12.4 两两独立，却非相互独立

工作纸

题目

STEP III 2021，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 2021 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696588-2021-past-papers.zip

提示

请先自行尝试解题。如果遇到困难，请逐一阅读提示，每读完一个便回到题目再试。标有停一停，试一试的提示是关键步骤。每个提示之后均留有空位供你演算。

想一想 $\mathrm{P}(X_{12}=1)$ 是多少？$\mathrm{P}(X_{12}=1 \text{ 及 } X_{23}=1)$ 又是多少？

阶段 1 两两独立（pairwise independence）

涵盖：第 (i) 部分：独立性

想一想 $\mathrm{P}(X_{12}=1)$ 是多少？$\mathrm{P}(X_{12}=1 \text{ 及 } X_{23}=1)$ 又是多少？把后者与两个概率的乘积比较一下。

阶段 2 得分的平均值（mean）与方差（variance）

涵盖：第 (i) 部分：平均值与方差

停一停，试一试 得分是 $\binom{n}{2}$ 个变量之和。期望值（expectation）永远可以相加。一个和的方差在甚么情况下等于各项方差之和？只有两两独立是否足够？

阶段 3 方差公式

涵盖：第 (ii) 部分

想一想 若各个 Y_i 的平均值为零，则和的方差就是其平方的 E。把平方展开：每个乘积 Y_iY_j 出现多少次？

阶段 4　单双游戏

涵盖：第 (iii) 部分：并非独立

想一想　求 Z_{12} 的分布（distribution）及其平均值。$Z_{12}Z_{23}$ 在甚么情况下不等于零，这时它的值是多少？比较 $\mathrm{E}(Z_{12}Z_{23})$ 与 $\mathrm{E}(Z_{12})\mathrm{E}(Z_{23})$。

阶段 5 方差

涵盖：第 (iii) 部分：方差

> **停一停，试一试** 运用第 (ii) 部分的结果。哪些“一对配对”会对交叉项（cross terms）有贡献：是没有共同玩家的，还是恰好有一名共同玩家的？两类各有多少对？

背后的数学

多个事件的独立性

如果对每一个子族，交集的概率都等于各概率之积，则称事件 $A_1, \ldots, A_n$ 独立。对三个事件来说，这意味着四个条件：三个是关于每一对的，另一个是 $\mathrm{P}(A \cap B \cap C) = \mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C)$。伯恩斯坦（Sergei Bernstein）在 1920 年代提出图 12.4.1（见网上该节的完整版）中的例子，说明两两之间的条件并不能推出三者的条件。（反过来，单凭三者的条件也不能推出两两之间的条件。）题目中“点数相同”的事件 X_{12}、X_{23}、X_{13} 与伯恩斯坦的硬币如出一辙。

不相关不等于独立

协方差为零，即 $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$，是比独立更弱的条件，称为不相关（uncorrelated）。STEP III 2000 Q14（见练习）说明，对于只取两个值的变量，这两个概念是一致的，但取三个值时则不然。一个简单的例子：设 X 在 $\{-1, 0, 1\}$ 上均匀分布，并设 $Y = X^2$。那么 $\mathrm{E}(XY) = \mathrm{E}(X^3) = 0 = \mathrm{E}(X)\mathrm{E}(Y)$，然而 Y 完全由 X 决定。协方差只量度线性关联。

廉价的随机性

对电脑来说，产生随机数的成本很高。1979 年，Larry Carter 和 Mark Wegman 证明，对许多算法而言，例如把数据储存在伪随机位置的杂凑（hashing），两两独立的随机选

择与完全独立的同样有效。两两独立的数值可以用极低的成本产生：例如，取随机的 a 和 b，则 $a + bk \pmod{p}$ 对 $k = 0, 1, \ldots, p-1$ 给出 p 个数值，其中任何两个都是独立的。本题的方差计算正是它行得通的原因：切比雪夫（Chebyshev）不等式只需要方差，而方差只需要看成对的变量。

延伸练习

1 **两个值。**试做 STEP III 2000，第 14 题（step.maths.org/questions/00-s3-q14）。该题证明，对于各自只取两个值的随机变量，$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$ 必然推出独立，并要求给出一个取三个值的反例。

2 **人数更多的队伍。**在第 (i) 部分中取 $n = 7$，求得分的平均值和标准差，并把例子中的得分 7 与它们比较。

3 **再谈生日。**利用指示变量，求 k 个人当中生日相同的人数对的平均值和方差（Section 9.9）。这些指示变量是否两两独立？

4 **三者成立但两两不独立。**找出三个事件，每个的概率都是 $\frac{1}{2}$，满足 $\mathrm{P}(A \cap B \cap C) = \frac{1}{8}$，但并非两两独立。

延伸阅读

- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press），开首几章讨论独立性。
- **Michael Mitzenmacher & Eli Upfal, *Probability and Computing***（Cambridge University Press）。两两独立及其在算法中的应用。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press；可于网上免费阅读），关于指示变量与协方差。

完整的引导式解答见第 12.4 节：

https://beyondstep.keithmcnulty.org/zh-hans/ch12/pairwise-independence.html

每星期一道 STEP 题目，完整讲解：The STEP Up，以电子邮件免费发送，订阅网址：keithmcnulty.substack.com