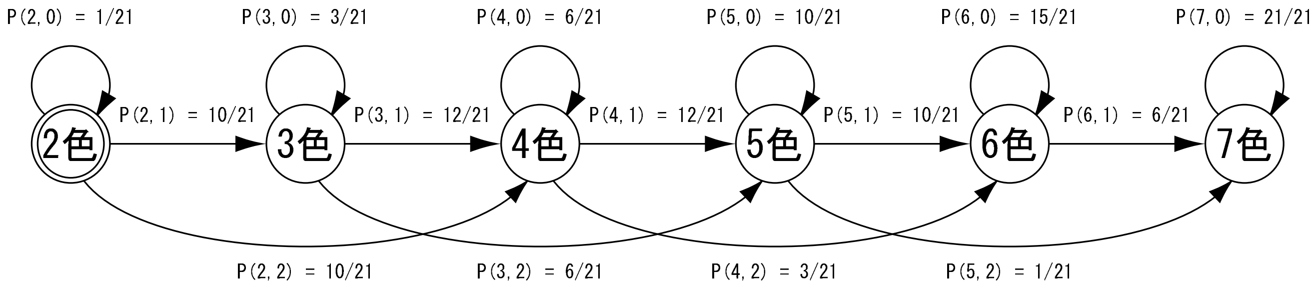
(1) 4 箱から取り出したときの色数の期待値

一般に n 箱から合計 $2n$ 個取り出したときの色数の期待値を求める。
1 箱目は必ず 2 色取り出されるので、その状態を初期状態としてさらに m 箱 $(m = n - 1)$ から取り出すと考える。
現在 x 色取り出しているときに新しい箱から 2 個取り出して y 色増える確率を $P(x, y)$ とすると、

$$
\begin{aligned}
P(x,0) &= \frac{{}_x\mathrm{C}_2}{{}_7\mathrm{C}_2} = \frac{x(x-1)}{42} \\
P(x,1) &= \frac{{}_x\mathrm{C}_1 \times {}_{7-x}\mathrm{C}_1}{{}_7\mathrm{C}_2} = \frac{x(7-x)}{21} \\
P(x,2) &= \frac{{}_{7-x}\mathrm{C}_2}{{}_7\mathrm{C}_2} = \frac{(7-x)(6-x)}{42}
\end{aligned}
$$

箱から 2 個の玉を取り出すごとの色数の変化は以下の非決定性有限状態オートマトンとして表すことができる。



遷移確率を行列で表すと、

$$
\mathbf{P} = \begin{pmatrix}
P(2,0) & 0 & 0 & 0 & 0 & 0 \\
P(2,1) & P(3,0) & 0 & 0 & 0 & 0 \\
P(2,2) & P(3,1) & P(4,0) & 0 & 0 & 0 \\
0 & P(3,2) & P(4,1) & P(5,0) & 0 & 0 \\
0 & 0 & P(4,2) & P(5,1) & P(6,0) & 0 \\
0 & 0 & 0 & P(5,2) & P(6,1) & P(7,0)
\end{pmatrix}
= \frac{1}{21}\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
10 & 3 & 0 & 0 & 0 & 0 \\
10 & 12 & 6 & 0 & 0 & 0 \\
0 & 6 & 12 & 10 & 0 & 0 \\
0 & 0 & 3 & 10 & 15 & 0 \\
0 & 0 & 0 & 1 & 6 & 21
\end{pmatrix}
$$

となる。$\mathbf{P}_{i,j}$ は $(j+1)$ 色の状態で新しい箱から 2 個取り出したときに $(i+1)$ 色になる確率である。
初期状態（2 色）から始めて m 個の箱から $2m$ 個の玉を取り出したときに r 色になっている確率を $P_r(m)$ とすると、

$$
\begin{pmatrix} P_2(m) \\ P_3(m) \\ P_4(m) \\ P_5(m) \\ P_6(m) \\ P_7(m) \end{pmatrix}
= \mathbf{P}^m \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
$$

である。（$\mathbf{P}^m$ のあとの列ベクトルは初期状態が 2 色であることを表している）
$\mathbf{P}$ を対角化して $\mathbf{P}^m$ を計算すると $P_r(m)$ は以下のようになる。（途中計算省略）

$$
\begin{pmatrix} P_2(m) \\ P_3(m) \\ P_4(m) \\ P_5(m) \\ P_6(m) \\ P_7(m) \end{pmatrix}
= \mathbf{P}^m \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
= \frac{1}{21^m}\begin{pmatrix}
1 \\
-5 + 5 \cdot 3^m \\
10 - 20 \cdot 3^m + 10 \cdot 6^m \\
-10 + 30 \cdot 3^m - 30 \cdot 6^m + 10 \cdot 10^m \\
5 - 20 \cdot 3^m + 30 \cdot 6^m - 20 \cdot 10^m + 5 \cdot 15^m \\
-1 + 5 \cdot 3^m - 10 \cdot 6^m + 10 \cdot 10^m - 5 \cdot 15^m + 21^m
\end{pmatrix}
\quad (m \geqq 0)
$$

よって、最初の 1 箱も含めて n 箱から合計 $2n$ 個取り出したときの色数の期待値 $E_{\mathrm{Color}}(n)$ は、

$$
\begin{aligned}
E_{\mathrm{Color}}(n) &= \sum_{i=2}^{7} iP_i(n-1) \\
&= \Big(2\cdot 1 \\
&\quad + 3\left(-5 + 5\cdot 3^{n-1}\right) \\
&\quad + 4\left(10 - 20\cdot 3^{n-1} + 10\cdot 6^{n-1}\right) \\
&\quad + 5\left(-10 + 30\cdot 3^{n-1} - 30\cdot 6^{n-1} + 10\cdot 10^{n-1}\right) \\
&\quad + 6\left(5 - 20\cdot 3^{n-1} + 30\cdot 6^{n-1} - 20\cdot 10^{n-1} + 5\cdot 15^{n-1}\right) \\
&\quad + 7\left(-1 + 5\cdot 3^{n-1} - 10\cdot 6^{n-1} + 10\cdot 10^{n-1} - 5\cdot 15^{n-1} + 21^{n-1}\right)\Big)/21^{n-1} \\
&= \frac{7^n - 5^n}{7^{n-1}} \quad (n \geqq 1)
\end{aligned}
$$

となる。$n = 4$ の場合は、

$$
\begin{aligned}
E_{\mathrm{Color}}(4) &= \frac{7^4 - 5^4}{7^3} \\
&= \frac{1776}{343} \\
&\simeq 5.178(\text{色})
\end{aligned}
$$

である。

(2) 7 色そろうまでの箱数の期待値

初期状態（2 色）から始めて、m 箱から $2m$ 個の玉を取り出したときに 7 色になっている確率が $P_7(m)$ なので、m 箱目で初めて 7 色になる確率 $Q_7(m)$ は、

$$
\begin{aligned}
Q_7(m) &= P_7(m) - P_7(m-1) \\
&= \left(-1 + 5\cdot 3^m - 10\cdot 6^m + 10\cdot 10^m - 5\cdot 15^m + 21^m\right)/21^m \\
&\quad + \left(-1 + 5\cdot 3^{m-1} - 10\cdot 6^{m-1} + 10\cdot 10^{m-1} - 5\cdot 15^{m-1} + 21^{m-1}\right)/21^{m-1} \\
&= 20\cdot\frac{1}{21^m} - 30\cdot\frac{1}{7^m} + 25\cdot\frac{2^m}{7^m} - 11\cdot\frac{10^m}{21^m} + 2\cdot\frac{5^m}{7^m} \quad (m \geqq 1)
\end{aligned}
$$

よって、最初の 1 箱を除いて 7 色そろうのに必要な箱数の期待値 $E_{\mathrm{Box}}(7)$ は、

$$
\begin{aligned}
E_{\mathrm{Box}}(7) &= \sum_{i=1}^{\infty} iQ_7(i) \\
&= \sum_{i=1}^{\infty}\left(20\cdot\frac{i}{21^i} - 30\cdot\frac{i}{7^i} + 25\cdot\frac{i}{\left(\frac{7}{2}\right)^i} - 11\cdot\frac{i}{\left(\frac{21}{10}\right)^i} + 2\cdot\frac{i}{\left(\frac{7}{5}\right)^i}\right)
\end{aligned}
$$

$\displaystyle\sum_{i=1}^{\infty}\frac{i}{x^i} = \frac{x}{(x-1)^2}$ より、

$$
\begin{aligned}
E_{\mathrm{Box}}(7) &= \frac{21}{20} - \frac{35}{6} + 14 - \frac{210}{11} + \frac{35}{2} \\
&= \frac{5033}{660}
\end{aligned}
$$

ゆえに、最初の 1 箱を含めた期待値 $E^{+}_{\mathrm{Box}}(7)$ は

$$
\begin{aligned}
E^{+}_{\mathrm{Box}}(7) &= E_{\mathrm{Box}}(7) + 1 \\
&= \frac{5693}{660} \\
&= 8.62\dot{5}\dot{7}(\text{箱})
\end{aligned}
$$

付録

$\mathbf{P}^m$ の残りの列は以下のようになる。

$$\mathbf{P}^m \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{21^m} \begin{pmatrix} 0 \\ 3^m \\ -4\cdot 3^m + 4\cdot 6^m \\ 6\cdot 3^m - 12\cdot 6^m + 6\cdot 10^m \\ -4\cdot 3^m + 12\cdot 6^m - 12\cdot 10^m + 4\cdot 15^m \\ 3^m - 4\cdot 6^m + 6\cdot 10^m - 4\cdot 15^m + 21^m \end{pmatrix}$$

$$\mathbf{P}^m \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{21^m} \begin{pmatrix} 0 \\ 0 \\ 6^m \\ -3\cdot 6^m + 3\cdot 10^m \\ 3\cdot 6^m - 6\cdot 10^m + 3\cdot 15^m \\ -6^m + 3\cdot 10^m - 3\cdot 15^m + 21^m \end{pmatrix}$$

$$\mathbf{P}^m \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{21^m} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 10^m \\ -2\cdot 10^m + 2\cdot 15^m \\ 10^m - 2\cdot 15^m + 21^m \end{pmatrix}$$

$$\mathbf{P}^m \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{21^m} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 15^m \\ -15^m + 21^m \end{pmatrix}$$

$$\mathbf{P}^m \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{21^m} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 21^m \end{pmatrix}$$