詳細解答

STEP III 2023，第 12 題

概率與統計 • 組合概率

STEP III 2023，第 12 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2023 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

www.ocr.org.uk/Images/707102-2023-past-papers.zip

關鍵思路

雖然同一對襪子的兩隻看起來一樣，但抽出的是從 $2n$ 隻*實際存在*的襪子中均勻隨機（uniformly random）選取的 $2k$ 隻所組成的集合，所以我們為每隻襪子編上標號，然後數集合的個數。這樣一來，一切都歸結為分階段構造一個選法：先選出哪些對是完整出現的，再選出哪些其他的對只貢獻一隻襪子，最後選出每一個這樣的對中取哪一隻。第 (iii) 部分用的是熟悉的技巧 $r\binom{n}{r} = n\binom{n-1}{r-1}$，它把一個加權和轉化為一個概率（probability）之和。

解答

為襪子編上標號，使全部 $2n$ 隻互不相同。共有 $\binom{2n}{2k}$ 種機會均等的選法，而下面每個概率都等於（有利選法的數目）$/\binom{2n}{2k}$。

第 (i) 部分

一個沒有完整一對的選法，從每一對中最多只含一隻襪子，所以它的 $2k$ 隻襪子來自 $2k$ 個不同的對。要構造這樣的選法，先選出有代表出現的 $2k$ 對，有 $\binom{n}{2k}$ 種方法（因為 $2k \leq n$，這是可行的），然後從每一對的兩隻中選一隻，有 2^{2k} 種方法。每一種這樣的選擇都給出不同的集合，而每個沒有成對的集合都以這種方式產生，所以

$$\mathrm{P}(\text{沒有成對}) = \frac{\binom{n}{2k}\,2^{2k}}{\binom{2n}{2k}}.$$

第 (ii) 部分

固定 r，其中 $0 \le r \le k$。一個恰好有 r 對完整襪子的選法由以下部分組成：

- 整對取出的 r 對（共 $2r$ 隻襪子），以及
- 另外 $2k - 2r = 2(k-r)$ 隻襪子，全部來自不同的對（否則成對的數目會多於 r），而且都不來自已用的 r 對。

因此，我們以 $\binom{n}{r}$ 種方法選出 r 個完整的對；從餘下的 $n-r$ 對中，以 $\binom{n-r}{2(k-r)}$ 種方法選出貢獻一隻襪子的 $2(k-r)$ 對；再從這些對中各取一隻，有 $2^{2(k-r)}$ 種方法。同樣地，每個恰好有 r 對的集合都恰好被產生一次，所以

$$\mathrm{P}(X_{n,k} = r) = \frac{\binom{n}{r}\binom{n-r}{2(k-r)}\,2^{2(k-r)}}{\binom{2n}{2k}},$$

此即所求。（當 $r = 0$ 時，這就是第 (i) 部分。）

第 (iii) 部分

設 $1 \le r \le k$。由於

$$r\binom{n}{r} = \frac{r \cdot n!}{r!\,(n-r)!} = n \cdot \frac{(n-1)!}{(r-1)!\,(n-r)!} = n\binom{n-1}{r-1},$$

由第 (ii) 部分得

$$r\,\mathrm{P}(X_{n,k} = r) = \frac{n\binom{n-1}{r-1}\binom{n-r}{2(k-r)}\,2^{2(k-r)}}{\binom{2n}{2k}}.$$

現在把第 (ii) 部分應用於 $n-1$ 對襪子、抽出 $2(k-1)$ 隻、找到 $r-1$ 對的情況。這是容許的，因為由 $2k \le n$ 可推出 $2(k-1) \le n-1$，而且 $0 \le r-1 \le k-1$。由於 $(n-1)-(r-1) = n-r$ 及 $(k-1)-(r-1) = k-r$，

$$\mathrm{P}(X_{n-1,k-1} = r-1) = \frac{\binom{n-1}{r-1}\binom{n-r}{2(k-r)}\,2^{2(k-r)}}{\binom{2n-2}{2k-2}}.$$

兩者的分子除了因子 n 之外完全相同，所以

$$\begin{aligned}\frac{r\,\mathrm{P}(X_{n,k} = r)}{\mathrm{P}(X_{n-1,k-1} = r-1)} &= \frac{n\binom{2n-2}{2k-2}}{\binom{2n}{2k}} = n \cdot \frac{(2n-2)!}{(2k-2)!\,(2n-2k)!} \cdot \frac{(2k)!\,(2n-2k)!}{(2n)!} \\ &= n \cdot \frac{2k(2k-1)}{2n(2n-1)} = \frac{k(2k-1)}{2n-1}.\end{aligned}$$

（上面的除法是安全的：$n - r \ge 2k - r \ge 2(k-r)$，所以這裏每個二項式係數（binomial coefficient）都是正數。）兩邊相乘即得所求結果。

想一想 這個恆等式對求 $\mathrm{E}(X_{n,k})$ 有甚麼幫助？

期望值（expectation）正好是左邊各項之和，而右邊各項加起來是一個完整的概率分佈（probability distribution）。期望值中 $r=0$ 的一項為零，所以

$$\mathrm{E}(X_{n,k}) = \sum_{r=1}^{k} r\,\mathrm{P}(X_{n,k}=r) = \frac{k(2k-1)}{2n-1}\sum_{r=1}^{k}\mathrm{P}(X_{n-1,k-1}=r-1)$$
$$= \frac{k(2k-1)}{2n-1}\sum_{s=0}^{k-1}\mathrm{P}(X_{n-1,k-1}=s).$$

變量 $X_{n-1,k-1}$ 數的是 $2(k-1)$ 隻襪子中的對數，所以它只取值 $0,1,\ldots,k-1$，因此最後的和等於 1。故此

$$\mathrm{E}(X_{n,k}) = \frac{k(2k-1)}{2n-1}.$$

驗算

- 當 $k=1$（兩隻襪子）時，答案是 $\frac{1}{2n-1}$：無論先抽到哪一隻，餘下的 $2n-1$ 隻中恰好有一隻是它的配對。
- 用指示變量（indicator variables）可以直接求出平均值：第 i 對完整出現的概率是 $\binom{2n-2}{2k-2}/\binom{2n}{2k} = \frac{2k(2k-1)}{2n(2n-1)}$，乘以 n 對，同樣得到 $\frac{k(2k-1)}{2n-1}$。
- 當 $n=6$、$k=3$ 時：第 (i) 部分給出 $\frac{64}{924} = \frac{16}{231}$，而平均值是 $\frac{15}{11}$。以電腦對所有 $n \le 8$ 枚舉全部選法，證實了 (ii) 中的分佈、(iii) 中的恆等式以及平均值。

常見錯誤

- 把同一對的兩隻襪子當作真正相同，然後去數「顏色組合」：這些組合並非機會均等，所以必須對有標號的襪子計數。
- 忘記了因子 $2^{2(k-r)}$，它對應每個單隻出現的對中取哪一隻。
- 在 (iii) 中沒有留意到 $\sum_s \mathrm{P}(X_{n-1,k-1}=s)=1$，而嘗試硬算這個和。

STEP 之外

第二項驗算用的是期望值的線性性（linearity of expectation）配合指示變量，它無需知道分佈就能求出平均值；即使「第 i 對完整出現」這些事件並不獨立，這方法依然成立。(iii) 中的恆等式是一個尺寸偏倚（size-biasing）關係：以 r 為分佈加權，會得到一個較小問題的平移版本；這個技巧同樣可以求出二項分佈（binomial distribution）和超幾何

分佈（hypergeometric distribution）的平均值。Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）全書都採用這類指示變量和故事式證明。

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。