详细解答

STEP II 2024，第 12 题

概率与统计 • 离散顺序统计量

STEP II 2024，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2024 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

关键思路

在第一个游戏中，胜出者是持有唯一最小数字的玩家，而这样的玩家最多只有一个。因此“有胜出者”可以按照谁胜出、以哪个数字胜出，分拆成互斥（disjoint）的情况，而每一种情况都只是若干独立（independent）概率的简单乘积。题目给出 $\sum i^2$ 和 $\sum i^3$ 的公式，正是提示所得的和是多项式求和。第二个游戏的做法相同，只是现在要考虑一对有序的玩家（一个最低、一个最高），并要对两个数字求和。

解答

在整个解答中，k 个数字互相独立，且每个都在 $\{1,2,\ldots,n\}$ 上均匀分布（uniform），因此任何一位玩家得到任何一个特定数字的概率都是 $\frac{1}{n}$，而得到某个含 s 个值的集合中的数字的概率是 $\frac{s}{n}$。

第 (i) 部分

固定 a，其中 $1 \le a \le n-1$。Ada 必须得到 a（概率为 $\frac{1}{n}$），而其余 $k-1$ 位玩家每人都必须得到 $a+1,\ldots,n$ 这 $n-a$ 个数字之一（每人的概率为 $\frac{n-a}{n}$）。由独立性，

$$\mathrm{P}(\text{Ada 得到}a\text{, 其他人全部较大}) = \frac{1}{n}\left(\frac{n-a}{n}\right)^{k-1} = \frac{(n-a)^{k-1}}{n^k}.$$

（当 $a = n$ 时此概率为零，这就是 a 只取到 $n-1$ 的原因。）

设 W 为有胜出者的事件。胜出者所持的数字严格小于其他所有人的数字，所以不可能有两个胜出者：否则两者都必须比对方小。因此 W 是以下事件的互斥并集：对 k 位玩家及 $a = 1,\ldots,n-1$，事件“这位玩家得到 a，而其他人都得到较大的数字”。由对称性

（symmetry），每位玩家的贡献都与 Ada 相同，所以

$$\mathrm{P}(W) = k\sum_{a=1}^{n-1}\frac{(n-a)^{k-1}}{n^k} = \frac{k}{n^k}\sum_{j=1}^{n-1} j^{k-1},$$

其中我们代入了 $j = n-a$，它同样由 1 取到 $n-1$。

当 $k=4$ 时，利用题目给出的 $\sum i^3$ 公式，并以 $n-1$ 代替 n，

$$\mathrm{P}(W) = \frac{4}{n^4}\cdot\frac{1}{4}(n-1)^2n^2 = \frac{(n-1)^2}{n^2},$$

即为所求。

第 (ii) 部分

固定 d 和 a，其中 $1 \le d \le n-2$ 且 $1 \le a \le n-d-1$；这些条件保证 $a+d+1 \le n$，并且 a 与 $a+d+1$ 之间至少严格夹着一个数字。Ada 必须得到 a，Bob 必须得到 $a+d+1$（各自的概率为 $\frac{1}{n}$），而其余 $k-2$ 位玩家每人都必须得到 $a+1,\ldots,a+d$ 这 d 个数字之一。由独立性，

$$\mathrm{P}(\text{Ada 得到}a,\ \text{Bob 得到}a+d+1,\ \text{其他人严格介乎两者之间}) = \frac{1}{n^2}\left(\frac{d}{n}\right)^{k-2} = \frac{d^{k-2}}{n^k}.$$

> **想一想** 为甚么在对玩家以及对 a 和 d 求和时，每一种出现两个胜出者的方式都恰好出现一次?

有两个胜出者，即是一位玩家持有唯一最小的数字，而另一位玩家持有唯一最大的数字。与第 (i) 部分一样，低位胜出者是唯一的，高位胜出者亦然，因此结果决定了一对有序的不同玩家（低位胜出者，高位胜出者），连同较小的数字 a 以及到较大数字的差距 $d \ge 1$：差距必须至少为 1，因为其余 $k-2 \ge 2$ 位玩家的数字都严格介乎两者之间。这些选择不同，所得的事件就互斥。有序玩家对共有 $k(k-1)$ 对，每一对的贡献都与（Ada，Bob）相同，所以

$$\mathrm{P}(\text{两个胜出者}) = k(k-1)\sum_{d=1}^{n-2}\sum_{a=1}^{n-d-1}\frac{d^{k-2}}{n^k} = \frac{k(k-1)}{n^k}\sum_{d=1}^{n-2}(n-1-d)\,d^{k-2},$$

因为内层的和有 $n-d-1$ 个相等的项。

现在取 $k=4$，并记 $m = n-1$，于是 d 由 1 取到 $m-1$。利用题目给出的两条公式，并以 $m-1$ 代替 n，

$$\begin{aligned}\sum_{d=1}^{m-1}(m-d)\,d^2 &= m\cdot\tfrac{1}{6}(m-1)m(2m-1) - \tfrac{1}{4}(m-1)^2m^2\\ &= \tfrac{1}{12}m^2(m-1)[2(2m-1)-3(m-1)]\\ &= \tfrac{1}{12}m^2(m-1)(m+1).\end{aligned}$$

因此，由于 $m-1=n-2$ 及 $m+1=n$，

$$\mathrm{P}(\text{两个胜出者})=\frac{12}{n^4}\cdot\frac{1}{12}(n-1)^2(n-2)n=\frac{(n-2)(n-1)^2}{n^3},$$

这正是所求的结果。

当 $k=4$ 时比较一个与两个胜出者

设 L 为有玩家持有唯一最小数字的事件，H 为有玩家持有唯一最大数字的事件。第 (i) 部分给出 $\mathrm{P}(L)=\frac{(n-1)^2}{n^2}$。把每个数字 x 换成 $n+1-x$，会把均匀分布的数字变成均匀分布的数字，并把“最小”与“最大”互换，所以同样有 $\mathrm{P}(H)=\mathrm{P}(L)$。

恰好两个胜出者就是事件 $L\cap H$，因此记 $T=\mathrm{P}(L\cap H)=\frac{(n-2)(n-1)^2}{n^3}$。恰好一个胜出者，即是 L 或 H 发生但不同时发生，所以

$$\mathrm{P}(\text{恰好一个})=\mathrm{P}(L)+\mathrm{P}(H)-2\,\mathrm{P}(L\cap H)=\frac{2(n-1)^2}{n^2}-2T.$$

两个胜出者比一个胜出者更有可能，当且仅当

$$\begin{aligned}T>\frac{2(n-1)^2}{n^2}-2T &\iff 3\cdot\frac{(n-2)(n-1)^2}{n^3}>\frac{2(n-1)^2}{n^2}\\ &\iff 3(n-2)>2n \iff n>6,\end{aligned}$$

其中我们除以正数 $\frac{(n-1)^2}{n^3}$（因为 $n\geq 3$，所以它是正数）。当 $n=6$ 时两个概率相等，因此最小值是

$$n=7.$$

验算

- 当 $n=3$ 及 $k=4$ 时共有 81 个结果。以 1 胜出需要其余三个数字都在 $\{2,3\}$ 中：共 $4\cdot 2^3=32$ 个结果。以 2 胜出需要其余三个都是 3：共 4 个结果。所以 $\mathrm{P}(W)=\frac{36}{81}=\frac{4}{9}=\frac{(3-1)^2}{3^2}$。
- 当 $n=6$ 时，“恰好一个”和“恰好两个”的概率都是 $\frac{25}{54}\approx 0.463$；当 $n=7$ 时，两者分别是 $\frac{144}{343}\approx 0.420$ 和 $\frac{180}{343}\approx 0.525$。这些数值，以及两部分中在 $n\leq 8$ 和 $k=2,3,4$ 时的公式，都已经用电脑逐一枚举（enumeration）所有结果加以核实。
- 当 $n\to\infty$ 时，数字相同的情况变得罕见，所以有四位玩家时，几乎总会同时有唯一最低和唯一最高的数字：的确 $\frac{(n-2)(n-1)^2}{n^3}\to 1$。

常见错误

- 在第 (i) 部分漏掉因子 k，或在第 (ii) 部分漏掉因子 $k(k-1)$。在第 (ii) 部分中，两位胜出者是有序的，因为一位是低位、一位是高位。
- 把“恰好一个胜出者”算成 $\mathrm{P}(L)+\mathrm{P}(H)-\mathrm{P}(L\cap H)$，但这其实是“至少一个”；重叠部分必须减去两次。
- 答 $n=6$：此时两个概率相等，并非严格的大小关系。

STEP 之外

把一个事件按“谁”和“在哪里”分拆成互斥的部分，是求顺序统计量（order statistics）——即样本中最小值和最大值——的分布的离散版本。对于连续样本，同样的论证会给出最小值与最大值的联合密度（joint density），而令本题变得有趣的数字相同情况亦会消失。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率导论》）是认识这些概念的好起点。

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。