# 12.4 兩兩獨立，卻非相互獨立

## 工作紙

### 題目

**STEP III 2021，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 2021 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

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

## 提示

請先自行嘗試解題。如果遇到困難，請逐一閱讀提示，每讀完一個便回到題目再試。標有停一停，試一試的提示是關鍵步驟。每個提示之後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？

### 階段 1　兩兩獨立（pairwise independence）

涵蓋：第 (i) 部分：獨立性

**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？把後者與兩個概率的乘積比較一下。

### 階段 2 得分的平均值（mean）與方差（variance）

涵蓋：第 (i) 部分：平均值與方差

**停一停，試一試** 得分是 $\binom{n}{2}$ 個變量之和。期望值（expectation）永遠可以相加。一個和的方差在甚麼情況下等於各項方差之和？只有兩兩獨立是否足夠？

### 階段 3 方差公式

涵蓋：第 (ii) 部分

**想一想** 若各個 $Y_i$ 的平均值為零，則和的方差就是其平方的 E。把平方展開：每個乘積 $Y_iY_j$ 出現多少次？

### 階段 4　單雙遊戲

涵蓋：第 (iii) 部分：並非獨立

? **想一想**　求 $Z_{12}$ 的分佈（distribution）及其平均值。$Z_{12}Z_{23}$ 在甚麼情況下不等於零，這時它的值是多少？比較 $\mathrm{E}(Z_{12}Z_{23})$ 與 $\mathrm{E}(Z_{12})\mathrm{E}(Z_{23})$。

**階段 5** **方差**

涵蓋：第 (iii) 部分：方差

> **停一停，試一試** 運用第 (ii) 部分的結果。哪些「一對配對」會對交叉項（cross terms）有貢獻：是沒有共同玩家的，還是恰好有一名共同玩家的？兩類各有多少對？

## 背後的數學

### 多個事件的獨立性

如果對每一個子族，交集的概率都等於各概率之積，則稱事件 $A_1, \ldots, A_n$ 獨立。對三個事件來說，這意味着四個條件：三個是關於每一對的，另一個是 $\mathrm{P}(A \cap B \cap C) = \mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C)$。伯恩斯坦（Sergei Bernstein）在 1920 年代提出圖 12.4.1（見網上該節的完整版）中的例子，說明兩兩之間的條件並不能推出三者的條件。（反過來，單憑三者的條件也不能推出兩兩之間的條件。）題目中「點數相同」的事件 $X_{12}$、$X_{23}$、$X_{13}$ 與伯恩斯坦的硬幣如出一轍。

### 不相關不等於獨立

協方差為零，即 $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$，是比獨立更弱的條件，稱為不相關（uncorrelated）。STEP III 2000 Q14（見練習）說明，對於只取兩個值的變量，這兩個概念是一致的，但取三個值時則不然。一個簡單的例子：設 $X$ 在 $\{-1, 0, 1\}$ 上均勻分佈，並設 $Y = X^2$。那麼 $\mathrm{E}(XY) = \mathrm{E}(X^3) = 0 = \mathrm{E}(X)\mathrm{E}(Y)$，然而 $Y$ 完全由 $X$ 決定。協方差只量度線性關聯。

#### 廉價的隨機性

對電腦來說，產生隨機數的成本很高。1979 年，Larry Carter 和 Mark Wegman 證明，對許多演算法而言，例如把數據儲存在偽隨機位置的雜湊（hashing），兩兩獨立的隨機選擇與完全獨立的同樣有效。兩兩獨立的數值可以用極低的成本產生：例如，取隨機的 $a$ 和 $b$，則 $a + bk \pmod{p}$ 對 $k = 0, 1, \ldots, p-1$ 給出 $p$ 個數值，其中任何兩個都是獨立的。本題的方差計算正是它行得通的原因：切比雪夫（Chebyshev）不等式只需要方差，而方差只需要看成對的變量。

## 延伸練習

1 **兩個值。**試做 STEP III 2000，第 14 題（step.maths.org/questions/00-s3-q14）。該題證明，對於各自只取兩個值的隨機變量，$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$ 必然推出獨立，並要求給出一個取三個值的反例。

2 **人數更多的隊伍。**在第 (i) 部分中取 $n = 7$，求得分的平均值和標準差，並把例子中的得分 7 與它們比較。

3 **再談生日。**利用指示變量，求 $k$ 個人當中生日相同的人數對的平均值和方差（Section 9.9）。這些指示變量是否兩兩獨立？

4 **三者成立但兩兩不獨立。**找出三個事件，每個的概率都是 $\frac{1}{2}$，滿足 $\mathrm{P}(A \cap B \cap C) = \frac{1}{8}$，但並非兩兩獨立。

## 延伸閱讀

- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes***（Oxford University Press），開首幾章討論獨立性。
- **Michael Mitzenmacher & Eli Upfal, *Probability and Computing***（Cambridge University Press）。兩兩獨立及其在演算法中的應用。
- **Joseph Blitzstein & Jessica Hwang, *Introduction to Probability***（CRC Press；可於網上免費閱讀），關於指示變量與協方差。

完整的引導式解答見第 12.4 節：

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

每星期一道 STEP 題目，完整講解：The STEP Up，以電郵免費發送，訂閱網址：keithmcnulty.substack.com