解答例

STEP III 2021、第 12 問

確率・統計 ・ 和の分散

> **STEP III 2021、問 12**
>
> 先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2021 年の STEP 過去問一式（zip ファイル）のうち、STEP III の問題用紙にある：
>
> www.ocr.org.uk/Images/696588-2021-past-papers.zip

この問題は、*Beyond STEP* の第 12.4 節「**二つずつなら独立、全体では独立でない**」で詳しく解説されている：https://beyondstep.keithmcnulty.org/ja/ch12/pairwise-independence.html

鍵となる考え方

チームの得点を、選手のペアごとに 1 つずつ定めた指示変数（indicator variable）型の確率変数の和として表し、その和の分散（variance）を展開する。すると、すべては 2 つの**異なる**ペアがどう関わり合うかにかかってくる。2 つのペアに共通の選手がいなければ明らかに独立（independent）であるから、本当に考えるべきなのは選手を 1 人共有するペアである。(i) ではそのようなペアは独立であることがわかり、交差項は消える。(iii) ではそうはならず、(ii) の恒等式がそれらの寄与を数え上げる道具となる。

解答

小問 (i)

各さいころは公正であり、選手たちは独立にさいころを振る。指定された 2 人の選手が同じ目を出す確率は、1 人目が何を出したかによらず $\frac{1}{6}$ である。よって

$$\mathrm{P}(X_{12}=1)=\mathrm{P}(X_{23}=1)=\tfrac{1}{6}.$$

$X_{12}=1$ と $X_{23}=1$ がともに起こるのは、選手 1、2、3 がすべて同じ目を出すときであり、かつそのときに限るから、

$$\mathrm{P}(X_{12}=1,\ X_{23}=1)=\tfrac{1}{6}\cdot\tfrac{1}{6}=\tfrac{1}{36}=\mathrm{P}(X_{12}=1)\,\mathrm{P}(X_{23}=1).$$

これを周辺確率（marginal probability）から引くと、残りの 3 つの場合の確率が得られる。たとえば $\mathrm{P}(X_{12}=1, X_{23}=0) = \frac{1}{6} - \frac{1}{36} = \frac{1}{6}\cdot\frac{5}{6}$ であり、同様に $\mathrm{P}(X_{12}=0, X_{23}=1) = \frac{5}{6}\cdot\frac{1}{6}$、$\mathrm{P}(X_{12}=0, X_{23}=0) = 1 - \frac{1}{6} - \frac{1}{6} + \frac{1}{36} = \frac{5}{6}\cdot\frac{5}{6}$ である。したがって同時分布（joint distribution）は周辺分布の積であり、X_{12} と X_{23} は独立である。

考えてみよう $X_{12}=1$ であると知っても、なぜ X_{23} について何もわからないのか。

選手 1 と 2 の目が一致したという条件のもとでも、選手 2 の目は依然として $\{1,\dots,6\}$ 上の一様分布（uniform distribution）に従い、選手 3 がそれに一致する確率は以前とまったく同じく $\frac{1}{6}$ である。どの目も等しく出やすいので、共有される選手は何の情報ももたらさないのである。

$T = \sum_{i<j} X_{ij}$ を総得点とする。これは $N = \binom{n}{2}$ 個の項の和である。各 X_{ij} はパラメータ $\frac{1}{6}$ のベルヌーイ分布（Bernoulli distribution）に従うので、$\mathrm{E}(X_{ij}) = \frac{1}{6}$、$\mathrm{Var}(X_{ij}) = \frac{1}{6}\cdot\frac{5}{6} = \frac{5}{36}$ である。ゆえに

$$\mathrm{E}(T) = \frac{N}{6} = \frac{n(n-1)}{12}.$$

分散を求めるために、任意の異なる 2 つのペアをとる。共通の選手がいなければ、それらは互いに素なさいころの集合に依存するので独立である。選手を 1 人共有する場合も、選手の番号を付け替えれば上の議論により独立である。よって X_{ij} は対ごとに独立であり、共分散（covariance）はすべて 0 となるので、和の分散は分散の和に等しい：

$$\mathrm{Var}(T) = N\cdot\frac{5}{36} = \frac{5n(n-1)}{72}.$$

小問 (ii)

$S = Y_1 + \cdots + Y_m$ とおく。期待値の線形性（linearity）より $\mathrm{E}(S) = 0$ であるから、$\mathrm{Var}(S) = \mathrm{E}(S^2) - 0^2 = \mathrm{E}(S^2)$ である。平方を展開すると、各 Y_i^2 は 1 回ずつ、$i<j$ である各積 Y_iY_j は 2 回ずつ現れる：

$$S^2 = \sum_{i=1}^{m} Y_i^2 + 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} Y_iY_j.$$

期待値をとって線形性を用いれば、求める式が得られる。

小問 (iii)

2 人の選手について、$\mathrm{P}(\text{同じ偶数}) = \frac{3}{36} = \frac{1}{12}$ であり、同様に $\mathrm{P}(\text{同じ奇数}) = \frac{1}{12}$ である。よって

$$\mathrm{P}(Z_{ij}=1) = \mathrm{P}(Z_{ij}=-1) = \tfrac{1}{12}, \qquad \mathrm{P}(Z_{ij}=0) = \tfrac{5}{6}.$$

独立でないこと。 $Z_{12}=1$ かつ $Z_{23}=-1$ であるとすると、選手 2 の目は偶数でも奇数でもなければならない。よって

$$\mathrm{P}(Z_{12}=1,\ Z_{23}=-1)=0\neq\tfrac{1}{12}\cdot\tfrac{1}{12}=\mathrm{P}(Z_{12}=1)\,\mathrm{P}(Z_{23}=-1),$$

となり、Z_{12} と Z_{23} は独立ではない。

平均。 対称性（symmetry）より $\mathrm{E}(Z_{ij})=\frac{1}{12}-\frac{1}{12}=0$ であるから、総得点 $U=\sum_{i<j}Z_{ij}$ の平均（mean）は 0 である。

分散。 どの Z_{ij} も平均が 0 なので、$N=\binom{n}{2}$ 個の変数 Z_{ij} に (ii) を適用できる。必要な期待値は 2 種類である。

- 平方：Z_{ij}^2 は選手 i と j の目が一致するときに限り 1 となるので、$\mathrm{E}(Z_{ij}^2)=\frac{1}{6}$ である。
- 異なる 2 つのペアの積：2 つのペアに共通の選手がいなければ、2 つの変数は独立で平均が 0 であるから、$\mathrm{E}(Z_{ij}Z_{kl})=\mathrm{E}(Z_{ij})\,\mathrm{E}(Z_{kl})=0$ である。選手を共有する場合、たとえば $Z_{12}Z_{23}$ では、積が 0 でないのは 3 人全員が同じ目を出すときに限る。このとき 2 つの因子はともに $+1$（偶数）かともに -1（奇数）であるから、積は 1 となる。ゆえに $\mathrm{E}(Z_{12}Z_{23})=\frac{1}{36}$ である。

> **考えてみよう** ちょうど 1 人の選手を共有するペアの組（順序なし）はいくつあるか。

共有される選手の選び方が n 通り、残りの $n-1$ 人からその 2 人の相手を選ぶ方法が $\binom{n-1}{2}$ 通りある。よってそのようなペアの組は $n\binom{n-1}{2}=\frac{1}{2}n(n-1)(n-2)$ 個であり、それぞれ 1 回ずつ数えられている。これがまさに (ii) の二重和で必要なものである。

以上をまとめると、

$$\begin{aligned}\mathrm{Var}(U)&=\binom{n}{2}\cdot\frac{1}{6}+2\cdot\frac{n(n-1)(n-2)}{2}\cdot\frac{1}{36}\\&=\frac{n(n-1)}{36}\big(3+(n-2)\big)\\&=\frac{n(n-1)(n+1)}{36}=\frac{1}{36}n(n^2-1),\end{aligned}$$

となり、題意が示された。

検算

- 例として出目が 3,3,3,3,6,6,2 のとき、ペアの数え上げでは $6+1=7$、符号付きの数え上げでは $-6+1=-5$ となり、問題文と一致する。

- $n = 2$ のときペアは 1 つしかないので $\mathrm{Var}(U) = \mathrm{E}(Z_{12}^2) = \frac{1}{6}$ であり、公式も $\frac{2\cdot 3}{36} = \frac{1}{6}$ を与える。
- $n = 3$ のとき、3 つのさいころの目がすべて一致すれば $U = \pm 3$ である（確率はそれぞれ $\frac{1}{72}$）。そうでないときはちょうど 1 つのペアが一致して $U = \pm 1$ となり、その確率の合計は $\frac{15}{36}$ である。よって $\mathrm{E}(U^2) = 9 \cdot \frac{1}{36} + \frac{15}{36} = \frac{2}{3}$ となり、$\frac{3\cdot 8}{36}$ と一致する。$n \leq 6$ についてコンピュータですべての結果を厳密に数え上げると、2 つの平均と 2 つの分散がいずれも確かめられる。たとえば $n = 6$ では、(i) で $\frac{25}{12}$、(iii) で $\frac{35}{6}$ となる。

よくある間違い

- 1 つの場合だけを確かめて、残りの場合がなぜ従うのかを述べないまま X_{12} と X_{23} の独立性を主張すること。2 値の変数では実際に従うが、そのことを明記すべきである。
- (iii) で、選手を共有するペアは共通の目が偶数でも奇数でも $+\frac{1}{36}$ を寄与することを忘れること。積をとるとマイナス符号は打ち消し合う。
- 重なり合うペアの組の個数を数え間違えること。たとえば（順序付きの）$n(n-1)(n-2)$ を用いたうえで、さらに 2 を掛けてしまう。

STEP のその先へ

小問 (i) は、**対ごとに独立**（pairwise independent）であるが**相互に独立**（mutually independent）ではない確率変数の典型例である。実際、$X_{12} = X_{23} = 1$ ならば $X_{13} = 1$ が強制される。分散の計算に必要なのは対ごとの独立性だけであり、この考え方は乱択アルゴリズム（randomised algorithm）や誕生日問題（birthday problem）で用いられる。小問 (iii) は、項どうしが従属するときの和の共分散を扱う一般的な方法を示している。どちらも Blitzstein and Hwang, *Introduction to Probability*（確率論入門）で詳しく展開されている。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。