勘誤表
資料分析的統計學基礎–應用 R 語言 (第 1 版)

陳旭昇

2024.12.29

摘 要

感謝王星宇來信告知錯誤與建議。

頁數	錯誤	修正
71	再透過相對頻率來找出機率值。這種給定機率值的觀點	再透過相對頻率來衡量機率值。這種詮釋機率值的觀點
79	並有比 Thomas Bayes 更為深入了討論與詮釋。	並有比 Thomas Bayes 更為深入的討論與詮釋。
97	上述的 Bernoilli 隨機變數與二項隨機變數的砥柱集合均為有限	上述的 Bernoulli 隨機變數與二項隨機變數的砥柱集合均為有限
160	$E(X) = \int_0^1 x dx = \left[\frac{x}{2}\right]_{x=0}^1 = \frac{1}{2}$	$E(X) = \int_0^1 x dx = \left[\frac{x^2}{2}\right]_{x=0}^1 = \frac{1}{2}$
235	$P(X = 2) = \sum_y P(X = 2, Y = y) = P(X = 2, Y = 2) + P(X = 2, Y = 4) + P(X = 2, Y = 6) = 0.12 + 0.24 + 0.24 = 0.6$	$P(X = 3) = \sum_y P(X = 3, Y = y) = P(X = 3, Y = 2) + P(X = 3, Y = 4) + P(X = 3, Y = 6) = 0.12 + 0.24 + 0.24 = 0.6$
266	本章探討當隨機樣本的樣本點變大時	本章探討當隨機樣本的樣本點增加且趨近無窮大時
311	因為一般而言, 100% 的涵蓋機率會建構出一個包含整個母體參數空間的區間, 這樣的區間沒有統計應用與決策上的價值。舉例來說, 如果我們有興趣的參數是支持率的母體比率 p, 則區間 $[0, 1]$ 會有 100% 的涵蓋機率, 但是這樣的區間毫無意義。	因為一般而言, 100% 的涵蓋機率會建構出一個包含整個母體參數空間的區間, 舉例來說, 如果我們有興趣的參數是支持率的母體比率 p, 則區間 $[0, 1]$ 會有 100% 的涵蓋機率, 但是這樣包牌式的區間沒有統計應用與決策上的價值。

定義 14.3 (分配收斂). 給定隨機變數之序列 $\{Y_n\}$, 其對應之分配函數為 $F_n(y)$, 另有一隨機變數 Y, 其對應之分配函數為 $F_Y(y)$, 若

$$\lim_{n\to\infty} F_n(y) = F_Y(y)$$

則我們稱 Y_n 分配收斂 *(converge in distribution)* 至隨機變數 Y, 並寫做

$$Y_n \xrightarrow{d} Y$$

簡而言之, 分配收斂意指隨著 n 變大, Y_n 的分配函數與 Y 的分配函數越來越靠近。因此, 我們可以在 n 變大時, 以 $F_Y(y)$ 來近似 Y_n 的分配。換句話說, $F_Y(y)$ 乃是 Y_n 在 n 變大時的近似分配 (approximate distribution), 或稱極限分配 (limiting distribution)。

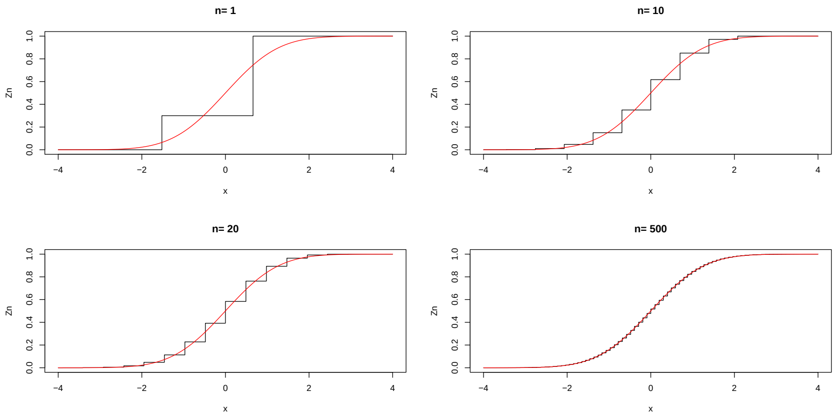
以下的 R 程式 14.2 以隨機變數 $X_n \sim \text{Binomial}(n,0.7)$ 為例, 介紹何謂近似分配。我們在圖 14.2 中畫出 $Y_n = \frac{X_n - n\times 0.7}{\sqrt{n\times 0.7(1-0.7)}}$ 的分配函數 $F_n(y)$, 並讓 $n = 1, 10, 20$, 以及 100。由於 Y_n 是離散隨機變數, 因此分配函數為階梯函數。我們在每一張圖中加入標準常態隨機變數 $N(0,1)$ 的分配函數, 注意到當 n 變大, 階梯函數就越來越平滑, 此時 $N(0,1)$ 提供 $F_n(y)$ 的良好近似。

R 程式 14.2 (分配收斂).

```
p = 0.7
par(mfrow = c(2,2))
for (n in c(1,10,20,500)) {
x = seq(-4,4,by=0.01)
z=(x*sqrt(p*(1-p)/n)+p)*n
Fn =  pbinom(z, n, p)
plot(x, Fn, main = paste("n=", n), ylab="Zn",  type="s")
curve(pnorm(x,0,1), add=TRUE, col="red")
}
```

注意到, CMT 在分配收斂上也成立, 亦即,

圖 14.2: 分配收斂



定理 14.2 (Continuous Mapping Theorem II). 給定 $X_n \xrightarrow{d} X$, 且 $g(\cdot)$ 為連續函數, 則

$$g(X_n) \xrightarrow{d} g(X)$$

以下定理連結動差生成函數與分配收斂。

定理 14.3. 給定 $X_1, X_2, \ldots, X_n$ 為隨機變數序列, X 為隨機變數, 且其 *MGF*, $M_{X_n}(t)$ 與 $M_X(t)$ 均存在, 若 $M_{X_n}(t) \longrightarrow M_X(t)$, 則

$$X_n \xrightarrow{d} X$$

亦即, 只要我們能證明 X_n 的 MGF 隨著樣本數 n 變大, 會趨近 X 的 MGF, 則 X_n 分配收斂 X。

14.3 什麼是弱大數法則

一個機率收斂最好的例子就是弱大數法則 (weak law of large numbers, WLLN)。弱大數法則的原型 (prototype) 來自 Jakob Bernoulli (1654-1705) 的 Bernoulli 法則。Jakob 死後 8 年, 於 1713 年, 在他的姪兒 Nicholas Bernoulli (1687-1759) 替他出版 Ars Conjectandi (The Art of Conjecturing) 一書中, Jakob 證明獨立且重複地觀測一發生機率為 p 之事件,