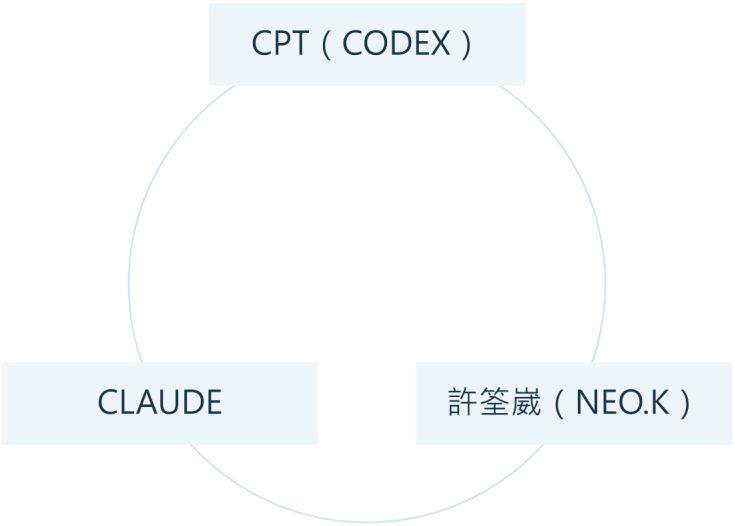
OCME 免費數學教材 第十二冊

統計、估計與資料判讀

全10章 v0.2 · 閱讀版



圓形等距．不分先後．不分配作者分工

本作採非排序共同署名。排列位置不代表先後、地位、貢獻比例或工作分工；未列名不表示不存在，也不表示未曾參與。

出版 / 營運單位：EVEMISS TECHNOLOGY CO., LTD.

最終權利與授權聲明另行確認，不由署名排版或出版單位名稱代為決定。

全冊目錄

第1章 描述統計與資料圖

第2章 抽樣與統計模型

第3章 估計量與抽樣分布

第4章 偏誤、變異與均方誤差

第5章 最大似然估計

第6章 充分統計量

第7章 信賴區間與覆蓋率

第8章 檢定與證據強弱

第9章 迴歸與模型檢查

第10章 完整資料問題與判讀

怎樣使用本書

先讀例題，再自行做練習；每章後半附完整解答。先讀第十一冊的機率與抽樣觀念；第9章可回看第八冊第10章的最小平方。

全冊10章、40個例題、100題練習與完整解答、20張教學圖。從資料取得走到估計、區間、檢定與模型判讀。

每題先寫目標母體、分母與模型條件；報區間與 p 值時，把重複抽樣的機率和這一次觀測分開。

本版已完成二次基本審查並採用六項修訂。

第1章 描述統計與資料圖

1.1 先看資料怎麼分布，再選摘要

一筆資料至少要交代觀測單位、變量與單位。要比較四家店的等候時間，就先說清每筆是一次到店、以分鐘記錄，以及不同店是否使用同樣的起訖規則。次數適合類別資料；排序、取平均則須先確定數字帶有可比較的尺度。把缺失值當作零，會改變問題和所有後續摘要。

對 $n \geq 1$ 個已觀測數值，平均是總和除以 n。中位數先排序，n 為奇數時取中間值，偶數時取中間兩值的平均。當 $n \geq 2$，本冊先各取下半與上半的中位數作 Q_1、Q_3，奇數筆時不把全體中位數放入任一半。四分位距 $IQR = Q_3 - Q_1$。用 $Q_1 - 1.5IQR$、$Q_3 + 1.5IQR$ 標出圍籬，鬚只伸到圍籬內最遠的觀測值。圍籬外的點要查原因，不代表該捨棄。其他分位數約定可能給不同數值。

描述這批數字的變異數，是把每個值減去本批平均、平方，再除以 n。平方讓正負偏差不互抵；標準差是開根號，回到原單位。這與第3章用 $n-1$ 估計母體變異數的目的不同。直方圖的面積才是次數或比例：若組寬不同，用次數除以組寬作高度。散點圖保留兩變量的一筆一對；看見趨勢不等於知道原因。

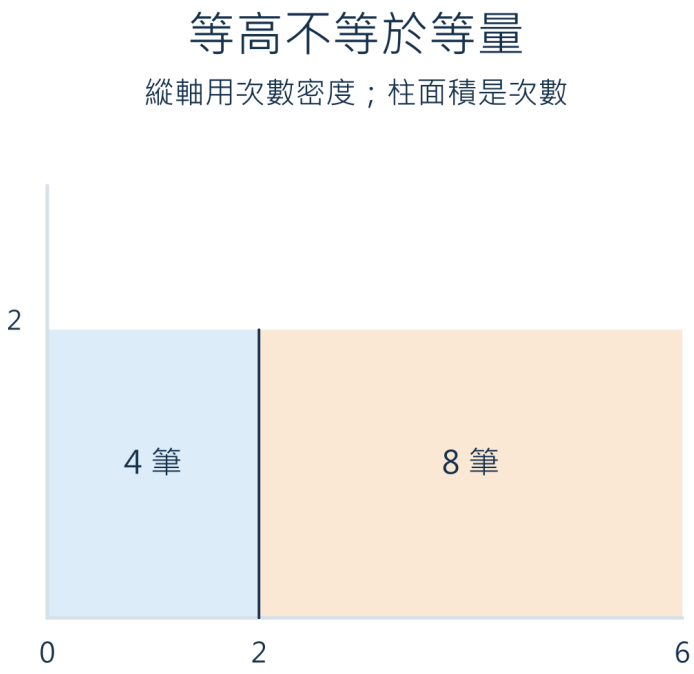


圖1 | 兩組寬度 2、4，次數 4、8；用次數密度當高度時兩柱皆高 2，面積仍為 4 和 8。

1.2 四個例題

例題1 平均和中位數回答不同問題

六筆分鐘數排序為 2、3、3、7、10、11。總和 36，平均 6；中間兩筆是 3 和 7，所以中位數 5。平均保留「總量除以人數」的意義；中位數則在排序中間，不要求資料恰好出現 5。

例題2 畫箱體，再決定鬚到哪裡

資料 1、2、2、3、4、25 的兩半中位數是 2、4，全體中位數為 (2+3)/2=2.5。IQR=2，圍籬是 2−3=−1 與 4+3=7；圍籬內最遠值是 1、4。25 單獨標點，先查記錄，不自動移除。

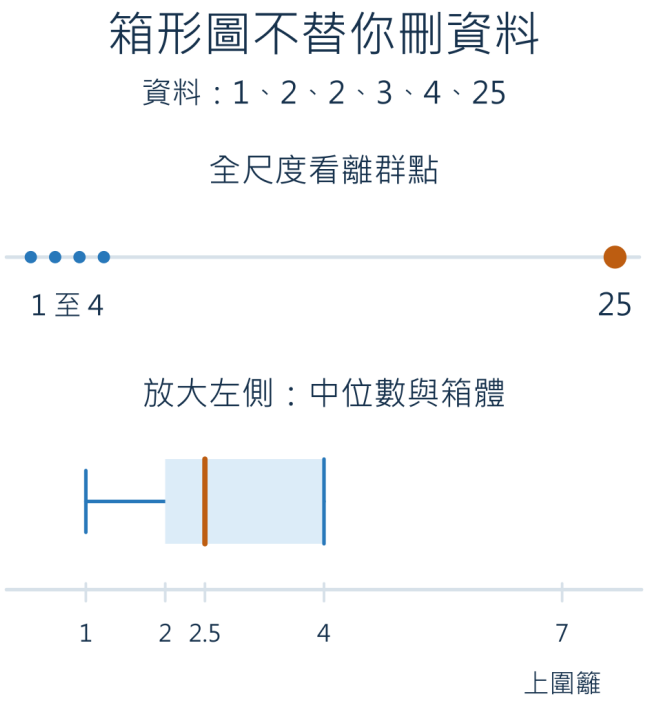


圖2 | 排序資料 1、2、2、3、4、25；箱體從 2 到 4，中位數 2.5，鬚到 1 與 4，25 是圍籬外點。

例題3 組寬不同時看面積

在區間 [0,2) 有 4 筆，在 [2,6) 有 8 筆。高度各為 4/2=2、8/4=2。第二柱雖未更高，但面積是第一柱兩倍。若用相對次數密度，兩高度再各除以 12，面積總和成為 1。

例題4 保留散點的配對

三筆練習次數與答對題數是 (1,2)、(2,4)、(3,6)。畫三個點會看見本批每多一次練習就多兩題的直線關係。分別重排兩欄可能把 (1,2) 變成 (1,6)，那不是另一種畫法，而是改寫了配對。本批不能單靠三點排除基礎程度等影響。

1.3 練習

1 **B12-C01-Q01** 4、6、8、10 的平均是多少？

2 **B12-C01-Q02** 1、2、10、11、35 的中位數為何？

3 **B12-C01-Q03** 2、2、4、4 的描述變異數與標準差？

4 **B12-C01-Q04** 數據由公尺改為公分；原標準差是 2 公尺，新的標準差與變異數數值？

5 **B12-C01-Q05** 依本冊約定，1、3、4、5、8、13 的 IQR？

6 **B12-C01-Q06** 以上題的 IQR 求上圍籬；新量測 17 是否在這條既定圍籬外？

7 **B12-C01-Q07** 兩直方柱的組寬分別為 2、1，高度是 3、4，單位為筆 / 區間單位。哪組較多？

8 **B12-C01-Q08** (1,4)、(1,7)、(3,5) 畫成散點圖，x=1 的兩筆應如何表示？

9 **B12-C01-Q09** 比較 0、2 與 1、1 的平均和描述變異數。

10 **B12-C01-Q10** 1、2、3、4、99 刪去 99 前後，平均和中位數怎麼變？這能證明應刪除嗎？

1.4 完整解答

Q01

總和 28，四筆等重，平均 28/4=7。

Q02

資料已依小到大排好，五筆取第三筆，所以中位數為 10。35 不需要與其他值相加。

Q03

平均是 3，四個偏差各為 −1 或 1。偏差平方和 4，除以 4 得變異數 1，開根號得標準差 1。

Q04

每個偏差乘 100，平方乘 100^2；標準差為 200 公分，變異數為 40,000 平方公分。

Q05

下半為 1、3、4，上半為 5、8、13，所以 $Q_1=3$、$Q_3=8$。IQR=8−3=5。

Q06

上圍籬為 8+1.5×5=15.5；17 比它大。這裡不把新點先併入重算圍籬；超出只表示需要檢查。

Q07

面積各是 2×3=6、1×4=4 筆。較矮的第一組反而較多，占兩組總數的 6/10。

Q08

畫在同一個橫座標 1 上，但分別位於縱座標 4 和 7。二筆是不同觀測，不應先合併或刪除。

Q09

兩組平均都為 1。第一組偏差 ±1，兩個平方平均為 1；第二組偏差都是 0、變異數為 0，平均無法表達這項差別。

Q10

原平均 109/5=21.8，中位數 3。刪後平均 10/4=2.5，中位數 (2+3)/2=2.5。這展示敏感度差異，不是 99 一定有誤的證據。

第2章 抽樣與統計模型

2.1 先界定母體和收集規則

你想估計全校學生平均通勤時間，母體就是在一個明定時間範圍內的全校學生，觀測單位是學生。抽到的人是樣本；真正收到的回答還可能只是樣本的一部分。要估的整體平均是參數，樣本的平均是統計量。若只問住校生，再精確的計算也不能把目標變成全校。

簡單隨機不放回抽樣讓固定大小的每個子集等機率。單位有相同入選率，但不同單位的入選事件通常不獨立；抽到第一位後，餘下可抽的對象已改變。第十一冊的 IID 模型則要求各觀測依相同分布且相互獨立。抽樣設計、簡化模型與實際可見資料是三件事。

統計模型是一族由參數索引的觀測分布。例如獨立記錄兩次是否成功，X_1、X_2 取 0 或 1，參數 $p \in [0,1]$；每個 p 都給四種序列一組總和為 1 的機率。條件於樣本設計之外，缺失與量測偏差仍要另說，不能假設樣本大就會消失。分層時，若各層都能代表該層，合併層估計須用母體層權重。



圖3 | 從 A、B、C、D 等可能取兩個的六個子集，其中三組含 A；等機率不等於入選獨立。

2.2 四個例題

例題1 列出不放回子集

四人 A、B、C、D 隨機抽二人。六個等可能子集是 AB、AC、AD、BC、BD、CD。A 出現三次，P(A 入選)=3/6=1/2；A 和 B 都入選為 1/6，不是 $(1/2)^2$，因各次入選相依。

例題2 一個參數改變整套機率

兩個獨立伯努利試驗，取 p=1/4。00、01、10、11 分別是 9/16、3/16、3/16、1/16，總和為 1。觀察到 01 並不表示 p 一定為 1/4，因其他參數也能給這個結果正機率。

例題3 分層資料要對回母體比重

母體兩層各有 80、20 人，各抽 10 人，層內贊成率 6/10 與 3/10。按母體權重估計為 (4/5)(3/5)+(1/5)(3/10)=27/50；若直接合併受訪者，則是 9/20。差別來自兩層抽樣比例不同。

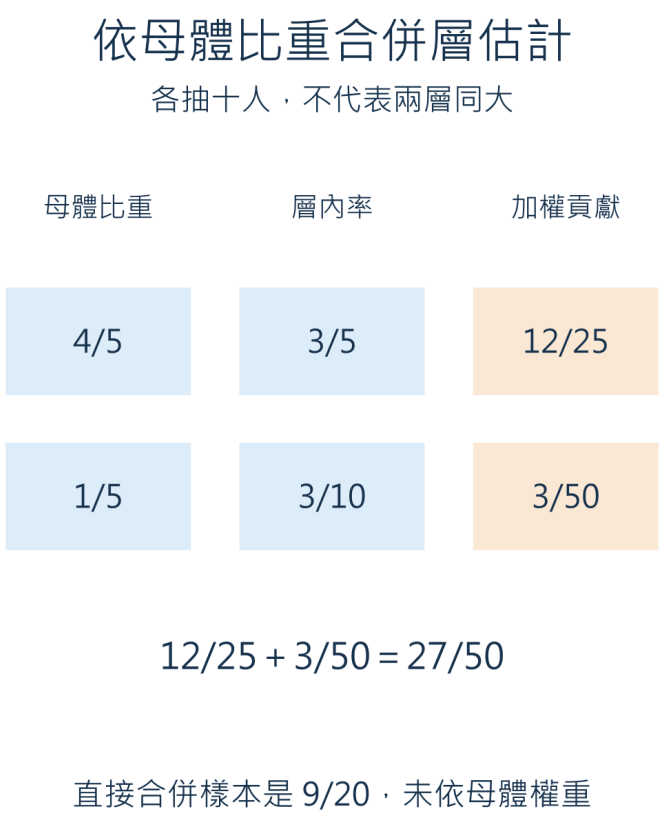


圖4 | 母體兩層權重 4/5、1/5，層內率 3/5、3/10；加權貢獻 12/25 與 3/50，合計 27/50。

例題4 回答者不是整個抽樣名單

抽到 100 人，60 人回覆，30 人回答「是」。答覆者中為 30/60=1/2；未知的 40 人若全否或全是，名單中贊成比例可能從 30/100 到 70/100。要再談母體精確度，還要說明誰被抽到及誰沒有回覆。

2.3 練習

1 **B12-C02-Q01** 想知道全校平均通勤時間，只收一班 30 人的分鐘數。哪一項是目標參數？

2 **B12-C02-Q02** 五名成員中等可能不放回抽二人，特定一人入選率？

3 **B12-C02-Q03** 延續上一題，已知 A 入選，B 入選率？

4 **B12-C02-Q04** 兩個獨立試驗成功率都是 3/10，先失敗再成功的機率？

5 **B12-C02-Q05** 兩個 IID 伯努利試驗 p=1/4，成功數是 0 的機率？

6 **B12-C02-Q06** 母體兩層占 3/4、1/4，其層內率估計為 2/5、4/5。分層合併值？

7 **B12-C02-Q07** 上一題若兩層各抽同樣多人，直接合併樣本時比例是多少？

8 **B12-C02-Q08** 名單 80 人，50 人回覆，20 人贊成。名單內的贊成率可能在哪個範圍？

9 **B12-C02-Q09** 把某人的同一個回答抄成兩行，能不能當成兩筆 IID 觀測？

10 **B12-C02-Q10** 用自願填寫的校網問卷得到 2,000 筆，能只憑筆數聲稱代表全校嗎？

2.4 完整解答

Q01

目標參數是全校學生通勤時間的平均；30 人的平均是可算的統計量。只有一班不等於全校隨機樣本。

Q02

共有 C(5,2)=10 組，指定一人後另一人有 4 種，入選率 4/10=2/5。

Q03

餘下名額只有一個，四人等可能，所以為 1/4；不同於無條件的 2/5。

Q04

獨立使路徑機率相乘，故為 (7/10)(3/10)=21/100。

Q05

一定是兩次都失敗，機率 $(1-1/4)^2=9/16$。

Q06

用母體比重相乘再加總：(3/4)(2/5)+(1/4)(4/5)=3/10+1/5=1/2。

Q07

兩層的樣本比重都是 1/2，直接合併為 (2/5+4/5)/2=3/5，不等於按母體層比重的 1/2。

Q08

未答 30 人。全否時下界 20/80=1/4；全是時上界 (20+30)/80=5/8。這是只用名單資訊得到的界，不是對更廣母體的信賴區間。

Q09

不能。第一行決定第二行，並非獨立重抽。表格筆數雖增加，取得的隨機資訊未增加。

Q10

不能。回覆可能與通勤時間等變量有關；即使同類回覆很多，樣本的選擇機制仍可能偏向特定學生。應先釐清抽樣框及未回覆情況。

第3章 估計量與抽樣分布

3.1 把算出的數字和算數字的規則分開

估計量 T 是從樣本算出數字的規則，估計值 t 是某一次資料代入後的結果。事先不知道會抽到什麼，所以 T 有自己的抽樣分布。這與資料值的分布不同，也不是把已看到的樣本整理成直方圖。要找小樣本的抽樣分布，可以列出所有可能樣本、算出各自 T，再把同一 T 的機率相加。

設 $X_1,\dots,X_n$ 獨立同分布，共同平均 μ、有限變異數 σ^2。把樣本平均記作 $M_n=(X_1+\cdots+X_n)/n$。期望的線性性給 $E[M_n]=\mu$，獨立使交叉協方差為零，故 $\mathrm{Var}(M_n)=\sigma^2/n$。抽樣分布的標準差稱為標準誤；此處為 $\sigma/\sqrt{n}$。若 σ 未知，常把下一段的樣本標準差 s 代入，得到估計標準誤 $s/\sqrt{n}$，不是單筆資料波動。

在 $n\ge 2$ 時，$S^2=\Sigma(X_i-M_n)^2/(n-1)$ 是母體變異數的無偏估計量。展開 $X_i-\mu=(X_i-M_n)+(M_n-\mu)$，交叉項因 $\Sigma(X_i-M_n)=0$ 消失，所以 $\Sigma(X_i-M_n)^2=\Sigma(X_i-\mu)^2-n(M_n-\mu)^2$。取期望得 $n\sigma^2-n(\sigma^2/n)=(n-1)\sigma^2$；這是除以 $n-1$ 的理由。它不表示開根號後的 S 也無偏。

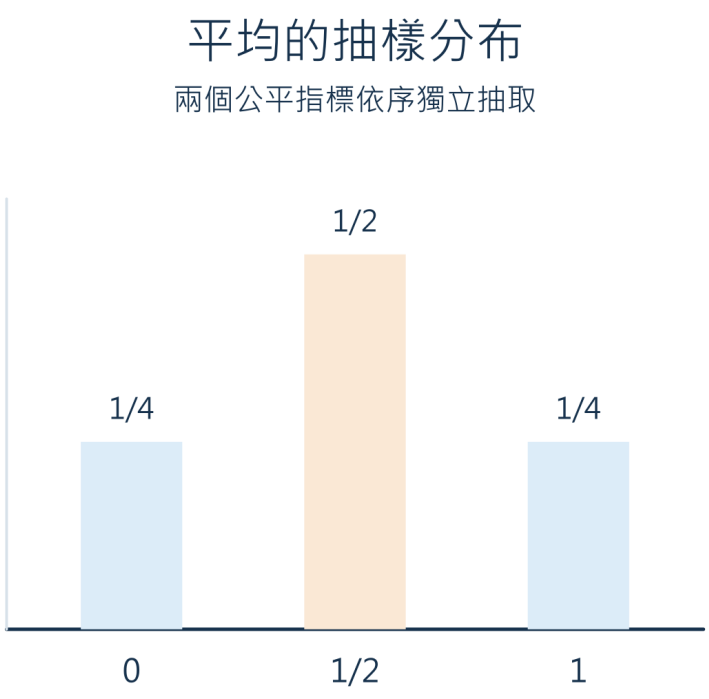


圖5 | 兩次獨立公平指標的平均可為 0、1/2、1，質量各為 1/4、1/2、1/4。

3.2 四個例題

例題1 列出平均的抽樣分布

兩次獨立公平伯努利試驗的四個有序序列各重 1/4。平均 0 來自 00，1/2 來自 01 與 10，1 來自 11。加權平均為 1/2，偏差平方加權為 (1/4)(1/4)+(1/4)(1/4)=1/8，而單個指標的變異數是 1/4。

例題2 同一批樣本可問兩種變異數

三個觀測 2、4、6，平均 4，偏差平方和為 4+0+4=8。描述這三筆時，變異數 8/3；在 IID 模型下估計母體變異數時，估計值 8/(3−1)=4。應先說目的，不把兩個分母混用。

例題3 平均的標準誤不是單次標準差

單次母體標準差為 6。獨立取 9 筆時，平均的標準誤是 6/√9=2；取 36 筆時是 6/√36=1。筆數四倍，平均的標準誤減半，單次波動仍是 6。

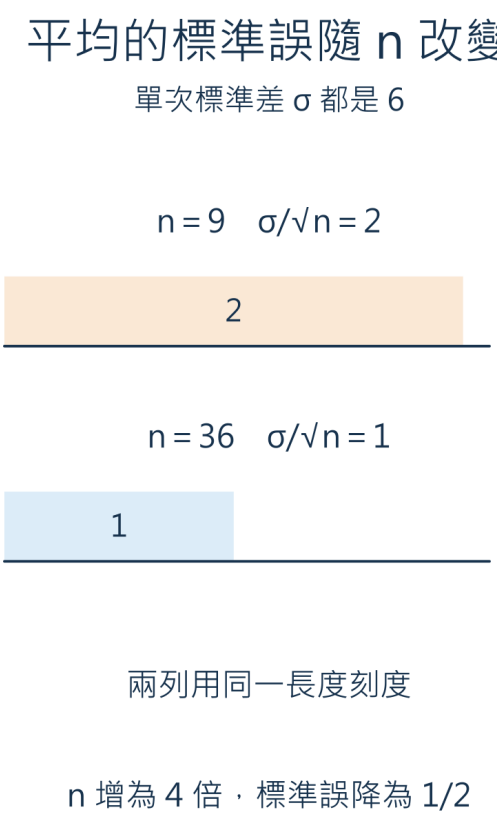


圖6 | 母體標準差 6；獨立抽樣 n=9、36 時，平均的標準誤分別為 2、1。

例題4 用一筆還是用四筆

四筆 IID 觀測，每筆變異數 9。$T=X_1$ 與 M_4 的期望都為 μ，但 T 的變異數為 9；M_4 為 $(4\times9)/4^2=9/4$。這是在同模型下比較兩條規則，並非保證此次 M_4 一定較近。

3.3 練習

1 **B12-C03-Q01** 兩次 IID 伯努利試驗的 $p=1/4$。平均 0、1/2、1 各有多少機率？

2 **B12-C03-Q02** 上一題的平均估計量期望、變異數為何？

3 **B12-C03-Q03** 觀測為 3、7，用 $n-1$ 的樣本變異數為何？

4 **B12-C03-Q04** 只有一筆觀測，能不能照用 S^2 定義？

5 **B12-C03-Q05** IID 每筆變異數 16，樣本量 16，平均的標準誤？

6 **B12-C03-Q06** 9 筆 IID 資料算出的 $s=12$，平均的估計標準誤？

7 **B12-C03-Q07** 為什麼 $E[M_5]=\mu$ 的結論不需要五筆獨立？

8 **B12-C03-Q08** 將一筆資料原樣複製 8 次，平均的抽樣變異數會縮小嗎？

9 **B12-C03-Q09** IID 資料 $\sigma^2=25$，$n=5$；以 n 當分母的偏差平方平均，其期望為何？

10 **B12-C03-Q10** 在同一母體的 IID 抽樣下，若樣本量從 n 增為 $4n$，平均的變異數與標準誤各變為幾倍？

3.4 完整解答

Q01

零次成功為 $(3/4)^2=9/16$；恰一次為 $2(1/4)(3/4)=6/16$；兩次為 $1/16$。這三項和為 1。

Q02

$E[M_2]=p=1/4$。兩次獨立，$Var(M_2)=p(1-p)/2=(1/4)(3/4)/2=3/32$。亦可用上題三個質量重算。

Q03

平均 5，偏差 −2、2，平方和 8。樣本大小 2，分母 1，故 $s^2=8$。

Q04

不能。此時 $n-1=0$。以 n 作分母的描述變異數是 0，但它只描述單筆，不能說母體變異數也為 0。

Q05

平均的變異數為 $16/16=1$，標準誤取非負平方根為 1。

Q06

代入 $s/\sqrt{n}$ 得 $12/3=4$。這用樣本波動估計抽樣波動，不代表母體 σ 已知。

Q07

只要每筆期望皆為 μ，期望的線性性給 $(\mu+\mu+\mu+\mu+\mu)/5=\mu$。獨立用在消除協方差，不用在這一步。

Q08

不會。平均仍是那一筆，分布與原觀測相同。複製出的 8 筆完全相依，不能套 $\sigma^2/8$。

Q09

偏差平方和期望為 $(n-1)\sigma^2=4\times25$。再除以 5 得 20，比真正的 25 小。

Q10

當原波動非零，變異數乘 1/4、標準誤乘 1/2，因方差分母是 n，標準差分母是 $\sqrt{n}$。若 $\sigma=0$，兩者一開始就為零，不談比值。

第4章 偏誤、變異與均方誤差

4.1 用平方損失衡量兩種誤差

偏誤是估計量在重複抽樣下的平均偏離：Bias(T)=E[T]−θ。偏誤可以隨真參數 θ 改變，也可能是常數。單次差值 T−θ 不等於偏誤。變異數量的是 T 繞 E[T] 的波動。小變異數不保證接近目標，例如總猜固定值就沒有抽樣變異。

若 E[T²] 有限，均方誤差 MSE=E[(T−θ)²]。把 T−θ 拆成 (T−E[T])+(E[T]−θ)，平方後有兩個平方和一個交叉項。交叉項的期望是 2(E[T]−θ)E[T−E[T]]=0，故 MSE=Var(T)+Bias(T)²。偏誤有正負，平方損失中的偏誤項總是非負。

比較要在同一目標、同一資料機制和同一損失下進行。小樣本把估計往一個基準值拉，可能減少變異、增加偏誤；是否有利須看參數值。沒有資料之外的理由時，不能從一個 θ 上的優劣推到所有 θ。

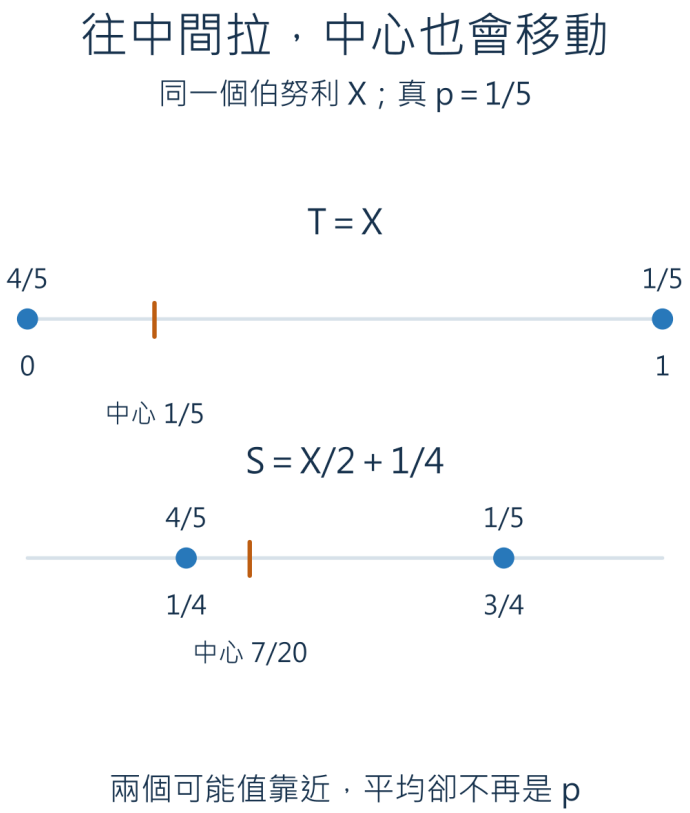


圖7 | 在 p=1/5，T=X 的中心為 1/5；S=X/2+1/4 的中心為 7/20，兩個可能值 1/4、3/4。

4.2 四個例題

例題1 小變異數可以補償小偏誤

只觀測一個伯努利 X，真 p=1/5。T=X 無偏，MSE=(1/5)(4/5)=4/25。S=X/2+1/4 的期望 7/20，偏誤 3/20；變異數是 (4/25)/4=1/25，所以 MSE=1/25+(3/20)²=1/16。對這個 p，S 的平方損失平均較小。

例題2 換一個參數，優劣會換

沿用上例，在 p=0，X 永遠是 0。T 正好等於真值，MSE=0；S 不再波動卻定在 1/4，MSE=(1/4−0)²=1/16。這否定「縮小變異就能處處較好」。

例題3 加上常數不會降低波動

假定 4 筆 IID，單筆變異數 16。M_4 無偏、MSE=16/4=4；T=M_4+2 的變異數一樣是 4，偏誤變成 2，MSE=4+2²=8。平移只改中心。

例題4 反例：較集中的估計反而較差

在 θ=10 的模型，第一法的結果是 11 或 13，各半；平均 12、變異數 1，MSE=1+2²=5。第二法是 8 或 12，各半；平均 10、變異數 4，MSE=4。第一法較集中，卻集中在偏離目標的地方。

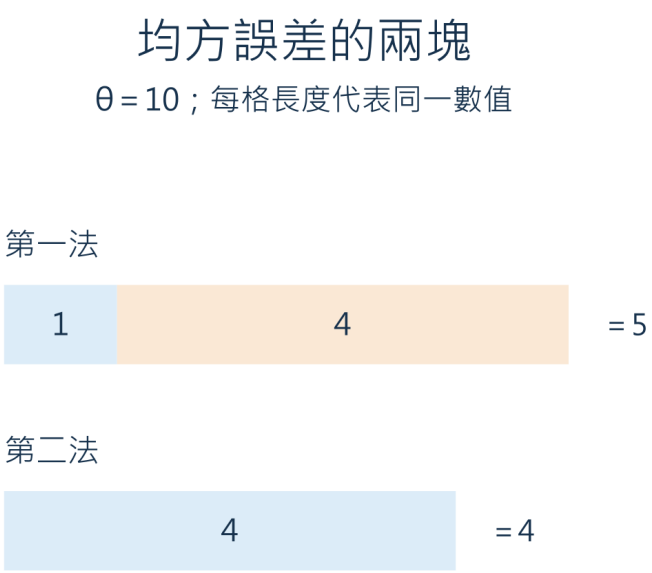


圖8 | θ=10 時第一法變異數 1、偏誤平方 4，MSE=5；第二法變異數 4、偏誤平方 0，MSE=4。

4.3 練習

1 **B12-C04-Q01** E[T]=12、θ=10、Var(T)=9，偏誤與 MSE？

2 **B12-C04-Q02** 偏誤 −3、變異數 4，MSE？

3 **B12-C04-Q03** 若 T−θ 取 −1 或 3，各半，求偏誤、變異數與 MSE。

4 **B12-C04-Q04** 估計規則總是輸出 4，在 θ=6 時的 MSE？

5 **B12-C04-Q05** 25 筆 IID，每筆變異數 100，以樣本平均估計共同平均 μ，MSE 為何？

6 **B12-C04-Q06** 五個公平獨立伯努利指標，用成功比例估 p；MSE？

7 **B12-C04-Q07** 9 筆 IID，單筆變異數 9，用 M_9-1 估 μ，MSE？

8 **B12-C04-Q08** 兩估計量在相同 θ 都無偏，能只用變異數比較此處的平方損失嗎？

9 **B12-C04-Q09** 例題 1 的 p 若是 1/2，兩估計量 MSE 各為何？

10 **B12-C04-Q10** 用分鐘估平均等待時間；MSE 單位是什麼？改成秒會怎樣？

4.4 完整解答

Q01
偏誤是 12−10=2。MSE=9+2²=13，不是加上有符號的 2。

Q02
負偏誤平方為 9，與變異數相加得 13。

Q03
誤差平均 1，所以偏誤 1；兩偏差為 −2、2，變異數 4。誤差平方平均 (1+9)/2=5，亦為 4+1²。

Q04
沒有抽樣波動；每次誤差為 −2。MSE=(4−6)²=4。

Q05
對 μ 無偏，MSE 就是平均的變異數 100/25=4。

Q06
比例無偏，MSE=p(1−p)/5=(1/2)(1/2)/5=1/20。

Q07
平均的變異數 9/9=1，減 1 不改變異數，但偏誤 −1，故 MSE=1+1=2。

Q08
能。由 MSE=Var(T)+Bias(T)²，兩偏誤都為零，同一 θ 下 MSE 正好等於變異數；這不涉及其他損失函數。

Q09
T 無偏，MSE=1/4；S 期望 (1/2)/2+1/4=1/2 也無偏，變異數縮為 (1/4)/4=1/16。

Q10
誤差用分鐘，平方後是平方分鐘。改成秒時每個誤差乘 60，MSE 數值乘 60²=3,600，估計方法本身並未變差。

第5章 最大似然估計

5.1 固定資料，問哪個模型參數最能解釋它

在本章的離散觀測模型中，每個參數 θ 都給觀測資料 x 一個機率。觀察到 x 後固定 x，把 L(θ;x)=Pθ(X=x) 視為 θ 的函數，稱為似然。固定 θ 時，對所有觀測資料 x 的機率求和為 1。似然在參數上不需要求和為 1。選使 L 最大的參數稱為最大似然估計值。這不是參數的後驗機率。

在有限參數集合上，可逐列比較；有並列就保留全部最大值。有些觀測在所有候選參數下都不可能，這要先回頭檢查模型或資料，不能用一個任意並列點解讀。若參數連續，則要先寫參數空間；上確界可能不在空間內，不能把接近最大叫做已達到最大。

有 n≥1 次 IID 伯努利，觀察到的有序序列共 k 次成功。由獨立性，$L(p)=p^k(1-p)^{(n-k)}$。當 0<k<n，兩端的似然為零，在 0<p<1 可對 log L 微分：導數 $k/p-(n-k)/(1-p)$ 的符號與 $k-np$ 相同。它在 $p=k/n$ 左正右負，所以最大值在 k/n。當 k=0 或 n，直接用 $(1-p)^n$ 或 p^n 查端點，不對零取對數。若只觀察成功數，似然乘與 p 無關的 C(n,k)，最大點不變。

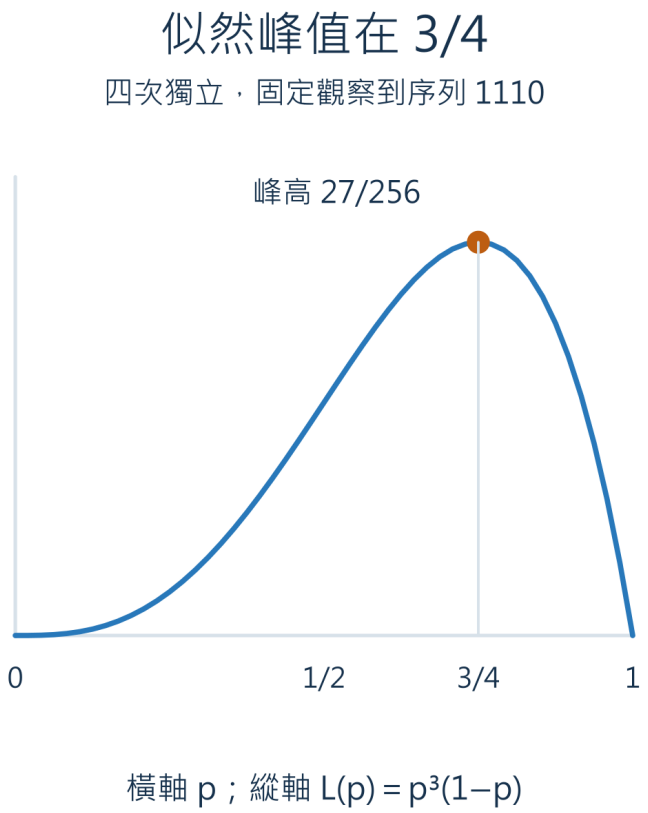


圖9 | 四次 IID 伯努利觀察到三次成功的有序序列，似然 $p^3(1-p)$，峰值在 $p=3/4$。

5.2 四個例題

例題1 四次三成一敗

觀察 1110。$L(p)=p^3(1-p)$，從導數的符號 $3-4p$ 知 L 先升後降。$p=3/4$ 時 $L=(3/4)^3(1/4)=27/256$；$p=1/2$ 時 $L=1/16$，比峰值小。

例題2 全部失敗，要回頭看參數空間

四次都是 0。$L(p)=(1-p)^4$ 隨 p 減少而上升。在 [0,1] 上最大點 $p=0$，$L=1$；在 (0,1) 上任何 p 都可以再減半而提高 L，故不存在 MLE，只能說上確界為 1。

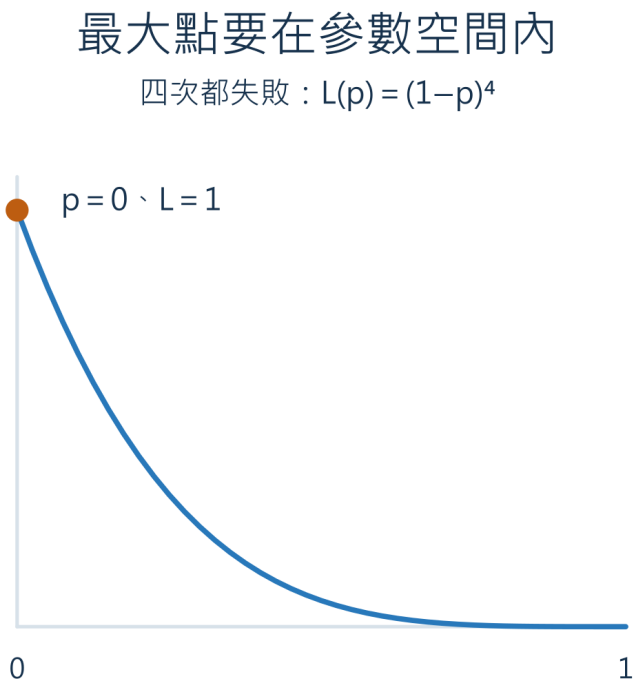


圖10 | 四次都失敗時 L(p)=(1−p)⁴。閉區間容納 p=0 的最大值；開區間只有上確界。

例題3 有限格點可能並列

只允許 p=1/4 或 3/4，觀察 01。兩個似然都為 p(1−p)=3/16，兩者都達最大；不能把 1/2 取作估計，因它不在模型的參數集合。

例題4 似然較大不等於後驗機率較大

候選 p=2/5 或 4/5，獨立兩次皆成功。似然分別為 4/25、16/25，後者大四倍，所以有限模型的 MLE 為 4/5。若要說參數的機率，還需指定另一套對參數的概率模型。

5.3 練習

1 **B12-C05-Q01** 五次 IID 伯努利有兩次成功，p∈[0,1]。MLE 為何？

2 **B12-C05-Q02** 四次 IID 伯努利試驗，p=1/2。固定序列 1110 與「共三次成功」的機率各為何？

3 **B12-C05-Q03** IID 伯努利的有序資料，n=4、k=1。比較 p=1/4 與 1/2 的似然。

4 **B12-C05-Q04** 三次 IID 全成，p∈[0,1] 的 MLE？

5 **B12-C05-Q05** 上題若限制 p∈(0,1)，把 p=0.999 寫成 MLE 合理嗎？

6 **B12-C05-Q06** 格點 p=1/4、1/2、3/4；兩次 IID 伯努利試驗均失敗，哪一格最好？

7 **B12-C05-Q07** 伯努利模型輸入一筆觀測值 2；應報一個 MLE 嗎？

8 **B12-C05-Q08** 把整條似然乘一個與參數無關的正數，最大點集合會變嗎？

9 **B12-C05-Q09** 兩次獨立成功率分別為 p、p^2，見 10，p∈[0,1]。MLE？

10 **B12-C05-Q10** 只有相異標籤的兩個參數，但它們給完全同一觀測分布，能透過最大似然分辨嗎？

5.4 完整解答

Q01

此為內點情形 $0<2<5$。導數符號取決於 $2-5p$，故最大點 $p=2/5$。

Q02

前者是單一序列，機率 $(1/2)^4=1/16$。後者含四種失敗位置，總機率 $4/16=1/4$。

Q03

$p=1/4$ 給 $(1/4)(3/4)^3=27/256$；$1/2$ 給 $1/16=16/256$。前者大，亦恰是連續參數空間的最大點。

Q04

$L(p)=p^3$ 隨 p 增加，且 1 可取，所以最大點是 $p=1$，最大似然 1。

Q05

不合理。任取 $p<1$，$(1+p)/2$ 更大且仍小於 1，給更高似然。可接近 1，不存在最大點。

Q06

三個似然依次為 9/16、1/4、1/16。最大的在 $p=1/4$。

Q07

不應。2 不在該觀測空間；若硬寫每個 p 的似然都給零，也只是所有候選皆不相容。應先查資料編碼或模型。

Q08

不會。正數乘法保序，也保留等號，所以所有大小比較、並列點均相同。若倍率依參數而變則不能如此判定。

Q09

似然 $p(1-p^2)=p-p^3$，導數 $1-3p^2$。在 $1/\sqrt{3}$ 左方正、右方負，兩端似然皆零，所以最大點為 $1/\sqrt{3}$，而非 $1/2$。

Q10

同一資料的機率在兩標籤下相同，似然永遠並列。重複使用同一不可辨識的觀測模型也不能只靠似然把它們區分。

第6章 充分統計量

6.1 哪些彙總在指定模型下不丟失參數資訊

統計量 T 把整筆資料 x 變成較簡短的 t。在有限觀測空間中，若每個可能的 t 都有一組固定的條件分布 P(X=x|T=t)，而且它在所有讓 Pθ(T=t)>0 的參數 θ 下都相同，就說 T 對 θ 充分。條件機率分母為零的參數不拿來比較。充分不表示能重建看見的資料順序，更不表示一個統計量對所有模型都充分。

若機率可以用非負因子寫成 Pθ(X=x)=gθ(T(x))h(x)，其中 h 與 θ 無關，在 T=t 的一組資料內加總得 Pθ(T=t)=gθ(t)Σh(x)。當這個分母為正，條件分布只剩 h(x)/Σh(x)，不再含 θ，所以 T 充分。反過來，若 T 充分，可將每一組的共同條件分布當作 h，取 gθ(t)=Pθ(T=t)。若一組對所有 θ 都不可能，可任取組內分布，乘上零仍給正確機率。

伯努利 IID 序列的似然只由成功總數 K 決定。在相同 k 的每個有序序列，機率均為 $p^k(1-p)^{(n-k)}$；只要該組有正機率，組內就均分為 1/C(n,k)。如果第二次成功率改為 p^2，即使總成功數沒變，成功出現在哪一格也可能提供關於 p 的資訊。

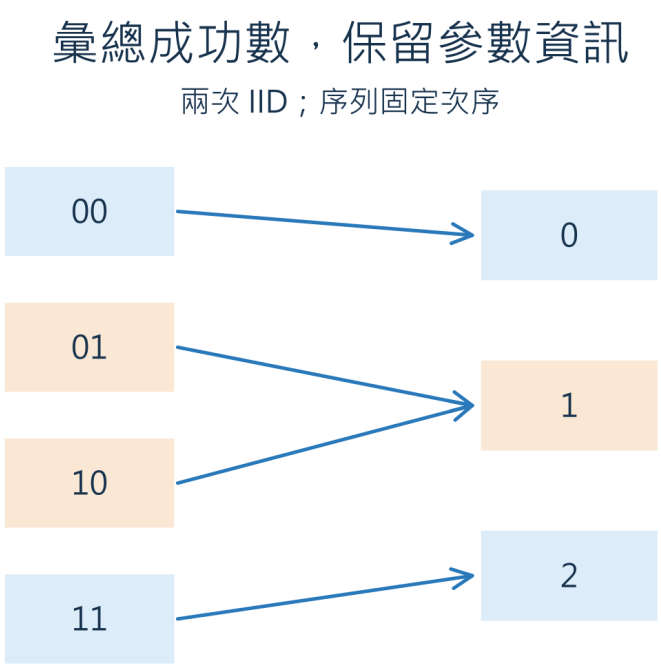


圖11 | 兩個 IID 伯努利試驗由 00、01、10、11 彙總成成功數 0、1、2；已知 K=1 時 01、10 各半。

6.2 四個例題

例題1 知道總數後順序不再透露 p

兩次 IID，給定 K=1。01 和 10 都重 $p(1-p)$，且 K=1 的質量為 $2p(1-p)$。$0<p<1$ 時兩條件率皆 1/2。K=0、2 皆只有一個序列，於是 K 充分。

例題2 只保留第一筆不充分

在相同的兩次 IID 模型，保留 $T=X_1$。給定 T=0 後第二步成功率仍是 p。比較 p=1/4 和 3/4 得不同條件率，否定充分性。

例題3 總數可能漏掉資訊

獨立但兩次的成功率為 p、p^2。當 $0<p<1$ 且 K=1，10 的權重 $p(1-p^2)$，01 的權重 $(1-p)p^2$。把共同因子 $p(1-p)$ 約去，10 的條件率為 $(1+p)/(1+2p)$。它在 p=1/4 是 5/6，p=3/4 是 7/10，故 K 不充分。

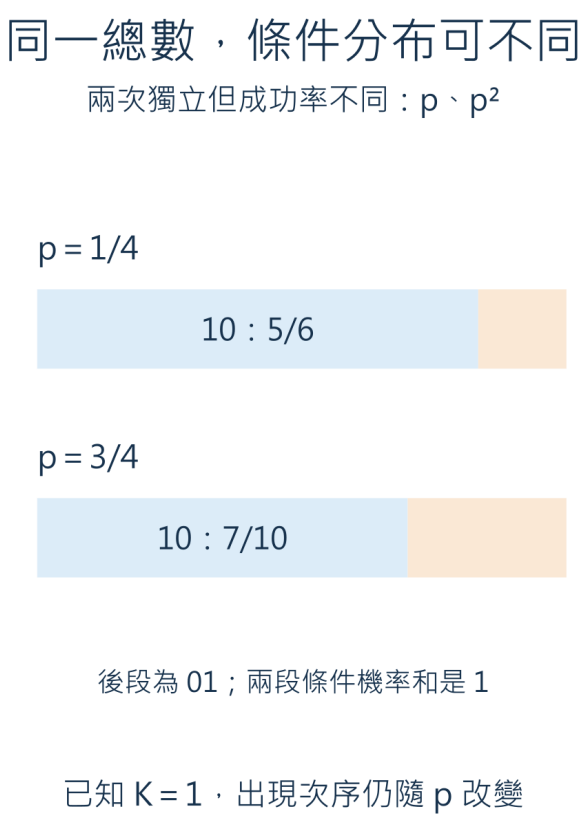


圖12 | 若兩次成功率是 p、p^2，已知 K=1 時 P(10|K=1) 在 p=1/4 為 5/6，在 p=3/4 為 7/10。

例題4 從計數式重建相對權重

四次 IID，給定 K=2 且該觀測有正機率。有 C(4,2)=6 個可能序列；由共同似然因子可知每個占 1/6。保留 K 足以做關於 p 的似然比較，但不能從 K 知道人實際看到哪個序列。

6.3 練習

1 **B12-C06-Q01** 三次 IID，$0<p<1$ 且 $K=1$。指定序列 001 的條件機率？

2 **B12-C06-Q02** 三次 IID，$p<1$。給定 $K=0$，資料為 000 的條件率？

3 **B12-C06-Q03** $p=0$ 時能用 $K=1$ 作條件嗎？

4 **B12-C06-Q04** 四次 IID 公平試驗，給定 $K=2$ 時指定有序序列 1010 的機率？

5 **B12-C06-Q05** 兩次 IID，用 $3K+2$ 代替 K，會喪失充分性嗎？

6 **B12-C06-Q06** 只觀察一次伯努利，用常數 $T=0$ 當統計量，是否對 p 充分？

7 **B12-C06-Q07** 三次 IID 有 $K=2$，參數 3/4 對 1/4 的似然比？

8 **B12-C06-Q08** 兩次 IID 只留 K 的奇偶。用『偶數』組說明它不充分。

9 **B12-C06-Q09** 因子分解中某筆的 $h(x)=0$，而它所在的組有正機率，該筆條件率是多少？

10 **B12-C06-Q10** 兩次 IID 的 $P_p(X=x)$ 可寫成 $g_p(K)h(x)$。取哪個 h，並算 $K=1$ 時的比值？

6.4 完整解答

Q01

三個成功位置同重 $p(1-p)^2$。給定 K=1，就剩三個等可能位置，各 1/3。

Q02

K=0 只能由 000 產生，且 $p<1$ 使該事件正機率，故比值 $(1-p)^3/(1-p)^3=1$。

Q03

不能。此時試驗都失敗，P(K=1)=0。充分性的比較只在這個 t 確實可能的參數上進行。

Q04

單一序列原質量 1/16，而 K=2 總質量 6/16，兩者相除為 1/6。

Q05

不會。知道 3K+2 就能反算 K，再取同一組內的條件分布。這是可逆重標記，不是合併不同的 K。

Q06

不充分。T=0 是必然事件，此條件下 X=1 的機率仍是 p，並未去除參數。

Q07

共同的排列乘數抵消。比值為 $[(3/4)^2(1/4)]/[(1/4)^2(3/4)]=3$；比值不是參數機率。

Q08

偶數組為 00 或 11，後者條件率 $p^2/[(1-p)^2+p^2]$。$p=1/4$ 得 1/10，$p=3/4$ 得 9/10；條件規則仍依 p。

Q09

是 0。分子含 h(x)=0，分母是該組各 h 之和且為正；不是把零除零硬定義為零。

Q10

取所有允許序列的 $h(x)=1$，$g_p(k)=p^k(1-p)^{(2-k)}$，端點可直接以原機率讀取。K=1 且組有正質量時，組內兩個 1 正規化成 1/(1+1)=1/2。

第7章 信賴區間與覆蓋率

7.1 覆蓋率是區間規則的長期表現

資料尚未抽到時，區間 C(X) 的端點會隨樣本變動；參數 θ 固定。覆蓋率 c(θ)=Pθ(θ∈C(X))，計算時在該 θ 的抽樣分布下，把能包住 θ 的結果加總。說規則具有 γ 的信賴水準，是要求在所宣告的模型下每個 θ 的覆蓋率至少為 γ；離散模型的實際值可以高於目標。

例如兩次 IID 伯努利只考慮 p=1/4 或 3/4。成功數 K=0 時報 [1/4,1/4]，K=1 時報 [1/4,3/4]，K=2 時報 [3/4,3/4]。在 p=1/4，只有 K=2 漏蓋；在 p=3/4，只有 K=0 漏蓋。可對這個模型逐一核對，而不需要常態近似。

若資料為 IID 常態，單次標準差 σ>0 已知，$(M_n-\mu)/(\sigma/\sqrt{n})$ 服從標準常態。代數上把 $|M_n-\mu|\le z\sigma/\sqrt{n}$ 改寫成 μ 落在 $[M_n-z\sigma/\sqrt{n},\ M_n+z\sigma/\sqrt{n}]$。取標準常態中央機率 0.95 的臨界值 z≈1.96，可得約 95% 覆蓋。若 σ 未知或模型不符合，代入一個 s 或照抄常態臨界值並不自動保留此結論。抽完看到區間後，不能把頻率覆蓋率改說成參數在這一段裡的機率。



圖13 | 真 p=1/4 時，K=0、1 的區間覆蓋，K=2 不覆蓋；合計 15/16。

7.2 四個例題

例題1 兩個參數都要算

對前述兩點模型，p=1/4 漏蓋機率 $(1/4)^2=1/16$，所以覆蓋 15/16。p=3/4 漏蓋機率 $(1–3/4)^2=1/16$，覆蓋亦 15/16。只算其中一個還不能代表整個參數空間。

例題2 更寬的集合可以較常蓋到

模型只允許 p=1/10 或 9/10，觀察一個伯努利。看到 0 就給左端單點，看到 1 給右端單點，兩個真參數下各有 9/10 覆蓋。總給 [1/10,9/10] 則必覆蓋，但不隨資料縮小。高覆蓋應和區間長度一起看。

例題3 從標準誤取得區間半寬

已知 σ=6、n=9，M_9=20。半寬 1.96×6/√9=3.92，區間為 [16.08,23.92]。若樣本量 36、另一批平均也為 20，半寬 1.96，區間縮為 [18.04,21.96]。模型假設相同才可如此比較。

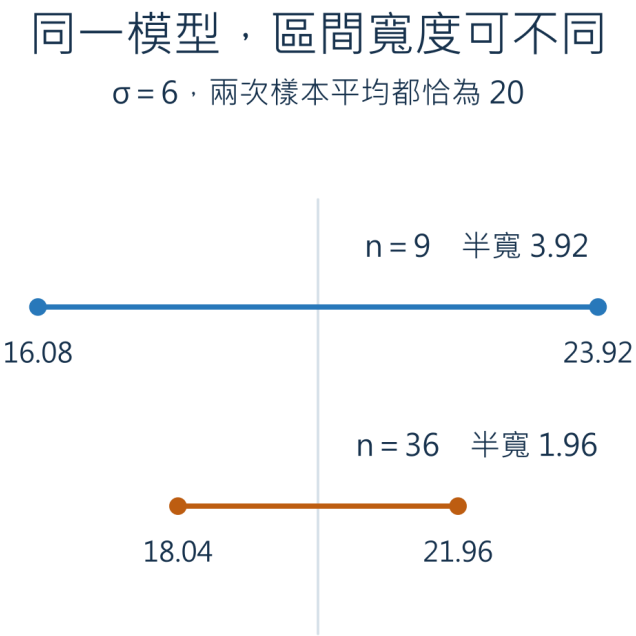


圖14 | 已知 σ=6，中心同取 20；n=9 的半寬是 3.92，n=36 的半寬是 1.96，比較半寬時固定臨界值。

例題4 倍數不同，覆蓋率也不同

在已知 σ 的常態規則下，把半寬從 1 個標準誤改為 2 個，涵蓋事件分別是 |Z|≤1 和 |Z|≤2。標準常態機率約 0.6827、0.9545；用較寬的區間換到較多的覆蓋。

7.3 練習

1 **B12-C07-Q01** 前述兩點規則觀測 K=0 時，報什麼區間？

2 **B12-C07-Q02** 在 p=1/4 時，前述規則漏蓋機率？

3 **B12-C07-Q03** 在 p=3/4 時，前述規則覆蓋率？

4 **B12-C07-Q04** 常態模型 σ=4 已知、n=16，取 z=1.96。區間長度？

5 **B12-C07-Q05** 上一題若觀測到平均 8，兩端是什麼？

6 **B12-C07-Q06** 固定 σ>0 與臨界值、把 n 增四倍，長度比例是多少？

7 **B12-C07-Q07** 報出某個 95% 信賴區間，能說固定 μ 有 95% 機率在它裡面嗎？

8 **B12-C07-Q08** 把一個常態觀測複製成 16 筆，用 σ/4 算平均的標準誤合理嗎？

9 **B12-C07-Q09** 相同資料與已知 σ>0，把中央信賴水準從 90% 提高到 95%，區間如何變？

10 **B12-C07-Q10** 模擬五次，有四次區間蓋到已知真值，這等於規則的覆蓋率嗎？

7.4 完整解答

Q01

依事先訂好的映射，報 [1/4,1/4]。這是 C(0)，不證明真 p 必為 1/4；p=3/4 時也有機會看到 0。

Q02

只有 K=2 漏掉 1/4，須兩次都成功，機率 $(1/4)^2=1/16$。

Q03

只有 K=0 漏掉 3/4，為 $(1/4)^2=1/16$；覆蓋是其補集 $1-1/16=15/16$。

Q04

標準誤 $4/\sqrt{16}=1$。1.96 是半寬，總長度為 $2\times1.96=3.92$。

Q05

用 8 分別減、加 1.96，得 [6.04,9.96]。兩端對稱，合計長 3.92。

Q06

長度為 $2z\sigma/\sqrt{n}$；新長度分母變成 $2\sqrt{n}$，故在 $z>0$ 時為原來的一半。區間中心仍要看新資料。

Q07

不能用這個頻率解釋說成單次參數的機率。95% 說的是在假設成立時，固定 μ，重抽樣重做區間的覆蓋比例。

Q08

不合理。16 筆完全相依，平均仍等於那個觀測，標準誤是 σ。過窄的半寬不保證原覆蓋。

Q09

中央區間要包更多標準常態機率，正的臨界值 z 會上升。樣本平均和標準誤不變，兩側較遠，區間較長。

Q10

本次觀測的命中比例 4/5=0.8 是模擬估計。覆蓋率是模型下所有可能資料的機率，少量模擬的命中數本身仍會波動。

第8章 檢定與證據強弱

8.1 事先定義什麼結果算反對虛無模型

檢定先提出虛無假設 H_0，再說明對哪些替代 H_1 敏感。例如五次 IID 伯努利試驗，H_0:p=1/2，H_1:p>1/2，成功數 K 愈大愈反對 H_0。拒絕區域須先選；若 H_0 真，錯誤拒絕是第一類錯誤。顯著水準 α 是對這項錯誤機率的上限，離散模型的實際錯誤率可能比 α 小。

若 H_1 中真參數為 p，未拒絕的機率 β(p) 是該點的第二類錯誤機率，檢定力為 1−β(p)。它取決於真 p、樣本量和拒絕區域。把 α 壓低而資料量不變，往往也更難偵測某些偏離。一次未拒絕不表示兩種情況相等。

對固定樣本計畫與單一虛無模型，一側 p 值是在 H_0 下，統計量達到這次觀測或更偏向 H_1 的機率。本例就是 $P_0(K\geq k)$。雙側需先說何謂兩邊都極端，不能所有檢定都直接把一側乘二。p 值不等於 H_0 為真的機率，也不告訴你差異有多大。看過資料才改方向、改停止規則或多次挑最小值，原來的錯誤控制未必適用。



圖15 | H_0:p=1/2、n=5，六柱質量為 1、5、10、10、5、1 除以 32；α=0.05 時右端 K=5 才拒絕。

8.2 四個例題

例題1 把 α 和真正的第一類錯誤分開

五次 IID，H_0:p=1/2。若只在五次全成才拒絕，第一類錯誤為 $(1/2)^5=1/32$，即 0.03125，不超過 α=0.05。因 K 的機率是離散的，這個不隨機化規則不必剛好用完 0.05。

例題2 觀測四次成功的右尾 p 值

H_0 下 K=4 有 5 個序列，K=5 有 1 個，各序列重 1/32。故 p 值為 (5+1)/32=3/16=0.1875，高於 0.05；本次不拒絕。只取 K=4 的機率會漏掉更極端結果。

例題3 若 p=3/4，規則抓得到幾次

仍只在 K=5 拒絕。當 p=3/4，檢定力為 $(3/4)^5=243/1024$，第二類錯誤機率為 781/1024。即使真的偏高，小樣本多數時候仍未拒絕。



圖16 | α=0.05 的右尾規則 K=5 才拒絕；H_0:p=1/2 與 p=3/4 下拒絕率為 1/32、243/1024。

例題4 雙側先指定極端的排序

四次 IID 與 H_0:p=1/2，預先以 |K−2| 判斷偏離。觀測 K=4 時，至少同樣極端的是 K=0 或 4，各 1/16，故雙側 p 值 1/8。若另換不對稱的 H_0 或排序，不能沿用這個加倍。

8.3 練習

1 **B12-C08-Q01** 五次 IID，H_0:p=1/2，觀測 K=3，右尾 p 值？

2 **B12-C08-Q02** 同一檢定觀測 K=5，α=0.05 時該拒絕嗎？

3 **B12-C08-Q03** 若把拒絕區域改成 K≥4，還是 0.05 水準嗎？

4 **B12-C08-Q04** 五次 IID 伯努利試驗，採 K=5 才拒絕的規則，真 p=3/5 時的第二類錯誤機率？

5 **B12-C08-Q05** 四次 IID，全失敗；H_0:p=1/2，事先訂左尾。p 值？

6 **B12-C08-Q06** 沿用五次 IID 伯努利試驗且只在 K=5 時拒絕的規則。H_0:p=1/2 為真時，未拒絕機率？

7 **B12-C08-Q07** p 值 0.03 能說效應大小是 3% 嗎？

8 **B12-C08-Q08** 比較「P_0(結果至少這麼極端)」與「給定結果後 H_0 為真」。哪一項是 p 值？

9 **B12-C08-Q09** 兩個獨立檢定的虛無皆真，實際第一類錯誤率各 0.05。至少一個錯誤拒絕的機率？

10 **B12-C08-Q10** 例題 3 未拒絕時，可寫「已證明 p=1/2」嗎？

8.4 完整解答

Q01

加總 K=3、4、5 的三項，得 (10+5+1)/32=1/2。

Q02

右尾只剩 K=5，p 值 1/32=0.03125≤0.05，照事先規則拒絕；這不是保證任何替代為真。

Q03

不是。H_0 的拒絕機率 3/16 高於 0.05，實際第一類錯誤超過指定上限。

Q04

拒絕須五次全成功，機率 $(3/5)^5=243/3125$。其補集是 $1-243/3125=2882/3125$。

Q05

觀測 K=0，左尾 K≤0 只含零，機率 $(1/2)^4=1/16$。若事先訂的是右尾，答案會不同。

Q06

是 $1-1/32=31/32$。H_0 真時未拒絕不叫第二類錯誤；第二類錯誤只在替代為真時定義。

Q07

不能。0.03 是虛無模型下的尾機率，大小受樣本量與波動影響；效應大小要從估計值及其單位來報。

Q08

前者是 p 值，機率方向從假設指向資料。後者另需對假設與資料的機率模型，不能把條件方向直接交換。

Q09

兩個都不拒絕的機率 $(19/20)^2$；取補集得 39/400=0.0975。不因每個單獨是 0.05 而使整組仍是 0.05。

Q10

不行。在 p=3/4 時同一規則也有 781/1024 不拒絕。未拒絕只表示資料未達到指定拒絕門檻。

第9章 迴歸與模型檢查

9.1 用垂直殘差的平方和選擇直線

觀測為 (x_i,y_i)。用直線預測 y 為 $a+bx$，定義殘差 $e_i=y_i-(a+bx_i)$。最小平方選 a、b 讓殘差平方和 $SSE=\Sigma e_i^2$ 最小。給定 b，讓截距取平均 y 減 b 倍平均 x：這時殘差和為零；把截距再平移 c 只會增加 nc^2。因此先將兩欄中心化，就剩下選 b 的一維問題。

以 m_1、m_2 分別表示 x 欄和 y 欄的平均，令 $Sxx=\Sigma(x_i-m_1)^2$、$Sxy=\Sigma(x_i-m_1)(y_i-m_2)$、$Syy=\Sigma(y_i-m_2)^2$。代入截距後 $SSE=Syy-2bSxy+b^2Sxx$。當 $Sxx>0$，把這個二次式配方，平方項為 $Sxx(b-Sxy/Sxx)^2$，故唯一最小斜率為 Sxy/Sxx。若所有 x 相同，$Sxx=0$，資料只決定該 x 高度，不能分辨截距與斜率。

有截距時，若 $Syy>0$，可用 $R^2=1-SSE/Syy$ 比較直線對樣本變動的描述。它不檢查殘差曲線、資料是否獨立或因果方向。應回畫殘差對 x 圖，看是否有弧形、異常點或波動隨 x 改變；模型外推和泛化另看新資料。計算最小平方本身不要求常態誤差；要據此做不確定度推論，就要補模型假設。

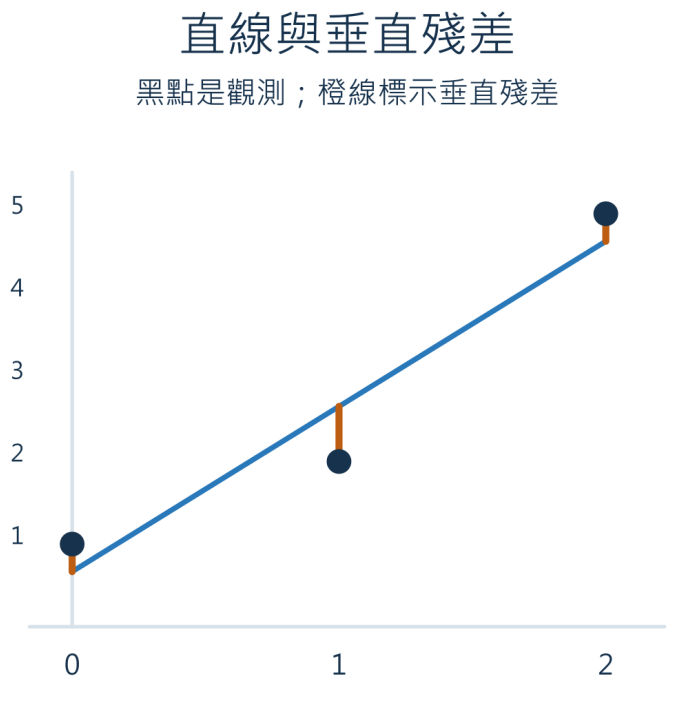


圖17 | 點 (0,1)、(1,2)、(2,5) 擬合直線 y=2x+2/3，三條垂直殘差為 1/3、−2/3、1/3。

9.2 四個例題

例題1 中心化後求斜率與截距

三點 (0,1)、(1,2)、(2,5)，平均 x=1、平均 y=8/3。Sxx=2、Sxy=(−1)(−5/3)+(1)(7/3)=4，故 b=2、a=8/3−2=2/3。殘差依序 1/3、−2/3、1/3，平方和 (1+4+1)/9=2/3。

例題2 所有 x 相同時斜率不可辨

三筆 x 都為 2，y 分別是 1、3、5。所有直線只透過 a+2b 在資料上作用，最佳值是 3。所有滿足 a+2b=3 的係數組合都給相同 SSE=(1−3)²+(5−3)²=8，不能挑一條宣稱獨特。

例題3 零斜率也會漏掉明顯關係

取 x=−2、−1、0、1、2，y=x²。平均 y=2，中心化乘積依正負相抵，Sxy=0，預測直線恆為 2。殘差依 x 為 2、−1、−2、−1、2，呈 U 形，顯示不是純粹沒有關係，而是直線未捕捉曲線。

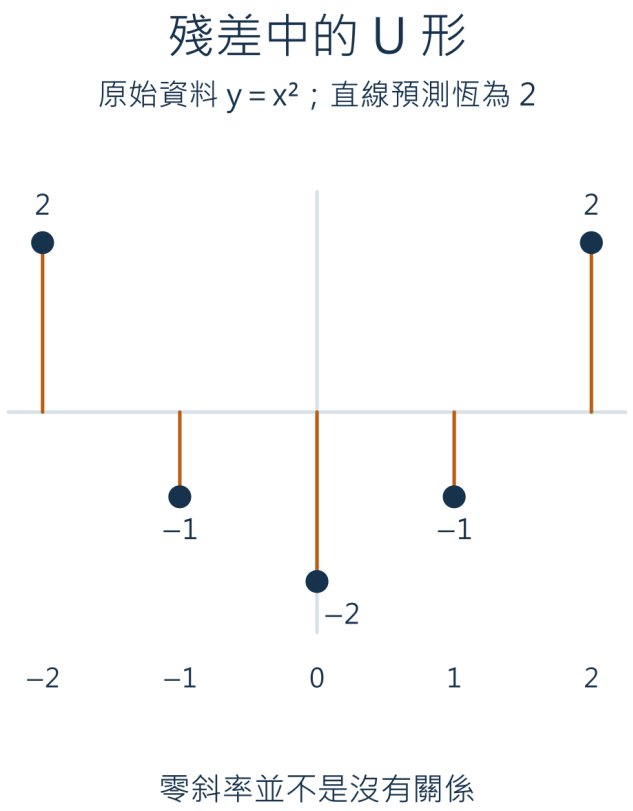


圖18 | 五點 y=x² 在 x=−2、−1、0、1、2 的直線殘差為 2、−1、−2、−1、2，呈 U 形。

例題4 高 R² 仍要看適用範圍

例題 1 的 Syy=(25+4+49)/9=26/3，R²=1−(2/3)/(26/3)=12/13。這僅說直線相對只報平均，能降低這三筆的平方誤差；不因分數高就能外推或判定改變 x 導致 y 改變。

9.3 練習

1 **B12-C09-Q01** 資料 (0,3)、(1,5)、(2,7) 的最小平方直線？

2 **B12-C09-Q02** 用上一題直線預測 x=3 的 y，可不可以保證準？

3 **B12-C09-Q03** 資料 (0,1)、(1,4)、(2,4)，求斜率、截距及 SSE。

4 **B12-C09-Q04** 為什麼有截距的最小平方線的殘差和為零？

5 **B12-C09-Q05** 兩筆 x 都為 1，y 為 0、2。預測值在這個 x 應為多少？斜率唯一嗎？

6 **B12-C09-Q06** 所有 y 都相同，還能用 $R^2=1-SSE/Syy$ 報數字嗎？

7 **B12-C09-Q07** 例題 3 改用另一對對稱點 $x=-3$、3 及 $x=0$，$y=x^2$。斜率與截距？

8 **B12-C09-Q08** 若 x 是件數、y 是分鐘，斜率 b=2 的單位？

9 **B12-C09-Q09** 在某線性擬合，x 用公尺時斜率 5。改用公分但 y 單位不變，新斜率？

10 **B12-C09-Q10** 練習時間與成績的斜率為正，能說延長練習一定提高成績嗎？

9.4 完整解答

Q01

平均 x=1、平均 y=5，Sxx=2、Sxy=4，得到 b=2、a=3。三筆都在線上，所以 SSE=0。

Q02

代入得到 3+2×3=9。但 x=3 超出觀測範圍 [0,2]，9 是模型外推，不是觀測事實。

Q03

中心 (1,3)，Sxx=2、Sxy=3，b=3/2、a=3/2。三個預測 1.5、3、4.5，殘差 −0.5、1、−0.5，SSE=1/4+1+1/4=3/2。

Q04

對任意最優斜率 b，截距取平均 y−b×平均 x。把 $\Sigma[y_i-a-bx_i]$ 展開為 n(平均 y−a−b×平均 x)，恰為零。

Q05

最佳共同預測是平均 y=1，所以 a+b=1。無論 b 取何值，只要截距配合，兩筆的預測不變。斜率不唯一。

Q06

這裡 Syy=0，公式未定義。截距可以完美擬合常數，但「解釋相對變動的比例」沒有非零分母；不要由軟體慣例推成定理。

Q07

平均 x=0，Sxy 由對稱相消，b=0；平均 y=(9+0+9)/3=6，所以 a=6。零斜率仍不等於 y 與 x 無關。

Q08

單位是分鐘 / 件：在這條直線上，x 增一件，預測 y 增二分鐘。不是說每次都實際多兩分鐘。

Q09

同一距離的新數值 x′=100x。代入 x=x′/100，原斜率 5 變為 5/100=0.05，預測不改變。

Q10

不能由這張觀察圖推出因果或必然。基礎程度、課程安排等可能同時影響兩欄；還要檢查收集設計與替代解釋。

第10章 完整資料問題與判讀

10.1 把估計值連同取得方式和限制一起說

一份可讀的資料報告先寫明：想回答誰的什麼問題、在何時與什麼尺度上。接著交代誰被納入、誰沒有回覆、量測怎麼做；必要時列出分母、缺失數或合理的上下界。只要抽樣機制改變，結論能推到的對象就可能改變，不要把圖表上的樣本比例當成母體真值。

先定分析規則再看結果：想估差多少，就報估計值、單位及不確定度；想檢定某個明確模型，就先寫假設、方向、拒絕規則。已知 σ 的常態區間不足以處理所有資料，數字很精密也無法補救不合適的模型。用另一種合乎問題的分組、畫殘差或改變缺失假設，檢查結論是否敏感。預測要在未拿來調模型的資料上檢查。

最後把結果句與限制句接在一起。讀者應知道看到的是描述、關聯、預測還是因果證據；也應知道哪一段是資料直接支持、哪一段來自模型。下面四個情境都用自編資料，目的不是找一個最好看的數字，而是報出可以重算的判讀。



圖19 | 從問題、收集分母、估計、檢查假設到陳述限制；缺失假設不穩時，回去補資料或縮小結論。

10.2 四個例題

例題1 小型問卷：讓缺失留在分母裡

想知道受邀的 100 人是否贊成延長圖書館時段。60 人回覆、45 人贊成，回覆者中為 75%。40 人答案未知，名單內可能從 45% 到 85%。可報「回覆者 45/60 贊成；全名單比例只能由現有資訊界在 45% 到 85%」，不把此界稱作信賴區間。

例題2 比較兩個隨機分派組，先報差多少

自編實驗把 40 名參與者隨機分成兩組，沒有失訪。A 的完成率 12/20、B 為 8/20，樣本差 0.20，即 20 個百分點。隨機分派有助於避免基礎差異系統性集中在某一組。即使固定這 40 人，換一個分派仍可能得到不同的組間差。要推到更廣的母體，還要說明這 40 人如何納入。未計不確定度前不寫成「任何人都會提高 20%」。

例題3 用留出資料比較預測

兩種方法都未看三個留出真值 4、7、9。方法一預測 5、7、8，絕對誤差 1、0、1，平均 2/3。固定報 6 的基準誤差 2、1、3，平均 2。先在同一批、同一尺度下比，方法一這次較小；三筆檢查不足以保證所有未來樣本。

例題4 各層更高，合併卻更低

A 在易題完成 9/10、難題 20/40；B 在易題 32/40、難題 4/10。兩層中 A 都高 0.10，但 A 合併 29/50，B 為 36/50。若要在相同難度構成下比較，可各半加權，得 A 0.70、B 0.60。原加總和等重回答不同問題；分層本身不能取代隨機分派。



圖20 | A 在易組、難組皆比 B 高 10 個百分點，但兩組人數配置不同，原樣合併為 A 29/50、B 36/50。

10.3 練習

1 **B12-C10-Q01** 受邀 200 人，150 人回覆，90 人贊成。報回覆者比例及名單內上下界。

2 **B12-C10-Q02** 想估每次通勤平均時間，但表格混用分鐘與小時。先做什麼？

3 **B12-C10-Q03** 隨機分派與從全校隨機抽樣是同一件事嗎？

4 **B12-C10-Q04** 完成率 A 12/20、B 8/20，分別報絕對差及相對 B 的增加。

5 **B12-C10-Q05** 基準組完成率為 0 時，能照上一題算相對增加嗎？

6 **B12-C10-Q06** 留出真值 2、5，預測 3、2。平均絕對誤差？

7 **B12-C10-Q07** 用留出集挑出誤差最小的模型，再用同組誤差宣稱未參與選模。對嗎？

8 **B12-C10-Q08** 例題 4 用共同權重 3/4 易、1/4 難，A 與 B 的加權率？

9 **B12-C10-Q09** 缺失假設得到的上下界，等於 95% 信賴區間嗎？

10 **B12-C10-Q10** 把「75% 的所有受邀者贊成，已證明應延時」改寫成受例題 1 支持的句子。

10.4 完整解答

Q01

回覆者比例 90/150=0.6；其餘 50 人全否或全是時，名單內為 90/200=0.45 到 140/200=0.70。這些界不依賴未回覆者與回覆者分布相同。

Q02

先統一單位，確認每筆出發、抵達與缺失規則，例如把 0.5 小時改記 30 分鐘。未統一前平均值沒有一致量尺。

Q03

不同。分派決定已納入的人進哪個實驗組；抽樣決定哪些人進研究，關係到能推論哪個母體。一個研究可以只有其中一種。

Q04

A 為 0.60、B 為 0.40。差 0.20 是 20 個百分點；相對增加為 0.20/0.40=0.5，即 50%。兩個說法的分母不同。

Q05

不能用該比值，因分母是 0。可報兩組原始分子分母及絕對差；不要自行把未定義的量寫成有限百分比。

Q06

兩個絕對誤差分別為 |3−2|=1、|2−5|=3。等重平均為 (1+3)/2=2，與原值同單位。

Q07

不對。雖未用留出集更新係數，但它已影響選模。要評估選模後的預測表現，應保留另一批未參與選擇的資料。

Q08

A 為 (3/4)(9/10)+(1/4)(1/2)=4/5；B 為 (3/4)(4/5)+(1/4)(2/5)=7/10。共同權重消去兩邊難度比例不同的影響，但不保證是因果差異。

Q09

不等於。前者對未觀測值的可能取法給一個範圍，沒有用重複抽樣來分配 95%。後者必須說明樣本機制及隨機區間規則。

Q10

可寫「100 名受邀者中有 60 名作答，其中 45 名贊成（75% 的作答者）；尚有 40 名未答，名單內贊成率僅能界在 45% 至 85%。」是否延時還要考量需求與成本，資料沒有單獨證成政策。