# 復元、非復元抽出における期待値について

静岡県立高等学校教諭　石原　諭

これはよくある確率の期待値の問題ですが、自分自身不思議に感じたことがあったので、考えてみました。

次の異なる 3 つの場合について、期待値を考えます。

（Case 1）**赤球 4 個と白球 3 個が入っている袋から 3 個の球を 同時に取り出す とき、そこに含まれる赤球の個数の期待値** $E_1$

この場合の期待値 $E_1$ は次のように計算されます。

（解答）赤球が含まれる個数は $0,\ 1,\ 2,\ 3$ 個の 4 通りである。7 個から 3 個取り出す取り出し方は ${}_7\mathrm{C}_3 = \dfrac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 35$ 通りである。赤球 0 個を取り出す取り出し方は ${}_4\mathrm{C}_0 \cdot {}_3\mathrm{C}_3 = 1$ 通り、赤球 1 個を取り出す取り出し方は ${}_4\mathrm{C}_1 \cdot {}_3\mathrm{C}_2 = 12$ 通り、赤球 2 個を取り出す取り出し方は ${}_4\mathrm{C}_2 \cdot {}_3\mathrm{C}_1 = 18$ 通り、赤球 3 個を取り出す取り出し方は ${}_4\mathrm{C}_3 \cdot {}_3\mathrm{C}_0 = 4$ 通りであるから、赤球の個数を確率変数 $X$ とすれば、確率分布表は次のようになる。

| $X$ | 0 | 1 | 2 | 3 | 計 |
|---|---|---|---|---|---|
| $P$ | $\frac{1}{35}$ | $\frac{12}{35}$ | $\frac{18}{35}$ | $\frac{4}{35}$ | 1 |

したがって、期待値 $E_1$ は

$$E_1 = 0 \times \frac{1}{35} + 1 \times \frac{12}{35} + 2 \times \frac{18}{35} + 3 \times \frac{4}{35} = \frac{60}{35} = \frac{12}{7}\text{（個）}$$

となる。

では今の（Case 1）を少し変えた次の場合を考えてみます。

（Case 2）**赤球 4 個と白球 3 個が入っている袋から 1 個ずつ 毎回元の袋に戻しながら 3 個の球を取り出す とき、そこに含まれる赤球の個数の期待値** $E_2$

このような復元抽出の場合、確率変数 $X$ は二項分布 $B\left(3, \dfrac{4}{7}\right)$ に従うので、期待値の公式より、期待値 $E_2$ は

$$E_2 = 3 \times \frac{4}{7} = \frac{12}{7}\text{（個）}$$

となる。

では最後に（Case 1）を次のように変えた場合を考えてみます。

**（Case 3）赤球 4 個と白球 3 個が入っている袋から 1 個ずつ 取り出した球は元に戻さずに 3 個の球を取り出す とき、そこに含まれる赤球の個数の期待値 $E_3$**

これは非復元抽出です。確率を計算すると次のようになります。

$$P(X=0)=\frac{3}{7}\cdot\frac{2}{6}\cdot\frac{1}{5}=\frac{1}{35}$$

以下同様にして（少し省略しますが $\cdots$）

$$P(X=1)=\frac{12}{35},\ \ P(X=2)=\frac{18}{35},\ \ P(X=3)=\frac{4}{35}$$

となりますので、確率分布表は（Case 1）のものと一致し、よって期待値 $E_3$ は

$$E_3=0\times\frac{1}{35}+1\times\frac{12}{35}+2\times\frac{18}{35}+3\times\frac{4}{35}=\frac{60}{35}=\frac{12}{7}\text{（個）}$$

となります。

以上、「同時に取り出す場合」、「復元抽出の場合」、「非復元抽出の場合」のいずれの場合も同じ期待値になりました。どうしてこのようになるのか、また球の個数をいろいろ変えても同じような結果が得られるのだろうか？という疑問が湧きました。それがこのようなことを考えるようになったきっかけです。

さて結論から申しますと、球の個数をいろいろ変えても、3 つのいずれの場合に同じ期待値が得られます。以下ではその証明を与えます。

赤球 $m$ 個と白球 $n$ 個の入っている袋から $k$ 個の球を次のように取り出す。

(i)　$k$ 個同時に取り出す
(ii)　1 個ずつ元に戻して復元抽出で取り出す
(iii) 1 個ずつ元に戻さずに非復元抽出で取り出す

このときの赤球の数の期待値は、(i),(ii),(iii) のいずれも $\dfrac{mk}{m+n}$（個）である。

**（証明）**

(ii) の場合の期待値は先ほど述べたように、二項分布 $B\left(k,\dfrac{m}{m+n}\right)$ に従うので、公式より、$E_2=\dfrac{mk}{m+n}$ である。

(i) の場合と (iii) の場合の期待値が一致することは感覚的にも明らかであるが、式で表せば、(i) の場合 $P(X=l)=\dfrac{{}_m\mathrm{P}_l\cdot{}_n\mathrm{P}_{k-l}}{{}_{m+n}\mathrm{P}_k}\cdot{}_k\mathrm{C}_l$ であり、これを式変形すると、簡単な計算で $\dfrac{{}_m\mathrm{C}_l\cdot{}_n\mathrm{C}_{k-l}}{{}_{m+n}\mathrm{C}_k}$ となることから示される。（詳しい計算は省略）したがって (iii) の期待値が $\dfrac{mk}{m+n}$ になることを示せば証明は終わる。

袋から球を1個ずつ $k$ 回取り出す。ここで $i$ 回目の赤球の数を確率変数 $X_i$ とする。よって $X_i = 0,\ 1$ である。（ただし、$1 \leqq i \leqq k$）また、$k$ 回の非復元抽出の後の赤球の数を確率変数 $X$ とする。

このとき $X = \displaystyle\sum_{i=1}^{k} X_i$ が成り立つ。したがって $E(X) = \displaystyle\sum_{i=1}^{k} E(X_i)$ である。

ここで、$E(X_1) = E(X_2) = \cdots\cdots = E(X_k)$ である。なぜなら、「くじびきの公平性」より、非復元抽出の場合であっても、各回において、赤球を取り出す確率は常に $\dfrac{m}{m+n}$ であるから、

$$E(X_i) = 0 \times \frac{n}{m+n} + 1 \times \frac{m}{m+n} = \frac{m}{m+n}$$

となるので、$E(X_1) = E(X_2) = \cdots\cdots = E(X_k)$ である。

よって、$E(X) = \displaystyle\sum_{i=1}^{k} E(X_i) = \frac{mk}{m+n}$ となり示された。**（証明終）**

（補足）「くじびきの公平性」のところを詳しく説明すれば次のようになる。

袋から球を1個ずつ取り出して、取り出した順に1列に左から右へ並べる順列を考える。この順列の総数は $(m+n)!$ である。また、$1 \leqq i \leqq k$ とし、$i$ 番目の位置に赤球が置かれるような順列の総数は

$${}_mP_1 \times (m+n-1)! = m \times (m+n-1)!$$

である。したがって $i$ 回目に赤球を取り出す確率は

$$\frac{m \times (m+n-1)!}{(m+n)!} = \frac{m}{m+n}$$

となる。