详细解答

STEP I 2004，第 13 题

概率与统计 • 均匀分布变量的最大值

STEP I 2004，第 13 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

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

关键思路

若干个独立（independent）的数的最大值不超过 x，当且仅当其中每一个数都不超过 x。这样一来，棘手的“最大值”就化为若干个相同概率的乘积，一行便可写出 X 的分布函数（distribution function）。其余一切都由此而来：求导可得密度（density）和期望值（expectation），而令分布函数等于 0.05，便可求出检验的临界值（critical value）。

解答

第 (i) 部分

设 U_1, U_2, U_3 为该三个数，它们互相独立，且各自在 $[0,1]$ 上服从均匀分布（uniform distribution），所以当 $0 \le x \le 1$ 时，$\mathrm{P}(U_i \le x) = x$。由于 $X = \max(U_1, U_2, U_3)$，

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

由独立性，

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

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

求导后，X 在 $[0,1]$ 上的密度为 $f(x) = F'(x) = 3x^2$，所以

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

这个结果合情合理：三个数中最大的一个，理应远高于单独一个数的平均值 $\frac{1}{2}$。

第 (ii) 部分

现在有 N 个独立的数，各自在 $[0,a]$ 上服从均匀分布，所以当 $0 \le x \le a$ 时，$\mathrm{P}(U_i \le x) = x/a$，而同样的论证给出

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

而当 $x \ge a$ 时，$\mathrm{P}(X \le x) = 1$。由于 X 是连续的，$\mathrm{P}(X < x)$ 与此相同。

临界值。当 $x < c$ 时，检验拒绝 H_0；而 5% 的显著性水平（significance level）表示，在 H_0 为真（$a = 1$）时拒绝它的概率为 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 $a < 1$ 的证据，这就是拒绝域（rejection region）为 $x < c$ 的原因。

? **想一想** “$x \le 0.8$ 必然导致拒绝”怎样化为一个关于 N 的不等式?

每一个 $x \le 0.8$ 都被拒绝，当且仅当每一个这样的 x 都小于 c，即 $0.8 < c$。由于 $t \mapsto t^N$ 在 $t > 0$ 时是递增的，这等价于

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

所以我们要找满足 $0.8^N < 0.05$ 的最小整数 N。（留意当 N 增大时，c 会向 1 递增，所以 N 愈大，检验就愈倾向于拒绝。）

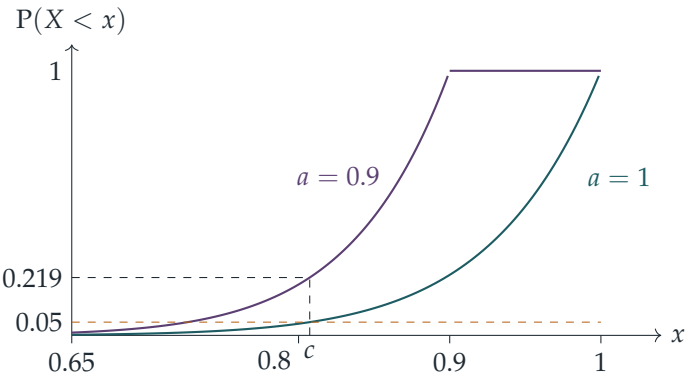
利用 $2^{10} \approx 10^3$。写成 $0.8 = 2^3/10$，于是 $0.8^N = 2^{3N}/10^N$。

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

由于 0.8^N 随 N 增大而递减，每一个 $N \le 13$ 同样不符合要求。所以最小的值是

$$N = 14.$$

（这个近似会略为低估 2 的幂，因为 $2^{10} = 1024$；真实数值 $0.8^{13} \approx 0.0550$ 和 $0.8^{14} \approx 0.0440$ 会得出相同的结论。）



当 $N = 14$ 时，临界值为 $c = 0.05^{1/14} \approx 0.807$，仅仅高于 0.8。

若 $a = 0.8$。 所有的数都落在 $[0, 0.8]$ 内,所以必然有 $X \leq 0.8 < c$。零假设(null hypothesis)被拒绝的概率为 1。

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

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

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

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

所以即使有 14 个观测值，这个检验通常也察觉不到 a 是 0.9 而不是 1。

验算

- 蒙地卡罗模拟（Monte Carlo simulation）（2×10^5 次试验）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 个均匀分布的数的最大值，其平均值为 $\frac{n}{n+1}$；当 $n = 1$ 时这是 $\frac{1}{2}$，理应如此，而当 $n = 3$ 时则是 $\frac{3}{4}$。
- 取 $N = 14$ 和 $c = 0.05^{1/14}$，模拟得出的拒绝率在 $a = 1$ 时约为 0.0505（即检验的大小），在 $a = 0.9$ 时约为 0.218，与 $0.05/0.9^{14} = 0.2186$ 吻合。

常见错误

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

- 把 $a = 0.9$ 时的拒绝概率当作 $(0.8/0.9)^{14}$。检验是在 $x < c$ 时拒绝，而 c 并不是 0.8。
- 因为 0.0512“大约是 0.05”而停在 $N = 13$。我们需要的是严格小于 0.05。

STEP 之外

最大值是一个*顺序统计量*（*order statistic*），而对于 $[0, a]$ 上的均匀分布，它是 a 的*充分统计量*（*sufficient statistic*）：它包含了样本中关于 a 的全部信息。把拒绝概率看成真实 a 的函数，就是该检验的*检定力函数*（*power function*）。在这里，当 $a \leq 1$ 时它是 $\min\{1, 0.05/a^N\}$，而统计学家正是通过比较检定力函数来在不同的检验之间作出选择。Grimmett 与 Welsh 的 *Probability: An Introduction* 涵盖了最大值和最小值的分布；Blitzstein 与 Hwang 的 *Introduction to Probability* 则详细讨论了顺序统计量。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。