詳細解答

STEP I 2004，第 13 題

概率與統計 • 均勻分佈變量的最大值

STEP I 2004，第 13 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/04-s1-q13

關鍵思路

若干個獨立（independent）的數的最大值不超過 x，當且僅當其中每一個數都不超過 x。這樣一來，棘手的「最大值」就化為若干個相同概率的乘積，一行便可寫出 X 的分佈函數（distribution function）。其餘一切都由此而來：求導可得密度（density）和期望值（expectation），而令分佈函數等於 0.05，便可求出檢驗的臨界值（critical value）。

解答

第 (i) 部分

設 U_1, U_2, U_3 為該三個數，它們互相獨立，且各自在 $[0,1]$ 上服從均勻分佈（uniform distribution），所以當 $0 \le x \le 1$ 時，$\mathrm{P}(U_i \le x) = x$。由於 $X = \max(U_1, U_2, U_3)$，

$$X \le x \iff U_1 \le x,\ U_2 \le x \text{ 及 } U_3 \le x.$$

由獨立性，

$$F(x) = \mathrm{P}(X \le x) = x^3 \qquad (0 \le x \le 1).$$

特別地，$\mathrm{P}(X \le 0.8) = 0.8^3 = 0.512$，正是所求。

求導後，X 在 $[0,1]$ 上的密度為 $f(x) = F'(x) = 3x^2$，所以

$$\mathrm{E}(X) = \int_0^1 x \cdot 3x^2 \,\mathrm{d}x = \left[\tfrac{3}{4}x^4\right]_0^1 = \frac{3}{4}.$$

這個結果合情合理：三個數中最大的一個，理應遠高於單獨一個數的平均值 $\frac{1}{2}$。

第 (ii) 部分

現在有 N 個獨立的數，各自在 $[0,a]$ 上服從均勻分佈，所以當 $0 \le x \le a$ 時，$\mathrm{P}(U_i \le x) = x/a$，而同樣的論證給出

$$\mathrm{P}(X \le x) = \left(\frac{x}{a}\right)^N \qquad (0 \le x \le a),$$

而當 $x \ge a$ 時，$\mathrm{P}(X \le x) = 1$。由於 X 是連續的，$\mathrm{P}(X < x)$ 與此相同。

臨界值。 當 $x < c$ 時，檢驗拒絕 H_0；而 5% 的顯著性水平（significance level）表示，在 H_0 為真（$a = 1$）時拒絕它的概率為 0.05：

$$\mathrm{P}(X < c \mid a = 1) = c^N = 0.05, \qquad \text{所以} \qquad c = 0.05^{1/N}.$$

x 的值偏小，是支持 $a < 1$ 的證據，這就是拒絕域（rejection region）為 $x < c$ 的原因。

? **想一想** 「$x \le 0.8$ 必然導致拒絕」怎樣化為一個關於 N 的不等式？

每一個 $x \le 0.8$ 都被拒絕，當且僅當每一個這樣的 x 都小於 c，即 $0.8 < c$。由於 $t \mapsto t^N$ 在 $t > 0$ 時是遞增的，這等價於

$$0.8^N < c^N = 0.05.$$

所以我們要找滿足 $0.8^N < 0.05$ 的最小整數 N。（留意當 N 增大時，c 會向 1 遞增，所以 N 愈大，檢驗就愈傾向於拒絕。）

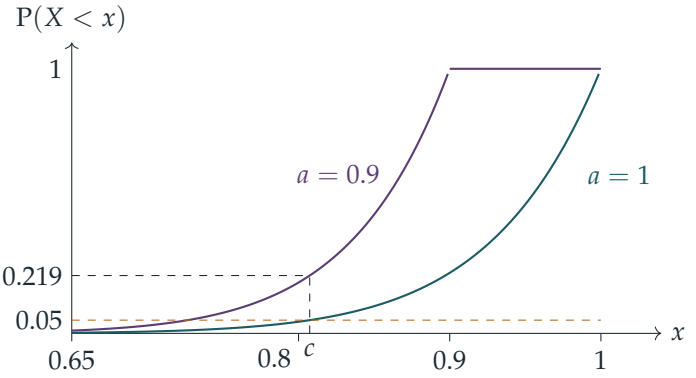
利用 $2^{10} \approx 10^3$。 寫成 $0.8 = 2^3/10$，於是 $0.8^N = 2^{3N}/10^N$。

- $N = 13$： $0.8^{13} = \dfrac{2^{39}}{10^{13}} = \dfrac{2^9 \cdot 2^{30}}{10^{13}} \approx \dfrac{512 \times 10^9}{10^{13}} = 0.0512 > 0.05$。
- $N = 14$： $0.8^{14} = \dfrac{2^{42}}{10^{14}} = \dfrac{2^2 \cdot 2^{40}}{10^{14}} \approx \dfrac{4 \times 10^{12}}{10^{14}} = 0.04 < 0.05$。

由於 0.8^N 隨 N 增大而遞減，每一個 $N \le 13$ 同樣不符合要求。所以最小的值是

$$N = 14.$$

（這個近似會略為低估 2 的冪，因為 $2^{10} = 1024$；真實數值 $0.8^{13} \approx 0.0550$ 和 $0.8^{14} \approx 0.0440$ 會得出相同的結論。）



當 $N = 14$ 時，臨界值為 $c = 0.05^{1/14} \approx 0.807$，僅僅高於 0.8。

若 $a = 0.8$。 所有的數都落在 $[0, 0.8]$ 內，所以必然有 $X \leq 0.8 < c$。零假設（null hypothesis）被拒絕的概率為 1。

若 $a = 0.9$。 此時 $c \approx 0.807 < 0.9$，所以

$$\mathrm{P}(\text{拒絕}) = \mathrm{P}(X < c \mid a = 0.9) = \left(\frac{c}{0.9}\right)^{14} = \frac{c^{14}}{0.9^{14}} = \frac{0.05}{0.9^{14}}.$$

這裏 $0.9^2 = 0.81$，$0.9^4 = 0.6561$，$0.9^8 \approx 0.4305$，因此 $0.9^{14} = 0.9^8 \cdot 0.9^4 \cdot 0.9^2 \approx 0.2288$。於是

$$\mathrm{P}(\text{拒絕} \mid a = 0.9) = \frac{0.05}{0.9^{14}} \approx 0.219.$$

所以即使有 14 個觀測值，這個檢驗通常也察覺不到 a 是 0.9 而不是 1。

驗算

- 蒙地卡羅模擬（Monte Carlo simulation）（2×10^5 次試驗）在第 (i) 部分得出 $\mathrm{P}(X \leq 0.8) \approx 0.512$ 和 $\mathrm{E}(X) \approx 0.750$。
- 一般而言，$[0, 1]$ 上 n 個均勻分佈的數的最大值，其平均值為 $\frac{n}{n+1}$；當 $n = 1$ 時這是 $\frac{1}{2}$，理應如此，而當 $n = 3$ 時則是 $\frac{3}{4}$。
- 取 $N = 14$ 和 $c = 0.05^{1/14}$，模擬得出的拒絕率在 $a = 1$ 時約為 0.0505（即檢驗的大小），在 $a = 0.9$ 時約為 0.218，與 $0.05/0.9^{14} = 0.2186$ 吻合。

常見錯誤

- 以 $1 - 0.8^N$ 代替 0.8^N：最大值偏小的概率是乘積 0.8^N，而不是它的補數。

- 把 $a = 0.9$ 時的拒絕概率當作 $(0.8/0.9)^{14}$。檢驗是在 $x < c$ 時拒絕，而 c 並不是 0.8。
- 因為 0.0512「大約是 0.05」而停在 $N = 13$。我們需要的是嚴格小於 0.05。

STEP 之外

最大值是一個順序統計量（*order statistic*），而對於 $[0, a]$ 上的均勻分佈，它是 a 的充分統計量（*sufficient statistic*）：它包含了樣本中關於 a 的全部資訊。把拒絕概率看成真實 a 的函數，就是該檢驗的檢定力函數（*power function*）。在這裏，當 $a \leq 1$ 時它是 $\min\{1, 0.05/a^N\}$，而統計學家正是透過比較檢定力函數來在不同的檢驗之間作出選擇。Grimmett 與 Welsh 的 *Probability: An Introduction* 涵蓋了最大值和最小值的分佈；Blitzstein 與 Hwang 的 *Introduction to Probability* 則詳細討論了順序統計量。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。