[東京工業大学 1989 年 $\boxed{5}$]

箱の中に，1 から n までの数字をそれぞれ 1 つずつ書いた n 枚のカードが入っている。箱から無作為に 1 枚のカードを取り出して，その数字を記録し，箱に戻す。この試行を k 回くり返し，それまでに記録された相異なる数字の個数を S_k とする。$S_k=r$ となる確率を $P(S_k=r)$ で表すとき，次の問に答えよ。

(1) $P(S_k=r)$ を $P(S_{k-1}=r)$ と $P(S_{k-1}=r-1)$ で表せ。

(2) S_k の期待値 $E_k=\displaystyle\sum_{r=1}^{k} r\,P(S_k=r)$ を求めよ。

---

(1) $S_k=r$ となるのは,「$k-1$回目までに r 個の相異なる数字が出ていて，k 回目にすでに出ている r 個の数字の中の 1 つが出るとき」または「$k-1$回目までに $r-1$個の相異なる数字が出ていて，k 回目にまだ出ていない $n-(r-1)$ 個の数字の中の 1 つが出るとき」である。

よって　$P(S_k=r)=\dfrac{r}{n}P(S_{k-1}=r)+\dfrac{n-(r-1)}{n}P(S_{k-1}=r-1)$

(2) $E_k=\displaystyle\sum_{r=1}^{k} r\,P(S_k=r)$

$$=\sum_{r=1}^{k} r\left\{\frac{r}{n}P(S_{k-1}=r)+\frac{n-(r-1)}{n}P(S_{k-1}=r-1)\right\}$$

$$=\sum_{r=1}^{k}\frac{r^2}{n}P(S_{k-1}=r)+\sum_{r=1}^{k}\frac{nr-r(r-1)}{n}P(S_{k-1}=r-1)$$

$$=\frac{1}{n}\sum_{r=1}^{k} r^2\,P(S_{k-1}=r)+\sum_{r=1}^{k}\frac{nr-r^2+r}{n}P(S_{k-1}=r-1)$$

$$=\frac{1}{n}\sum_{r=1}^{k} r^2\,P(S_{k-1}=r)+\sum_{r=1}^{k}\frac{nr-(r-1)^2-r+1}{n}P(S_{k-1}=r-1)$$

$$=\frac{1}{n}\sum_{r=1}^{k} r^2\,P(S_{k-1}=r)+\sum_{r=1}^{k}\frac{n(r-1)+n-(r-1)-(r-1)^2}{n}P(S_{k-1}=r-1)$$

$$=\frac{1}{n}\sum_{r=1}^{k} r^2\,P(S_{k-1}=r)+\sum_{r=1}^{k}\frac{(n-1)(r-1)+n-(r-1)^2}{n}P(S_{k-1}=r-1)$$

$$=\frac{1}{n}\sum_{r=1}^{k} r^2\,P(S_{k-1}=r)+\frac{n-1}{n}\sum_{r=1}^{k}(r-1)P(S_{k-1}=r-1)$$
$$+\sum_{r=1}^{k}P(S_{k-1}=r-1)-\frac{1}{n}\sum_{r=1}^{k}(r-1)^2P(S_{k-1}=r-1)$$

$$=\frac{1}{n}\sum_{r=1}^{k} r^2\,P(S_{k-1}=r)+\frac{n-1}{n}\sum_{r=1}^{k}(r-1)P(S_{k-1}=r-1)+1-\frac{1}{n}\sum_{r=1}^{k}(r-1)^2P(S_{k-1}=r-1)$$

ここで

$$\frac{1}{n}\sum_{r=1}^{k} r^2\, P(S_{k-1}=r)=\frac{1}{n}\sum_{r=1}^{k}(r-1)^2 P(S_{k-1}=r-1)\,,\quad \sum_{r=1}^{k}(r-1)P(S_{k-1}=r-1)=E_{k-1}$$

であることから

$$E_k=\frac{n-1}{n}E_{k-1}+1\,,\quad E_1=P(S_1=1)=1$$

である。

この漸化式を解くと

$$E_k=\frac{1}{1-\frac{n-1}{n}}+\left(1-\frac{1}{1-\frac{n-1}{n}}\right)\left(\frac{n-1}{n}\right)^{k-1}$$

$$=n+\left(1-n\right)\left(\frac{n-1}{n}\right)^{k-1}$$

$$=n-n\cdot\frac{n-1}{n}\left(\frac{n-1}{n}\right)^{k-1}$$

$$=n\left\{1-\left(\frac{n-1}{n}\right)^{k}\right\}$$