

68-95-99.7 rule
F-scale
Minitab
P-value
law of large numbers
lack of reliability
unbiased estimation
central limit theorem
median



Simpson's paradox
fixed market basket price index
seasonal variation
sampling distribution
histogram
social statistics
bar graph
non-sampling errors
confidence interval
insurance
constant dollars
correlation

2012年最新修訂版

boxplot
significance, statistical
stemplot (stem-and-leaf-plot)
processing error
least-squares regression
odds
scatterplot
Pew Research Center

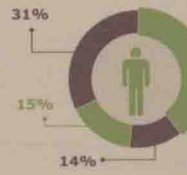
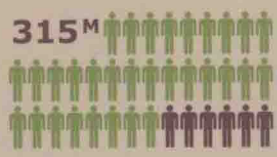
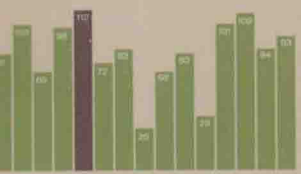
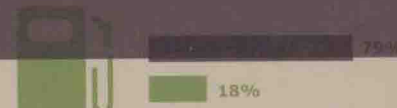
統計學的世界 III

機遇 · 推論

STATISTICS
CONCEPTS AND CONTROVERSIES
SEVENTH EDITION

BY DAVID S. MOORE,
WILLIAM I. NOTZ
淡江大學數學系教授 鄭惟厚 譯

79
折優惠
天下文化

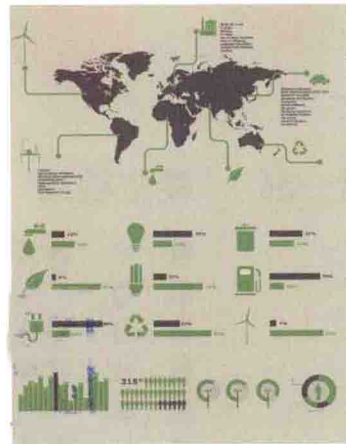


知識的世界 BW1106

統計學的世界 III

機遇 · 推論

STATISTICS
CONCEPTS AND CONTROVERSIES
SEVENTH EDITION



by David S. Moore, William I. Notz
淡江大學數學系教授 鄭惟厚 譯

統計學的世界 III

原 著 / 墨爾、諾茨
 譯 者 / 鄭惟厚
 顧 問 群 / 林和、牟中原、李國偉、周成功
 科學館總監 / 林榮崧
 責任編輯 / 林文珠 (第一版)、畢馨云 (第二版)
 封面設計暨美術編輯 / 江儀玲

出版者 / 天下遠見出版股份有限公司
 創辦人 / 高希均、王力行
 遠見·天下文化·事業群 董事長 / 高希均
 事業群發行人 / CEO / 王力行
 出版事業部總編輯 / 王力行
 版權部經理 / 張紫蘭
 法律顧問 / 理律法律事務所陳長文律師 著作權顧問 / 魏啟翔律師
 社址 / 台北市 104 松江路 93 巷 1 號 2 樓
 讀者服務專線 / (02) 2662-0012 傳真 / (02) 2662-0007 ; 2662-0009
 電子信箱 / cwpc@cwgv.com.tw
 直接郵撥帳號 / 1326703-6 號 天下遠見出版股份有限公司

排 版 廠 / 極翔企業有限公司
 製 版 廠 / 東豪印刷事業有限公司
 印 刷 廠 / 立龍藝術印刷股份有限公司
 裝 訂 廠 / 晨捷印製股份有限公司
 登 記 證 / 局版台業字第 2517 號
 總 經 銷 / 大和書報圖書股份有限公司 電話 / (02) 8990-2588
 出版日期 / 2002 年 10 月 30 日第一版
 2012 年 6 月 22 日第二版第 1 次印行
 定 價 / 400 元

國家圖書館出版品預行編目 (CIP) 資料

統計學的世界 / 墨爾 (David S. Moore), 諾茨 (William I. Notz) 作; 鄭惟厚譯. -- 第二版. -- 臺北市: 天下遠見, 2012.06

冊: 公分. -- (知識的世界; 1104-1106)

譯自: Statistics: concepts and controversies, 7th ed.

ISBN 978-986-216-959-9 (第 1 冊: 平裝). --

ISBN 978-986-216-960-5 (第 2 冊: 平裝). --

ISBN 978-986-216-961-2 (第 3 冊: 平裝)

1. 統計學

510

101009979

Statistics: Concepts and Controversies, Seventh Edition

by David S. Moore and William I. Notz

First published in the United States by W.H. FREEMAN AND COMPANY, New York

Copyright © 2009, 2006, 2001, 1997 by W. H. Freeman and Company

Complex Chinese Edition Copyright © 2002, 2012 by Commonwealth Publishing Co., Ltd.,

a member of Commonwealth Publishing Group

Published by arrangement with W. H. Freeman and Company through

Bardon-Chinese Media Agency.

All rights reserved.

ISBN: 978-986-216-961-2 (英文版 ISBN: 978-1-4292-2991-3)

書號: BW1106

BOOKZONE 天下文化書坊 <http://www.bookzone.com.tw>

簡明目錄 **Brief Contents**



第一部	產生數據	19
第 1 章	數據從何而來？	21
第 2 章	好樣本和壞樣本	43
第 3 章	樣本告訴我們什麼？	63
第 4 章	真實世界中的抽樣調查	89
第 5 章	實驗面面觀	121
第 6 章	真實世界中的實驗	147
第 7 章	資訊倫理	175
第 8 章	度量	201
第 9 章	數字合不合理？	229
第一部	復習	251

第二部	整合數據	265
第10章	好的圖及壞的圖	267
第11章	用圖形呈現分布	297
第12章	用數字描述分布	323
第13章	常態分布	355
第14章	描述關聯的方法：散布圖和相關係數	381
第15章	描述關聯：迴歸、預測及因果關係	409
第16章	消費者物價指數和政府統計	441
第二部	復習	467
第三部	機遇	511
第17章	考慮可能性	513
第18章	機率模型	537
第19章	模擬	555
第20章	賭場的優勢：期望值	577
第三部	復習	597
第四部	推論	607
第21章	什麼是信賴區間？	609
第22章	什麼是顯著性檢定？	641
第23章	統計推論的使用與濫用	669
第24章	雙向表及卡方檢定	689
第四部	復習	717

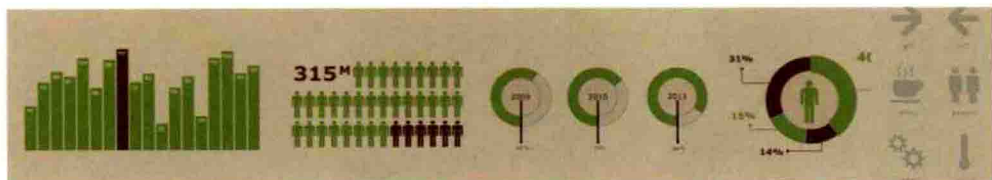


統計學的世界 III

目錄 Contents



第三部	機遇	511
第17章	考慮可能性	513
個案研究	513	機率概念 514
機遇的古代史	517	關於機遇結果之迷思 519
個人機率	525	機率與風險 527
本章重點摘要	528	個案研究隨堂考 529
第17章習題	530	網路尋奇 534
注釋與資料出處	535	
第18章	機率模型	537
個案研究	537	機率模型 538
機率規則	539	抽樣的機率模型 543
本章重點摘要	547	個案研究隨堂考 548
第18章習題	548	網路尋奇 554
注釋與資料出處	554	
第19章	模擬	555
個案研究	555	機率從何而來？ 556
模擬入門	557	考慮獨立性 561
更複雜的模擬	563	本章重點摘要 567
個案研究隨堂考	567	第19章習題 568
網路尋奇	575	注釋與資料出處 576
第20章	賭場的優勢：期望值	577
個案研究	577	期望值 578
大數法則	582	深入探討期望值 583
統計學上的爭議：合法賭博面面觀	584	用模擬計算期望值 585
本章重點摘要	587	個案研究隨堂考 587
第20章習題	588	網路尋奇 595
注釋與資料出處	596	



第三部復習 597

第三部重點摘要 598 第三部復習習題 599

第三部報告作業 604 注釋與資料出處 606

第四部 推論 607

第21章 什麼是信賴區間？ 609

個案研究 609 估計 610 有信心的估計 612

了解信賴區間 617 母體比例信賴區間之再補充 620

樣本平均數的抽樣分布 624 母體平均數的信賴區間 627

本章重點摘要 629 個案研究隨堂考 629

第21章習題 630 網路尋奇 638 注釋與資料出處 639

第22章 什麼是顯著性檢定？ 641

個案研究 641 統計顯著性檢定的論理基礎 642

假設及P值 646 統計顯著性 650 計算P值 651

母體平均數的檢定 653 本章重點摘要 658 個案研究隨堂考 658

第22章習題 659 網路尋奇 667 注釋與資料出處 667

第23章 統計推論的使用與濫用 669

個案研究 669 聰明做推論 670 顯著性檢定面臨的難處 672

信賴區間的優點 676 「水準5%之顯著性」並非魔術指標 677

統計學上的爭議：應不應該禁止統計檢定？ 678 提防刻意尋找的顯著性 679



本章重點摘要	681	個案研究隨堂考	681
第23章習題	682	網路尋奇	686
		注釋與資料出處	687
第24章 雙向表及卡方檢定			689
個案研究	689	雙向表	690
		雙向表的推論	692
卡方檢定	695	如何應用卡方檢定	701
		辛浦森詭論	704
本章重點摘要	707	個案研究隨堂考	707
第24章習題	708	網路尋奇	714
		注釋與資料出處	714
第四部復習			717
第四部重點摘要	719	第四部復習習題	721
第四部報告作業	729	注釋與資料出處	730
解決統計學上的爭議			733
「輪到你練習了」習題解答			740
習題解答 (單數題)			747
表A 隨機數字			755
表B 常態分布的百分位數			757

第三部 機遇

A stylized map of East Asia, including parts of China, Korea, and Japan, rendered in a light yellow color. Several green circular markers are placed on the map, with thin lines connecting them to circular callout boxes on the right. The callout boxes contain the numbers 95, 45, 26, and 18, arranged from top to bottom. The background of the entire page is a solid light yellow color.

95

45

26

18

「如果我有當國王的機會，那麼我就有戴上皇冠的命。」馬克白在莎士比亞劇中如是說。機會或機遇的確會作弄我們每個人，我們卻沒多少能耐去了解或者操控機會。不過有的時候，機遇被馴服了。擲骰子、簡單隨機樣本，甚至得自遺傳的眼珠顏色或者血型，代表已經安分下來的機遇，是我們可以了解並且操控的。不像馬克白或者我們的一生，骰子是可以一擲再擲的。結果是由機遇決定，但是在多次重複之後，會有某種模式（或型態）出現。而因為我們可以描述它的模式，也使得機遇不再神祕莫測了。

不論是幾何裡的圓和三角形，還是行星的運行，我們人類都會用數學來描述有規律的模式。當我們可以不斷重複某種機遇現象，使得機遇變得可掌握時，我們就會利用數學來了解這個機遇行為的規律模式。探討機遇的數學叫做機率。機率就是本書第三部的主題，不過我們會少講數學，而著重在實驗和思考。



考慮可能性

個案 研究

在2007年10月5日這一天，俄亥俄州瑪里斯維爾城的一位婦女生了她的第三胎。這個嬰兒的生日和她前兩個孩子相同。美聯社獲知此事，結果全國許多報紙都登載了這項令人驚異的巧合。如果生日是隨機的，而且彼此之間互相獨立的話，機率學家可以幫我們算出來，三個孩子同一天生日的機率，大約是百萬分之7.5。就是因為這樣的事鮮少發生，所以才值得當作新聞。

這件事到底有多麼稀奇呢？在這一章我們會學習如何解讀像百萬分之7.5這樣的機率。讀完本章之後，你就有能力評估像三個孩子同一天生日這類事件的巧合程度。這件事是否如同表面看來這樣令人驚訝呢？



可能性（或機會、機遇）是不大容易掌握的題目。我們就從「如果我們試很多次，會發生什麼狀況」開始思考。我們也會先討論像擲銅板時有二分之一機會得到正面這樣的例子，然後再探討更複雜的情境。

就連美式足球規則也都同意，擲銅板可以避免偏袒。抽樣調查挑選受訪對象時，或者醫學試驗將病人分配到處理組或安慰劑組時，如果有偏袒，就像美式足球比賽一開始要決定球先給哪一隊時有所偏袒一樣，都是不能接受的。這就是為什麼統計學家建議使用隨機樣本及隨機化實驗，這些只是擲銅板的花俏版本。如果我們仔細觀察擲銅板或者隨機樣本的結果，一件重要事實就會浮現：**短期機遇現象無法預測，但是長期下來，會呈現有規則且可預測的模式。**

擲一枚銅板或者選擇一個簡單隨機樣本，都無法在事前預測結果；因為你如果重複擲銅板或選樣本，結果本來就會次次不同。但是還是可以在結果裡面看到某種規律的模式，而且只有在重複許多次以後，這個模式才會清楚浮現。這個了不起的事實，就是機率概念的基礎。

例1 擲銅板

當你擲銅板的時候，結果只有兩種可能：正面或者反面。圖17.1顯示擲銅板1,000次的結果。對應從1到1,000次的每一擲，我們都將擲出正面的比例畫在圖上。第一次擲出正面，所以正面比例的第一個值是1。第二次擲出反面，所以在兩擲之後，正面比例降為0.5。再接下去的四次都是反面，之後擲出一個正面，所以擲七次之後的正面比例是 $\frac{2}{7}$ ，即0.286。

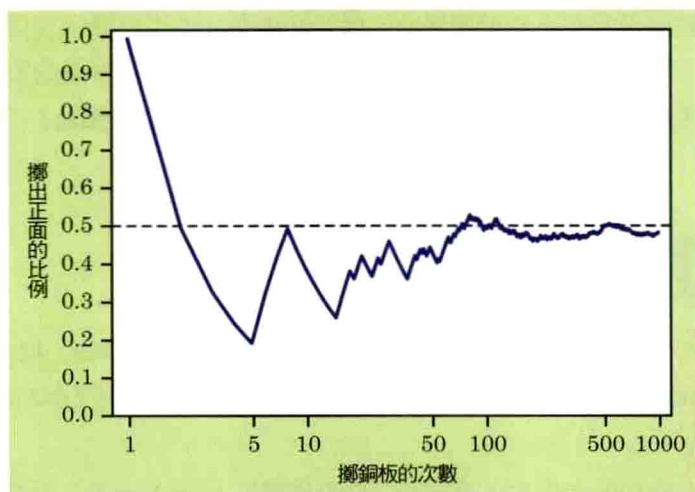


圖 17.1 擲銅板許多次，擲出正面的比例會隨著我們擲的次數而改變，但最終會非常接近0.5。這就是我們說「正面機率是一半」的意思。

剛開始的時候，擲出正面的比例變化很大，但是擲的次數愈來愈多，就慢慢穩定下來。這個比例到最後會靠近0.5，而且會一直維持在0.5附近。我們把這個0.5叫做得到正面的機率。0.5這個機率在圖上是以水平線表示出來。

在統計裡的「隨機」，並不是「偶然」的同義詞，而是在描述某種長期下來才會出現的規則。我們在每一天的生活經驗裡，都會碰到隨機不可預測的那一面，但我們很少有機會能重複觀察同一個隨機現象非常多次，次數多到能夠看出機率所描述的長期規律模式。你可以在圖 17.1 中看到規則出現，很長期來看，擲出正面的比例是0.5，這是機率的直覺概念。機率0.5代表「試驗很多很多次時，有一半時候會發生」。

我們可能僅僅因為銅板只有兩面，就會猜正面機率是0.5了。但是嬰兒的性別也只有兩種可能，機率卻不一樣——男嬰的機率差不多是0.51，而不是0.50。這個機率概念是根據經驗法則而來，也就是說，是

根據數據得來而不是根據理論。機率描述在許多次試驗中會發生什麼狀況，而且我們必須真的觀察許多次擲銅板的結果、許多的嬰兒，才能夠抓到這個機率。說到擲銅板，有些勤快的人還真的擲過上千次呢。

例2 幾位擲銅板出了名的人

十八世紀的法國自然學家布方伯爵（Count Buffon）把銅板擲了4,040次。結果：2,048次正面，或者說正面比例是 $2,048/4,040 = 0.5069$ 。

大約1900年左右，英國統計學家皮爾生（Karl Pearson）很神勇的擲一個銅板24,000次。結果：12,012次正面，比例0.5005。

南非數學家柯瑞屈（John Kerrich）在第二次世界大戰被德國人囚禁的時候，擲了銅板10,000次。結果：5,067次正面，比例0.5067。

▶ 隨機及機率

如果一個現象的個別結果無法預知，然而在多次重複之後，其結果會出現有規則的分布，則我們稱該現象為**隨機的（random）**。一個隨機現象任一結果的**機率（probability）**是在0和1之間的一個數字，該數字描述在重複許多次的情形下，該結果應會出現的比例。

機率為0的結果從來都不會發生，而機率為1的結果則每重複一次就發生一次。機率為1/2的結果，在一長串的試驗當中，大約有一半時間會發生。當然我們永遠沒辦法確實觀察出一個機率。比如說，銅板不管擲了多少次都可以再擲。而數學的機率是一種理想化的描述，根據的是推想在一串無休無止的試驗當中會發生的狀況。

我們沒有要在這裡做深入的研究。隨機現象的存在，只不過是我

們觀察這世界所得到的事實。機率也不過是用來描述隨機現象長期規律性的語言。擲一次銅板的結果、放射源發射出粒子的間隔時間，以及實驗室老鼠生的下一胎小老鼠的性別，都是隨機的。隨機樣本或者隨機化實驗的結果，也一樣是隨機的。一大群人的行為，也常常和擲多次銅板或取多個隨機樣本的結果差不多隨機。舉例來說，人壽保險根據的事實就是：在一大群人裡面，死亡是隨機發生的。

例3 死亡的機率

我們沒法預測某個人是不是明年會死亡。但是如果我們觀察好幾百萬人，死亡就是隨機的了。美國國家衛生統計中心公布，20歲到24歲的男性當中，在任何一年中會死亡的比例差不多是0.0014。這是一個年輕男性明年會死亡的機率。對於同年齡層的女性，死亡機率大約是0.0005。

如果一家保險公司賣出許多壽險給年齡在20到24歲之間的人，公司會知道：賣給男性的保險明年大約有0.14%要理賠，賣給女性的大約有0.05%要理賠。而因為男性理賠的比例比較高些，所以保費會收得多一點。



上帝擲骰子嗎？

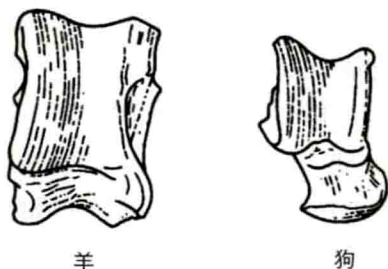
世界上很少有事情是百分之百的隨機，以致於不管我們有多少資訊，都沒法預測結果。比如說，理論上我們可以把物理定律應用在擲銅板上面，來計算會擲出正面還是反面。但是在個別的原子內部，隨機性的確規範了事件的發生狀況。愛因斯坦不大喜歡新興的量子論的這項說法。「上帝可沒有在和宇宙玩骰子，」這位偉大的科學家說。但許多年之後看來，愛因斯坦顯然錯了。



機遇的古代史

在玩需要多次重複的機遇遊戲如擲骰子、發洗好的牌、轉輪盤等時，我們最容易注意到隨機性。類似這些遊戲的機遇工具，曾在遠古時代用來找尋神的旨意。西方古時候最常用來隨機化的方法是「擲骨頭」，就是擲好幾塊距骨。距骨（圖17.2）是相當規則的實心骨頭，取

圖17.2 動物的腳跟骨（距骨）。
(From F. N. David, *Games, Gods and Gambling*, Charles Griffin & Company, 1962.
Reproduced by permission of the publishers.)



自動物腳跟。擲了之後，等距骨靜止，四面中的其中一面會朝上（其他兩面是圓形的）。用陶土或骨頭做的立方體骰子後來才有，但即使是骰子，在西元前2000年之前就已存在。跟占卜術比起來，用擲骨頭或擲骰子來賭博幾乎算是較近代的發展。大約西元前300年前的時候，還沒有這種「墮落行為」的明確紀錄。賭博是在羅馬時代大為風行，後來在基督教的不認同之下（和占卜術一起）暫時消退。

自有歷史以來，像距骨這樣的機遇工具就一直有人使用了。但是古代的偉大數學家，沒有一個人研究了擲骨頭或擲骰子很多次得到的規律模式。也許是因為距骨及大部分的古代骰子形狀不夠規則，使得每一個的結果都有不同的模式。又或者原因是更深層的，是因為傳統上不情願做有系統的實驗。

職業賭徒不像哲學家或數學家那麼壓抑自己，他們注意到擲骰子或發牌的結果有規則模式，並試著調整賭注來增加贏錢機會。「我該怎麼賭？」這個問題，就是機率理論的起源。對於隨機的有系統研究，始於（我們有一點過度簡化）十七世紀法國賭徒請法國數學家幫忙算出機遇遊戲的「公平」賭注之時。機率理論也就是探討隨機的數學研究，是十七世紀時從費馬（Pierre de Fermat）及巴斯卡（Blaise Pascal）開始的，到二十世紀統計學家接手的時候，機率理論已經發展得很完善。



關於機遇結果之迷思

機率的觀念似乎很直截了當。它對「如果我們這樣做很多次，會發生什麼狀況？」這個問題，提供了答案。但事實上不論是隨機現象的「表現」，還是機率的概念，都有很微妙的地方。我們不斷碰到機遇結果，而心理學家告訴我們，我們處理得不高明。

短期規律性的迷思

機率的概念是，隨機現象長期來說是有規則的。不幸的是，我們對於隨機的直覺卻是說，隨機現象應該在短期就有規則。當規則沒出現時，我們就會尋求解釋，而不把它當作是機遇變異。對了解機遇而言，我們的直覺會給我們很爛的指引。

例4 什麼看起來像隨機的？

把一個銅板擲六次，並且把每次是正或反面記錄下來。以下哪個結果比較可能發生？

正反正反反

反反反正正

幾乎每個人都說「正反正反反」比較容易發生，因為反反反正正「看起來不隨機」。事實上，兩者發生的機會一樣大。正面和反面機會均等的意思只是說：擲了很長一串的結果裡，應該大約有一半是正面；可沒有說只擲很少次時正反就應該差不多是間隔發生。銅板沒有記憶，又不知道前幾次擲了什麼，沒法試著製造出一串平衡的結果。

擲六次銅板得到「反反反正正」這樣的結果，看起來不尋常，是因為有連續3個反面和連續3個正面。連續出現同樣的東西好像在直覺上「不隨機」，但實際上常發生。以下這個例子比擲銅板還要令人印象深刻。