是要能「以簡馭繁」，以小樣本觀察、預測母群；但又必須「具體而微」，應有適當的樣本代表性。

而推論的樣本代表性，來自2個基本觀念：常態分配、與中央極限定理。

(一)常態分配

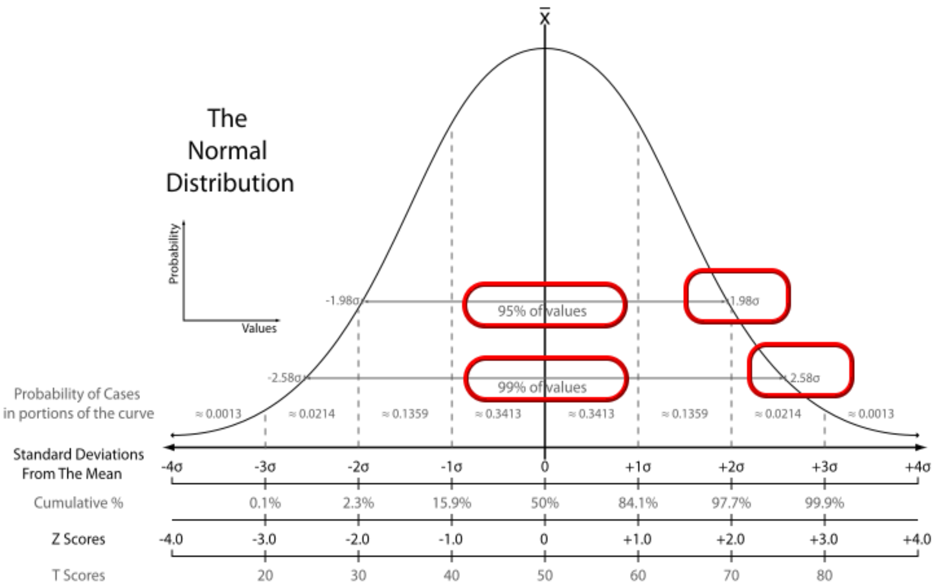
生物界的大多數特質，樣本之間都會呈現常態分配-亦即如下圖般：左右對稱的鐘型曲線。

底部 X 軸表示標準差(Z)。Y值就是觀察特徵值X值對應的樣本數。

(注[1])「母群」(population)即受調查的全體，不少中文書譯為「母體」，如果翻譯除了「信達」之外，也有「求雅」的必要，本研究者建議採用「母群」之譯文。登錄樣本的文件或虛擬架構即為母群清冊，又稱為抽樣對象（sample framework, 或sampling frame）。

函數曲線所覆蓋的面積，就是總樣本數。

底部= 號後的數值，為其垂直範圍占總樣本數的百分比(P)，如正負1個標準差內的草數佔全體68.26%。



Probability= 範圍內樣本數，占母群的百分比。

Standard Deviations= 標準差，亦稱離均差。

Z Scores= Z 分數，就是有幾個標準差，兩者其實相同。

常用Z值有2和2.5。

當Z= 2，單側P= .4772，左右合計為 P≒ .95，即95%的樣本，在2個標準差之內。

當Z= 2.5，單側P= .4938，左右合計為 P≒ .99，即99%的樣本，在2.5個標準差之內。

常態分配有很多深入的啟示，其中1項就是：不要把形象表面的差距，誤以為是真實的差距！

進一步解說，請參考：統雄-統計神掌機率與分配。

(二)中央極限定理。

中央極限定理是推論統計的基礎。

機率的估計值必須在「大數法則」下，才會實現；同時在「大數」時，觀察樣本會呈現「中央極限」現象，這就是我們解釋推論統計的基礎。

中央極限定理(Central limit theorem)係指從平均數為μ，標準差為σ的母群中，隨機/等機率地抽取大小為n 的獨立樣本。當樣本數很大

時，其樣本平均數 $\bar{X}$ 減掉母群平均數、再除以樣本標準差（特稱為標準誤），將會趨近平均數為0，標準差為1的常態分配。

也就是說：所有樣本平均數的集合，會形成一個「虛擬的」「樣本平均數分配的常態分配」。

其定義公式為：

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

	平均數	標準差
觀察對象,樣本	$\bar{X}$	s
（虛擬的） 樣本分配 的標準差	$\mu_{\bar{X}}$	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ $\sigma_{\bar{X}}$ 特稱標準誤 Standard error P (Degree of confidence) 為抽樣把握 ±Z 之間是為 Confidence interval
母群-事實不知的	μ	σ
作推論的參數 (^ 在此表示 estimate of)	$\hat{\mu} = \bar{X}$	$\hat{\sigma}$
母群-估計的區間	$\mu = \hat{\mu} \pm E$	

定義公式就是：

$\bar{x} - \mu = E$, E為觀察值與真實值的差距，即誤差。

Z= 誤差相當於幾個「標準誤」的值

定義公式移項以後就是：

$$E = z\sigma_{\bar{x}} = z.\frac{\sigma}{\sqrt{n}}$$

若 σ不知，則用 $\hat{\sigma}$ 代替 σ
(註：有些文獻上E作d）

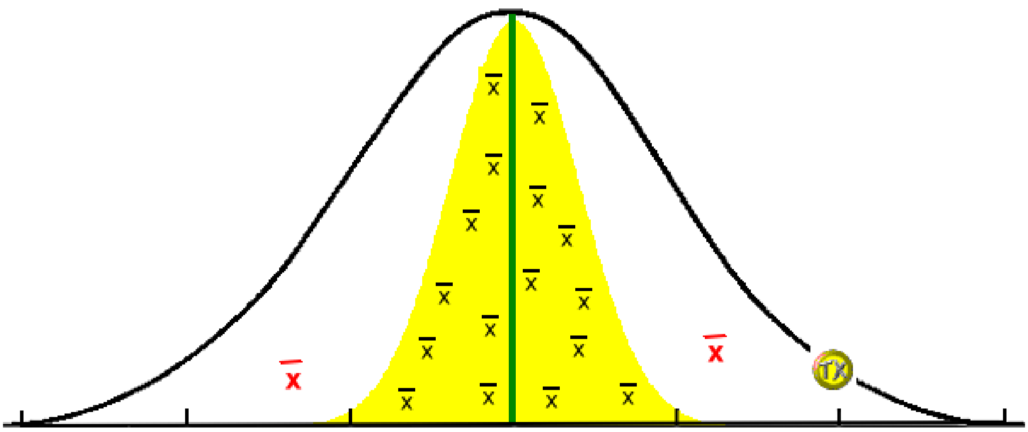
1、中央極限定理的逆向思考應用

以上讀得一頭霧水嗎？很正常！那就是一般教科書呈現的方式。
別慌！統雄神掌中央極限定理圖解來了！
全球獨一無二的視覺解說。

公主招親恐怖箱實驗

參加過統雄老師「公主招親恐怖箱實驗」的同學，曾經體驗過：6位駙馬候選人在同1個恐怖箱（母群）中抽出資料（樣本），彼此的資料並不相同，但事實上箱中的資料卻是相同。實驗與理論是一體的兩面，為了避免誤判，我們就要發展以下的檢定程序。

2、統雄圖解中央極限定理



母群（白色）的平均數為0（綠色），在其中任意抽出樣本，其平均數為x（黑色），其觀察值不為0，是因為抽樣誤差所造成，事實上還是為0；而且所有的樣本，也會呈現常態分配（黃色）。但如果抽到

x（紅色），其觀察值與母群相同的可能性已經小於百分之5或百分之1，這時，我們就推論樣本可能和母群不一樣。

亦即採用「樣本的平均數($\bar{x}$)」來「估計」「母群的平均數(μ)」，再檢定(test) $\bar{x}$和μ差異的程度，這個程度由「樣本平均數分配的標準差」或特稱為「標準誤($\sigma_{\bar{x}}$, standard error)」為單位，如果差異很大，就是樣本代表性較小，如果差異很小，就是樣本代表性較大。

平均數在平均數正負2個、或2.5個標準誤內（視把握度要求多大），常視為與母群平均數（綠線）相同，亦即「樣本不為0，母群為0」。

3、顯著性檢定：逆向思考法

1組樣本自母群隨機/等機率抽出，其特性應與母群一樣。

但實務上，樣本的平均數和母群的平均數不會完全一樣，而會呈常態分配差異。

這不是真的差異，而是抽樣過程的必然：有95%的樣本平均數和母群平均數差異最高可到±2E、有99%的樣本平均數和母群差異最高可到±2.5E，其實它們和母群完全相同。

所以，統計學開拓者 Fisher 就建立了「成對」「統計假設」的概念：[2]

統計假設由1對統計表示式組成，其專業符號為：

H_0：假設相同；如「母群的平均數(μ)」=「樣本的平均數($\bar{x}$)」，是為「反面假設」。

H_1：假設不相同；如「母群的平均數(μ)」≠「樣本的平均數($\bar{x}$)」，是為「正面假設」。

研究者優先假設H_0：「樣本與母群相同」，如果「樣本的平均數($\bar{x}$)」在「母群的平均數(μ)」的 2 個標準誤之內，即樣本看似與母群不同，但有95%的可能，其實與母群相同。同理，如果在 2.5 個標準誤之內，樣本還有99%的可能，仍與母群相同。

此時，研究者就「接受」H_0，推論「樣本與母群相同」。

（注[2]）H_0 為 Null Hypothesis，一般譯為虛無假設。

H_1 (有些書作 H_A) 為 Alternative Hypothesis，一般譯為對立假設。

這樣的直譯，似乎不能達意。

Fisher 當初創造 Null 這個詞，意指「預設『無關』（即為 0）」。而我們去作一個實驗，通常是希望發現「有關」。所以吳統雄建議譯為「反面假設」與「正面假設」。

但如果「樣本的平均數($\bar{x}$)」在「母群的平均數(μ)」的 2 個標準誤之外，即樣本仍與母群相同的機率已低於5%。同理，如果在 2.5 個標準誤之外，樣本還與母群相同的機率更低於1%。只作1次實驗，發生「樣本與母群相同」的可能性就偏低。

此時，研究者就「拒絕」H_0，而改為「接受」H_1，推論「樣本與母群不同」。

經由中央極限定理，可以推導：若無法在高機率下證明「樣本與母群相同」，則應可「反證」「樣本與母群不同」。

這也是統計假設必須是「反證法」的原因，這種逆向思考法，就稱為：顯著性檢定。

如果到達95%的把握(β)-相反就是低於5%的風險(α)-專業上稱為「到達.05顯著水準」；同理，99%的把握，專業上則稱為「到達.01顯著水準」，可能「樣本與母群不同」。

4、中央極限定理的功能：可否推論

衡量樣本和母群的差異是否到達「顯著水準」。

「顯著水準」就是檢定「樣本數會不會太少」，樣本觀察值是否是抽樣誤差所造成的。也所以，若未到達顯著水準，就沒有推論的意義。

相對的，若到達顯著水準，也不一定有「重要」 的意義。

若調查發現：NBA西區的球員平均身高比東區的差異到達「.01顯著水準」，只是表示兩者差異係因抽樣數導致的誤差小於百分之1，西區的球員平均身高極可能確實比東區高。但並不能預測西區就會贏球，即使西區贏球了，也不一定能證明身高是贏球的原因。

因為進一步分析，西區的球員平均身高只比東區高不到0.5公分，根據所有歷史資料，在這個高度差距下，身高不會是贏球的重要原因。

但若NBA的球員平均身高比泰國隊的差異也到達「.01顯著水準」，且差異超過15公分，那麼身高不僅是「顯著」、而且可能是「重要」贏球的原因。

許多論文寫「本研究假設甲與乙有顯著差異。」

這樣的敘述不僅不清楚，而且容易引起誤會，作者可能不知道什麼是「顯著」。

以上這個敘述如果放在「研究目的」章中，「顯著」2字有點畫蛇添足，好像研究目的是要證明樣本數不會太少。最好改為「理論敘述」形式。

如果放在「研究發現」章中，國際上完整的敘述是「本研究發現甲與乙在『某應變項』上的差異，到達.05(或.01)顯著水準。」顯著水準是個相對觀念，不是絕對觀念，不能省略其數值。

上段的敘述如果放在「研究結論」章中，而沒有進一步的詮釋，那就是 So what? 有差異，又怎樣？只是資料的展示，並沒有知識的產生。

5、雙尾與單尾檢定

本圖解將「把握」放中間，太大與太小的樣本平均數都視為到達顯著水準，稱為「雙尾檢定」。如果只檢定太大、或只檢定太小的樣本，則稱為「單尾檢定」。

6、型1錯誤 Type Ⅰ error 與 型2錯誤 Type Ⅱ error

不顯著就拒絕的方法，稱為避免犯「型1錯誤(Type Ⅰ error)」：把事實上樣本與母群相同，誤為與母群不同。

但如果事實上樣本與母群是有差異，但被錯誤拒絕了，則稱為犯「型2錯誤(Type Ⅱ error)」。即樣本與母群不同，卻誤為與母群相同。

研究均以保守為先，故通常避免「型1錯誤(Type Ⅰ error)」。這是由聶曼（Jerzy Neyman）和伊根・皮爾遜（Egon Pearson, 卡爾・皮爾遜之子）所共同建議的觀念。

7、進階討論

Ⓐ當然，「顯著水準」其實是灰階現象，訂定絕對切點有其困難；但常態分配在1個標準差點，有「反曲」-斜率正負相反現象，所以，在標準差整點切割已經是人為的、相對的最佳選擇。

Ⓐ中央極限定理是由「大數定律（又稱大數法則，Law of large numbers）」所發展出來的，意指數量越多，則其平均就越趨近期望值。最早是由瑞士 Jacob Bernoulli （雅客•博努力）所提出來的「博努力實驗 (Bernoulli Trail)」，就是投幣極多次，正反面出現的比例，各為50%，這是以二元資料的實驗證明，後人又發展至連續資料的證明。

Ⓐ在數學、統計、甚至物理研究上，許多地方會出現 Bernoulli 的名字，但不是同一個博努力，而是「博努力家族」，雅客與其 2 個弟

弟是第一代，其後至少 7 代都有出現著名的博努力，而多半出自 3 弟 Johann Bernoulli 雅漢後代。在政治世家奪權鬥爭不息，而在這個學術世家，為創作爭名也經常上演，不僅兄弟爭，連父子也爭，雅漢還和嫉妒天才兒子 Daniel Bernoulli 丹牛爭，和兒子斷絕關係。

丹牛最著名的是博努力定律（Bernoulli's principle），亦即流體力學的基礎定律。不過，統雄老師特別推薦他的「預期效用論（Expected Utility Theory）」，開啟了現代計量經濟學，也是人類行為研究的開拓性思想。