おまけが全５種類として、ｎ個購入したとき、何種類そろったかの確率を、

$p_5(n)$…５種類そろっている。

$p_4(n)$…４種類そろっている。

$p_3(n)$…３種類そろっている。

$p_2(n)$…２種類そろっている。

$p_1(n)$…１種類そろっている。

と表現すると、

$$p_5(n+1)=\frac{5}{5}p_5(n)+\frac{1}{5}p_4(n)$$
$$p_4(n+1)=\frac{4}{5}p_4(n)+\frac{2}{5}p_3(n)$$
$$p_3(n+1)=\frac{3}{5}p_3(n)+\frac{3}{5}p_2(n)$$
$$p_2(n+1)=\frac{2}{5}p_2(n)+\frac{4}{5}p_1(n)$$
$$p_1(n+1)=\frac{1}{5}p_1(n)$$

と書くことができる。これは行列で表記すると、

$$\begin{pmatrix} p_5(n+1) \\ p_4(n+1) \\ p_3(n+1) \\ p_2(n+1) \\ p_1(n+1) \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 5 & 1 & 0 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} p_5(n) \\ p_4(n) \\ p_3(n) \\ p_2(n) \\ p_1(n) \end{pmatrix}$$

となる。これを見ればぴーんときて、

$$\begin{pmatrix} p_5(n) \\ p_4(n) \\ p_3(n) \\ p_2(n) \\ p_1(n) \end{pmatrix} = \frac{1}{5^{n-1}}\mathbf{A}^{n-1}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{ただし、} \mathbf{A}=\begin{pmatrix} 5 & 1 & 0 & 0 & 0 \\ 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

であると、あっという間にわかる。

行列のべき乗をやっつけるには対角化しかない。
$\mathbf{A}$ の固有値を調べてみると、5,4,3,2,1 であることが簡単にわかる。
$\mathbf{A}$ の固有ベクトルも、$\mathbf{A}$ が上三角行列なので、すいすい求まる。
このとき階乗に似た計算がでてくるけれど、それがコンビネーション（組み合わせ$C_{n,r}$）であることは、すぐに気がつくだろう。

こうやって求めた固有ベクトルを都合よく並べることで、

$$\mathbf{P}=\begin{pmatrix}1 & -C_{1,1} & C_{2,2} & -C_{3,3} & C_{4,4}\\ 0 & 1 & -C_{2,1} & C_{3,2} & -C_{4,3}\\ 0 & 0 & 1 & -C_{3,1} & C_{4,2}\\ 0 & 0 & 0 & 1 & -C_{4,1}\\ 0 & 0 & 0 & 0 & 1\end{pmatrix} \text{ と、} \mathbf{P}^{-1}=\begin{pmatrix}1 & C_{1,1} & C_{2,2} & C_{3,3} & C_{4,4}\\ 0 & 1 & C_{2,1} & C_{3,2} & C_{4,3}\\ 0 & 0 & 1 & C_{3,1} & C_{4,2}\\ 0 & 0 & 0 & 1 & C_{4,1}\\ 0 & 0 & 0 & 0 & 1\end{pmatrix}$$

を算出する。$\mathbf{P}^{-1}$は、$\mathbf{P}$から計算して求めるのだけれども、具体的な数字を実際に計算してみると、なんとその計算結果は、$\mathbf{P}$の各要素のマイナス符号をはずしたものになっている！ 実は、おまけの種類を 5 種類に限定しない、すなわち、一般的な条件においても、$\mathbf{P}^{-1}$が、$\mathbf{P}$の各要素のマイナス符号をはずしたものになっていると、証明することができる。この証明は後で行なう。
さて、$\mathbf{P}$、$\mathbf{P}^{-1}$を利用すると、

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\begin{pmatrix}5 & 0 & 0 & 0 & 0\\ 0 & 4 & 0 & 0 & 0\\ 0 & 0 & 3 & 0 & 0\\ 0 & 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 0 & 1\end{pmatrix}$$

と対角化できる。これを利用すると、

$$\left(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right)^{n-1}=\mathbf{P}^{-1}\mathbf{A}^{n-1}\mathbf{P}=\begin{pmatrix}5 & 0 & 0 & 0 & 0\\ 0 & 4 & 0 & 0 & 0\\ 0 & 0 & 3 & 0 & 0\\ 0 & 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 0 & 1\end{pmatrix}^{n-1}=\begin{pmatrix}5^{n-1} & 0 & 0 & 0 & 0\\ 0 & 4^{n-1} & 0 & 0 & 0\\ 0 & 0 & 3^{n-1} & 0 & 0\\ 0 & 0 & 0 & 2^{n-1} & 0\\ 0 & 0 & 0 & 0 & 1^{n-1}\end{pmatrix}$$

である。

行列のべき乗をとりのぞいてやると、当初の求めたい確率は、

$$\begin{pmatrix} p_5(n) \\ p_4(n) \\ p_3(n) \\ p_2(n) \\ p_1(n) \end{pmatrix} = \frac{1}{5^{n-1}}\mathbf{A}^{n-1}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{5^{n-1}}\mathbf{P}\begin{pmatrix} 5^{n-1} & 0 & 0 & 0 & 0 \\ 0 & 4^{n-1} & 0 & 0 & 0 \\ 0 & 0 & 3^{n-1} & 0 & 0 \\ 0 & 0 & 0 & 2^{n-1} & 0 \\ 0 & 0 & 0 & 0 & 1^{n-1} \end{pmatrix}\mathbf{P}^{-1}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

であるから、あとはもう、数回の行列計算をがんばるだけでよく、たとえば、$p_5(n)$を求めてみると、

$$p_5(n) = \sum_{i=1}^{5}\left\{(-1)^{5-i}\mathrm{C}_{4,i-1}\left(\frac{i}{5}\right)^{n-1}\right\}$$

となる。

これまで、おまけの種類を５と限定して考えてきたけれど、おまけの種類を N と一般化しても、まったく同じ手順をとることができ、

$$p_m(n) = \sum_{i=1}^{m}\left\{(-1)^{m-i}\mathrm{C}_{N-i,m-i}\mathrm{C}_{N-1,i-1}\left(\frac{i}{N}\right)^{n-1}\right\}$$

を導出することが可能である。ただし、一般条件においては、対角化処理における$\mathbf{P}$を出すところまでは簡単でも、$\mathbf{P}^{-1}$を求めることがむずかしい。$\mathbf{P}$の各要素のマイナス符号をはずしたもの$\mathbf{P}^{?}$が、$\mathbf{P}^{-1}$になっていると予想するのだけれど、その予想は正しいのか？ それを証明するには、$\mathbf{PP}^{?}$が単位行列となることを示せばよい。

さて、一般条件における$\mathbf{P}$の要素（i 行 j 列）をa_{ij}とし、$\mathbf{P}^{?}$の要素をb_{ij}とすると、

$$a_{ij} = \begin{cases} (-1)^{i+j}\mathrm{C}_{j-1,i-1} & (j \ge i) \\ 0 & (j < i) \end{cases} \qquad b_{ij} = \begin{cases} \mathrm{C}_{j-1,i-1} & (j \ge i) \\ 0 & (j < i) \end{cases}$$

である。$\mathbf{PP}^{?}$の計算結果の要素をc_{ij}とすると、

$$c_{ij} = \begin{cases} \displaystyle\sum_{k=i}^{j} a_{ik}b_{kj} = \sum_{k=i}^{j}\left\{(-1)^{i+k}\mathrm{C}_{k-1,i-1}\mathrm{C}_{j-1,k-1}\right\} & (j \ge i) \\ 0 & (j < i) \end{cases}$$

であり、（i=j）のときを区分わければ、

$$c_{ij} = \begin{cases} \sum_{k=i}^{j} \left\{ (-1)^{i+k} \mathrm{C}_{k-1,i-1} \mathrm{C}_{j-1,k-1} \right\} & (j > i) \\ 1 & (j = i) \\ 0 & (j < i) \end{cases}$$

となる。**PP**? が単位行列であるためには、(j>i)のとき、

$$\sum_{k=i}^{j} \left\{ (-1)^{i+k} \mathrm{C}_{k-1,i-1} \mathrm{C}_{j-1,k-1} \right\} = 0$$

であることを示せればよい。左辺を計算すると、

$$\sum_{k=i}^{j} \left\{ (-1)^{i+k} \mathrm{C}_{k-1,i-1} \mathrm{C}_{j-1,k-1} \right\} = \sum_{k=i}^{j} \left\{ (-1)^{i+k} \frac{(k-1)!}{(i-1)!(k-i)!} \frac{(j-1)!}{(k-1)!(j-k)!} \right\}$$

$$= \frac{(j-1)!}{(i-1)!} \sum_{k=i}^{j} \left\{ (-1)^{i+k} \frac{1}{(k-i)!(j-k)!} \right\}$$

となる。ここで、

$$k = i + n \quad (0 \le n \le d, d = j - i > 0)$$

なる n、d を導入し、計算を続けると、

$$= \frac{(j-1)!}{(i-1)!} \sum_{n=o}^{d} \left\{ (-1)^{n} \frac{1}{n!(d-n)!} \right\}$$

$$= \frac{(j-1)!}{(i-1)!} \sum_{n=o}^{d} \left\{ (-1)^{n} \frac{1}{d!} \frac{d!}{n!(d-n)!} \right\}$$

$$= \frac{(j-1)!}{(i-1)!} \sum_{n=o}^{d} \left\{ (-1)^{n} \frac{1}{d!} \mathrm{C}_{d,n} \right\}$$

$$= \frac{(j-1)!}{(i-1)!d!} \sum_{n=o}^{d} \left\{ (-1)^{n} \mathrm{C}_{d,n} \right\}$$

を得るところまできた。さて、ここで二項定理を思い出そう。二項定理とは、自然数 n について、

$$(x+y)^{n} = \sum_{r=0}^{n} \mathrm{C}_{n,r} x^{r} y^{n-r}$$

が常に言えるという定理である。この式に対し、 $x = -1, y = 1$を代入すると、

$$(0) = \sum_{r=0}^{n} \mathrm{C}_{n,r} (-1)^{r}$$

ということがわかるのだけれど、これはつまり、さっきの式が０だということだ。これで、**P**の各要素のマイナス符号をはずしたものが、$\mathbf{P}^{-1}$になっていると証明できた。

後日の成果について

これまでの説明より、

$p_0(n)$　　０種類そろっている確率

も含めて考えたほうが、より一般化して考えることができる。
たとえば、以前、種類数５のとき５×５の正方行列で表現していたけれど、

$p_0(n)$を導入すると、６×６の正方行列で考えることになる。

そして、以下のような一般化的な解を求めることができた。

N 種類のおまけのうち、欲しいものが b 種類ある。そして、すでに a 種類入手しているとする。これから n 個のおかしを買い増ししたとき、手元に揃った種類数が m になる確率 p は、次のようになる。

$$p(N,a,b,m,n) = \sum_{i=b-m}^{b-a} \left\{ (-1)^{b-m+i} \, \mathrm{C}_{i,b-m} \, \mathrm{C}_{b-a,i} \left(\frac{N-i}{N} \right)^n \right\}$$

$$= \mathrm{C}_{b-a,m-a} \sum_{i=b-m}^{b-a} \left\{ (-1)^{b-m+i} \, \mathrm{C}_{m-a,m-b+i} \left(\frac{N-i}{N} \right)^n \right\}$$

ただし$(N > 0, n \geq 0, N \geq b \geq m \geq a \geq 0)$

通常、$m = b$の場合に興味を持つが、その場合に限定すると、

$$\sum_{i=0}^{b-a} \left\{ (-1)^i \, \mathrm{C}_{b-a,i} \left(\frac{N-i}{N} \right)^n \right\}$$

である。この式の(b-a)とは、「あと何種類欲しいか」を意味している。