# Some statistical method with finite memory

By

Kensuke TANAKA and Kōichi INADA

(Received October 1, 1972)

## 1. Introduction and Summary

The two-armed bandit problem with finite memory has been investigated by ROBBINS [10], ISBELL [9], SMITH and PYKE [11], and SAMUELS [12], under a time-invariant decision rule. But a limiting proportion of heads equal to max $\{p_1, p_2\}$ was not achieved by such rules. However if the choice of coin may depend on time, the above fact had been solved by COVER [1]. And some extensions of his rule were given by K. TANAKA and K. INADA [13]. Moreover COVER and HELLMAN [3] solved the two-armed bandit problem under the assumption of finite memory similar to that used by HELLMAN and COVER [7], using a time-invariant decision rule. Note that this problem combines the hypothesis testing problem investigated by COVER [2], and HELLMAN and COVER [7], with a time-invariant learning algorithm.

In this paper we shall describe the method of solving the following statistical problem with finite memory with a limiting probability of error zero, using a time-varying decision rule. That problem is, "*Select a coin with the same propability of coming up head as that of a given coin among many coins on the basis of an outcome of tossing the given coin and one coin among many coins in pairs.*"

This paper consists of three sections. In Section 2 the problem that a coin with the same probability of coming up head as that of a given coin among three coins is selected is solved. Moreover, in Section 3 the provlem that a coin with the same probability of coming up head as that of a given coin among more than four coins is selected is solved.

## 2. Case of three coins

We are given one coin and three coins (coin Ⓐ, coin ①, coin ②, coin ③) with unknown probabilities, $p=1-q$, $p_1=1-q_1$, $p_2=1-q_2$, and $p_3=1-q_3$ of coming up heads. We shall follow the procedure of tossing two coins successively in each test block $T_1$, $T_2$, $\cdots$. Each test block $T$ will be begun arbitrariry with coin ① as the favorite. (This precaution yields independence of the test blocks.) A test block will be broken into $m$ subblocks each consisting of $2s$ tosses. A subblock test will be said to be a success if $2s$ tosses yield an unbroken sequence of $HT$'s or $TH$'s. At the termination of each subblock, the new favorite coin is used to begin the next subblock until $m$ subblock tests have been performed. A test block consists of this collection of subblocks. Thus $ms$ tosses of the coins are made in the test block $T$. Now, we shall state the details of the test subblock.

Let the sequence of coins tossed $(\theta_A, \theta_1, \theta_2, \cdots, \theta_s)$, $\theta_A=$Ⓐ, $\theta_i \epsilon$ {①, ②, ③}, and outcomes observed $(X_A, X_1, X_2, \cdots, X_s)$, $X_A, X_i \epsilon \{H, T\}$, be divided into pairs

$$\begin{pmatrix}\theta_A, & \theta_1\\ X_A, & X_1\end{pmatrix}\begin{pmatrix}\theta_A, & \theta_2\\ X_A, & X_2\end{pmatrix}\cdots\begin{pmatrix}\theta_A, & \theta_s\\ X_A, & X_s\end{pmatrix}.$$

The memory of the past at time $n$ is the state

$$\begin{pmatrix}\theta_{n-1}, & \theta_n\\ X_{n-1}, & X_n\end{pmatrix} \text{ or } \begin{pmatrix}\theta_{n-1}\\ X_{n-1}\end{pmatrix}\begin{pmatrix}\theta_n\\ X_n\end{pmatrix}$$

accordingly as $n$ is even or odd. Thus the memory is of length $r=2$. In figure 1 we shall give the explicit description of this rule in which the details of the orderly transition from the current favorite coin to the new favorite coin are made clear.

coin Ⓐ: $p=1-q$
coin ①: $p_1=1-q_1$, coin ②: $p_2=1-q_2$, coin ③: $p_3=1-q_3$

| | | | |
|---|---|---|---|
| Ⓐ① (*HT*) | Ⓐ① Ⓐ① (*HT*)…(*HT*) | Ⓐ①→Ⓐ② | $\frac{1}{2}(pq_1)^s$ |
| | | ↘ Ⓐ③ | $\frac{1}{2}(pq_1)^s$ |
| (*TH*) | (*TH*)…(*TH*) | Ⓐ①→Ⓐ② | $\frac{1}{2}(qp_1)^s$ |
| | | ↘ Ⓐ③ | $\frac{1}{2}(qp_1)^s$ |
| | otherwise | Ⓐ①→Ⓐ① | $1-\{(pq_1)^s+(qp_1)^s\}$ |
| Ⓐ② (*HT*) | Ⓐ② Ⓐ② (*HT*)…(*HT*) | Ⓐ②→Ⓐ① | $\frac{1}{2}(pq_2)^s$ |
| | | ↘ Ⓐ③ | $\frac{1}{2}(pq_2)^s$ |
| (*TH*) | (*TH*)…(*TH*) | Ⓐ②→Ⓐ① | $\frac{1}{2}(qp_2)^s$ |
| | | ↘ Ⓐ③ | $\frac{1}{2}(qp_2)^s$ |
| | otherwise | Ⓐ②→Ⓐ② | $1-\{(pq_2)^s+(qp_2)^s\}$ |
| Ⓐ③ (*HT*) | Ⓐ③ Ⓐ③ (*HT*)…(*HT*) | Ⓐ③→Ⓐ① | $\frac{1}{2}(pq_3)^s$ |
| | | ↘ Ⓐ② | $\frac{1}{2}(pq_3)^s$ |
| (*TH*) | (*TH*)…(*TH*) | Ⓐ③→Ⓐ① | $\frac{1}{2}(qp_3)^s$ |
| | | ↘ Ⓐ② | $\frac{1}{2}(qp_3)^s$ |
| | otherwise | Ⓐ③→Ⓐ③ | $1-\{(pq_3)^s+(qp_3)^s\}$ |

**Fig. 1.** Coin transition in the test subblock.

Let M be the coin transition probability matrix in which $P_{ij}$ is the transition probability from the current favorite coin ⓘ to the new favorite coin ⓙ $(i, j=1, 2, 3)$.

$$\text{(2. 1)} \qquad M=\begin{pmatrix}P_{11}, & P_{12}, & P_{13}\\ P_{21}, & P_{22}, & P_{23}\\ P_{31}, & P_{32}, & P_{33}\end{pmatrix}$$

$$= \begin{pmatrix} 1-\{(pq_1)^s+(qp_1)^s\}, & \frac{1}{2}\{(pq_1)^s+(qp_1)^s\}, & \frac{1}{2}\{(pq_1)^s+(qp_1)^s\} \\ \frac{1}{2}\{(pq_2)^s+(qp_2)^s\}, & 1-\{(pq_2)^s+(qp_2)^s\}, & \frac{1}{2}\{(pq_2)^s+(qp_2)^s\} \\ \frac{1}{2}\{(pq_3)^s+(qp_3)^s\}, & \frac{1}{2}\{(pq_3)^s+(qp_3)^s\}, & 1-\{(pq_3)^s+(qp_3)^s\} \end{pmatrix}$$

And let $P_i$ be the stationary probability of coin ⓘ being the favorite ($i=1, 2, 3$). In order to calculate $P_i$, we at first consider the following

$$\text{(2. 2)} \quad \begin{aligned} P_1 &= P_1P_{11}+P_2P_{21}+P_3P_{31} \\ P_2 &= P_1P_{12}+P_2P_{22}+P_3P_{32} \\ P_3 &= P_1P_{13}+P_2P_{23}+P_3P_{33} \\ P_1&+P_2+P_3=1. \end{aligned}$$

From (2. 1) and (2. 2) we obtain

$$\text{(2. 3)} \quad \begin{aligned} P_1 =& \{(pq_2)^s+(qp_2)^s\}\{(pq_3)^s+(qp_3)^s\}/[\{(pq_2)^s+(qp_2)^s\}\{(pq_3)^s+(qp_3)^s\} \\ &+\{(pq_1)^s+(qp_1)^s\}\{(pq_3)^s+(qp_3)^s\}+\{(pq_1)^s+(qp_1)^s\}\{(pq_2)^s+(qp_2)^s\}] \\ P_2 =& \{(pq_1)^s+(qp_1)^s\}\{(pq_3)^s+(qp_3)^s\}/[\{(pq_2)^s+(qp_2)^s\}\{(pq_3)^s+(qp_3)^s\} \\ &+\{(pq_1)^s+(qp_1)^s\}\{(pq_3)^s+(qp_3)^s\}+\{(pq_1)^s+(qp_1)^s\}\{(pq_2)^s+(qp_2)^s\}] \\ P_3 =& \{(pq_1)^s+(qp_1)^s\}\{(pq_2)^s+(qp_2)^s\}/[\{(pq_2)^s+(qp_2)^s\}\{(pq_3)^s+(qp_3)^s\} \\ &+\{(pq_1)^s+(qp_1)^s\}\{(pq_3)^s+(qp_3)^s\}+\{(pq_1)^s+(qp_1)^s\}\{(pq_2)^s+(qp_2)^s\}]. \end{aligned}$$

That is, the stationary probabilities of coin ①, coin ② and coin ③ are $P_1$, $P_2$ and $P_3$ in (2. 3) respectively.

Here we shall assume $p_1>p_2>p_3$. And let $P_1{}^i$, $P_2{}^i$ and $P_3{}^i$ be the probabilities of coin ①, coin ② and coin ③ being the favorite coins in the test block $T_i$ respectively. Clearly $P_1{}^i$, $P_2{}^i$ and $P_3{}^i$ depend on $p$, $p_1$, $p_2$, $p_3$, $m_i$ and $s_i$, and approach $P_1$, $P_2$ and $P_3$ respectively as $m_i \to \infty$.

Now we must consider the following cases.

Case 1. If coin Ⓐ coincides with coin① (i.e. $p=p_1$), we hope to hold $\sum P_1{}^i=\infty$, $\sum P_2{}^i<\infty$ and $\sum P_3{}^i<\infty$.

Case 2. If coin Ⓐ coincides with coin ② (i.e. $p=p_2$), we hope to hold $\sum P_1{}^i<\infty$, $\sum P_2{}^i=\infty$ and $\sum P_3{}^i<\infty$.

Case 3. If coin Ⓐ coincides with coin ③ (i.e. $p=p_3$), we hope to hold $\sum P_1{}^i<\infty$, $\sum P_2{}^i<\infty$ and $\sum P_3{}^i=\infty$.

If one of the above three cases is realized, from the Borel zero-one law we may conclude that with probability one only a finite number of test block $T_i$ will result in an incorrect choice of coin as the block tests have been made independent, by letting coin ① be the favorite at the beginning of each test block.

So lastly we must indicate each of the above three cases being realized.

Case 1.

When coin Ⓐ coincides with coin ① (i.e. $p=p_1$), from (2. 3) we obtain

$$(2.\ 4)\qquad \begin{aligned} P_1^i &= A_1/(A_1+A_2+A_3) \\ P_2^i &= A_2/(A_1+A_2+A_3) \\ P_3^i &= A_3/(A_1+A_2+A_3) \end{aligned}$$

where

$$\begin{aligned} A_1 &= \{(p_1q_2)^{s_i}+(q_1p_2)^{s_i}\}\{(p_1q_3)^{s_i}+(q_1p_3)^{s_i}\} \\ A_2 &= 2(p_1q_1)^{s_i}\{(p_1q_3)^{s_i}+(q_1p_3)^{s_i}\} \\ A_3 &= 2(p_1q_1)^{s_i}\{(p_1q_2)^{s_i}+(q_1p_2)^{s_i}\}. \end{aligned}$$

And further we can transform the above equations into the following

$$(2.\ 5)\qquad \begin{aligned} P_1^i &= \frac{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})}{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_1^{s_i}+\beta_1^{s_i})} \\ P_2^i &= \frac{2(\alpha_2^{s_i}+\beta_2^{s_i})}{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_1^{s_i}+\beta_1^{s_i})} \\ P_3^i &= \frac{2(\alpha_1^{s_i}+\beta_1^{s_i})}{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_1^{s_i}+\beta_1^{s_i})} \end{aligned}$$

where

$$\alpha_1=p_2/p_1<1,\ \alpha_2=p_3/p_1<1$$

and

$$\beta_1=q_2/q_1>1,\ \beta_2=q_3/q_1>1.$$

Here we find

$$\begin{aligned} P_1^i &= 1-\frac{2(\alpha_2^{s_i}+\beta_2^{s_i})}{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_1^{s_i}+\beta_1^{s_i})} \\ &\quad -\frac{2(\alpha_1^{s_i}+\beta_1^{s_i})}{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_1^{s_i}+\beta_1^{s_i})} \\ &= 1-P_2^i-P_3^i. \end{aligned}$$

So that, from the above thing, if it is assured that $\sum P_2^i<\infty$ and $\sum P_3^i<\infty$, it follows that $\sum P_1^i=\infty$. For that reason, it is enough to show that $\sum P_2^i<\infty$ and $\sum P_3^i<\infty$. From (2. 5) we obtain

$$P_2^i = \frac{2(\alpha_2^{s_i}+\beta_2^{s_i})}{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_1^{s_i}+\beta_1^{s_i})}$$

$$=\frac{2}{\beta_1^{s_i}}\cdot\frac{1}{1+\alpha_1^{s_i}/\beta_1^{s_i}+2/\beta_1^{s_i}+2(\alpha_1^{s_i}+\beta_1^{s_i})/\beta_1^{s_i}(\alpha_2^{s_i}+\beta_2^{s_i})}<\frac{2}{\beta_1^{s_i}}$$

and

$$P_3^i=\frac{2(\alpha_1^{s_i}+\beta_1^{s_i})}{(\alpha_1^{s_i}+\beta_1^{s_i})(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_2^{s_i}+\beta_2^{s_i})+2(\alpha_1^{s_i}+\beta_1^{s_i})}$$

$$=\frac{2}{\beta_2^{s_i}}\cdot\frac{1}{1+\alpha_2^{s_i}/\beta_2^{s_i}+2/\beta_2^{s_i}+2(\alpha_2^{s_i}+\beta_2^{s_i})/\beta_2^{s_i}(\alpha_1^{s_i}+\beta_1^{s_i})}<\frac{2}{\beta_2^{s_i}}$$

As it is found from (2. 5) that $1/\beta_1<1$ and $1/\beta_2<1$, we have

$$\sum P_2^i<2\sum(1/\beta_1)^{s_i}<\infty$$

and

$$\sum P_3^i<2\sum(1/\beta_2)^{s_i}<\infty,\ \text{e.g.},\ s_i=i.$$

Case 2.

When coin Ⓐ coincides with coin ② (i.e. $p=p_2$), from (2. 3) we obtain

$$P_1^i=B_1/(B_1+B_2+B_3)$$

(2. 6) $$P_2^i=B_2/(B_1+B_2+B_3)$$

$$P_3^i=B_3/(B_1+B_2+B_3)$$

where

$$B_1=2(p_2q_2)^{s_i}\{(p_2q_3)^{s_i}+(q_2p_3)^{s_i}\}$$

$$B_2=\{(p_2q_1)^{s_i}+(q_2p_1)^{s_i}\}\{(p_2q_3)^{s_i}+(q_2p_3)^{s_i}\}$$

$$B_3=2(p_2q_2)^{s_i}\{(p_2q_1)^{s_i}+(q_2p_1)^{s_i}\}$$

And further we can transform the above equations into the following

$$P_1^i=\frac{2\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)}{2\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)+\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)+2\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)}$$

(2. 7) $$P_2^i=\frac{\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)}{2\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)+\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)+2\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)}$$

$$P_3^i=\frac{2\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)}{2\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)+\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)\left(\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\beta_2^{s_i}}{\beta_1^{s_i}}\right)+2\left(\frac{1}{\alpha_1^{s_i}}+\frac{1}{\beta_1^{s_i}}\right)}$$

where

$$\alpha_1=p_2/p_1<1,\quad \alpha_2=p_3/p_1<1$$

and

$$\beta_1 = q_2/q_1 > 1, \ \beta_2 = q_3/q_1 > 1.$$

Here, of course, $P_2{}^i = 1 - P_1{}^i - P_3{}^i$ holds. So that, from the above thing, if it is assured that $\sum P_1{}^i < \infty$ and $\sum P_3{}^i < \infty$, it follows that $\sum P_2{}^i = \infty$. For that reason, it is enough to show that $\sum P_1{}^i < \infty$ and $\sum P_3{}^i < \infty$. From (2. 7) we obtain

$$P_1{}^i = \frac{2}{1/\alpha_1{}^{s_i}} \cdot \frac{1}{2\alpha_1{}^{s_i} + \dfrac{\alpha_1{}^{s_i}}{\beta_1{}^{s_i}} + 1 + 2\alpha_1{}^{s_i} \cdot \dfrac{1/\alpha_1{}^{s_i} + 1/\beta_1{}^{s_i}}{\alpha_2{}^{s_i}/\alpha_1{}^{s_i} + \beta_2{}^{s_i}/\beta_1{}^{s_i}}}$$

$$< 2\alpha_1{}^{s_i}$$

and

$$P_3{}^i = \frac{2}{\beta_2{}^{s_i}/\beta_1{}^{s_i}} \cdot \frac{1}{\dfrac{2\beta_1{}^{s_i}(\alpha_2{}^{s_i}/\alpha_1{}^{s_i} + \beta_2{}^{s_i}/\beta_1{}^{s_i})}{\beta_2{}^{s_i}(1/\alpha_1{}^{s_i} + 1/\beta_1{}^{s_i})} + 1 + \dfrac{\alpha_2{}^{s_i}\beta_1{}^{s_i}}{\alpha_1{}^{s_i}\beta_2{}^{s_i}} + 2\dfrac{\beta_1{}^{s_i}}{\beta_2{}^{s_i}}}$$

$$< 2\frac{\beta_1{}^{s_i}}{\beta_2{}^{s_i}}.$$

As it is found from (2. 7) that $\alpha_1 < 1$ and $\beta_1/\beta_2 < 1$, we have

$$\sum P_1{}^i < 2\sum \alpha_1{}^{s_i} < \infty$$

and

$$\sum P_3{}^i < 2\sum (\beta_1/\beta_2)^{s_i} < \infty, \text{ e.g., } s_i = i.$$

Case 3.

When coin Ⓐ coincides with coin ③ (i.e. $p = p_3$), from (2. 3) we obtain

$$\begin{aligned} P_1{}^i &= C_1/(C_1 + C_2 + C_3) \\ P_2{}^i &= C_2/(C_1 + C_2 + C_3) \qquad (2.\ 8) \\ P_3{}^i &= C_3/(C_1 + C_2 + C_3) \end{aligned}$$

where

$$C_1 = 2(p_3q_3)^{s_i}\{(p_3q_2)^{s_i} + (q_3p_2)^{s_i}\}$$

$$C_2 = 2(p_3q_3)^{s_i}\{(p_3q_1)^{s_i} + (q_3p_1)^{s_i}\}$$

$$C_3 = \{(p_3q_1)^{s_i} + (q_3p_1)^{s_i}\}\{(p_3q_2)^{s_i} + (q_3p_2)^{s_i}\}.$$

And further we can transform the above equations into the following

$$P_1{}^i = \frac{2\left(\dfrac{\alpha_1{}^{s_i}}{\alpha_2{}^{s_i}} + \dfrac{\beta_1{}^{s_i}}{\beta_2{}^{s_i}}\right)}{2\left(\dfrac{\alpha_1{}^{s_i}}{\alpha_2{}^{s_i}} + \dfrac{\beta_1{}^{s_i}}{\beta_2{}^{s_i}}\right) + 2\left(\dfrac{1}{\alpha_2{}^{s_i}} + \dfrac{1}{\beta_2{}^{s_i}}\right) + \left(\dfrac{1}{\alpha_2{}^{s_i}} + \dfrac{1}{\beta_2{}^{s_i}}\right)\left(\dfrac{\alpha_1{}^{s_i}}{\alpha_2{}^{s_i}} + \dfrac{\beta_1{}^{s_i}}{\beta_2{}^{s_i}}\right)}$$

(2. 9)

$$P_2{}^i = \frac{2\left(\dfrac{1}{\alpha_2{}^{s_i}} + \dfrac{1}{\beta_2{}^{s_i}}\right)}{2\left(\dfrac{\alpha_1{}^{s_i}}{\alpha_2{}^{s_i}} + \dfrac{\beta_1{}^{s_i}}{\beta_2{}^{s_i}}\right) + 2\left(\dfrac{1}{\alpha_2{}^{s_i}} + \dfrac{1}{\beta_2{}^{s_i}}\right) + \left(\dfrac{1}{\alpha_2{}^{s_i}} + \dfrac{1}{\beta_2{}^{s_i}}\right)\left(\dfrac{\alpha_1{}^{s_i}}{\alpha_2{}^{s_i}} + \dfrac{\beta_1{}^{s_i}}{\beta_2{}^{s_i}}\right)}$$

$$P_3^i=\frac{\left(\frac{1}{\alpha_2^{s_i}}+\frac{1}{\beta_2^{s_i}}\right)\left(\frac{\alpha_1^{s_i}}{\alpha_2^{s_i}}+\frac{\beta_1^{s_i}}{\beta_2^{s_i}}\right)}{2\left(\frac{\alpha_1^{s_i}}{\alpha_2^{s_i}}+\frac{\beta_1^{s_i}}{\beta_2^{s_i}}\right)+2\left(\frac{1}{\alpha_2^{s_i}}+\frac{1}{\beta_2^{s_i}}\right)+\left(\frac{1}{\alpha_2^{s_i}}+\frac{1}{\beta_2^{s_i}}\right)\left(\frac{\alpha_1^{s_i}}{\alpha_2^{s_i}}+\frac{\beta_1^{s_i}}{\beta_2^{s_i}}\right)}$$

where

$$\alpha_1=p_2/p_1<1,\ \alpha_2=p_3/p_1<1$$

and

$$\beta_1=q_2/q_1>1,\ \beta_2=q_3/q_1>1.$$

Here, of course, $P_3^i=1-P_1^i-P_2^i$ holds. So that, from the above thing, if it is assured that $\sum P_1^i<\infty$ and $\sum P_2^i<\infty$, it follows that $\sum P_3^i=\infty$. For that reason, it is enough to show that $\sum P_1^i<\infty$ and $\sum P_2^i<\infty$. From (2. 9) we obtain

$$P_1^i=\frac{2}{1/\alpha_2^{s_i}}\cdot\frac{1}{2\alpha_2^{s_i}+\frac{2\alpha_2^{s_i}(1/\alpha_2^{s_i}+1/\beta_2^{s_i})}{(\alpha_1^{s_i}/\alpha_2^{s_i}+\beta_1^{s_i}/\beta_2^{s_i})}+\frac{\alpha_2^{s_i}}{\beta_2^{s_i}}+1}$$

$$<2\alpha_2^{s_i}$$

and

$$P_2^i=\frac{2}{\alpha_1^{s_i}/\alpha_2^{s_i}}\cdot\frac{1}{\frac{2\alpha_2^{s_i}/\alpha_1^{s_i}(\alpha_1^{s_i}/\alpha_2^{s_i}+\beta_1^{s_i}/\beta_2^{s_i})}{(1/\alpha_2^{s_i}+1/\beta_2^{s_i})}+2\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}+\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}\frac{\beta_1^{s_i}}{\beta_2^{s_i}}+1}$$

$$<2\frac{\alpha_2^{s_i}}{\alpha_1^{s_i}}$$

As it is found from (2. 9) that $\alpha_2<1$ and $\alpha_2/\alpha_1<1$, we have

$$\sum P_1^i<2\sum\alpha_2^{s_i}<\infty$$

and

$$\sum P_2^i<2\sum(\alpha_2/\alpha_1)^{s_i}<\infty,\text{ e.g., } s_i=i.$$

## 3. Case of $m$ coins (General case)

We are given one coin and $m$ coins (coin Ⓐ, coin ①, coin ②, $\cdots$, coin ⓜ) with unknown probabilities $p=1-q$, $p_1=1-q_1$, $p_2=1-q_2$, $\cdots$, $p_m=1-q_m$, of coming up heads. We shall follow the same procedure as used in Section 2. In Figure 2, we shall give the details of the orderly transition from the current favorite coin to the new favorite coin in the test subblock.

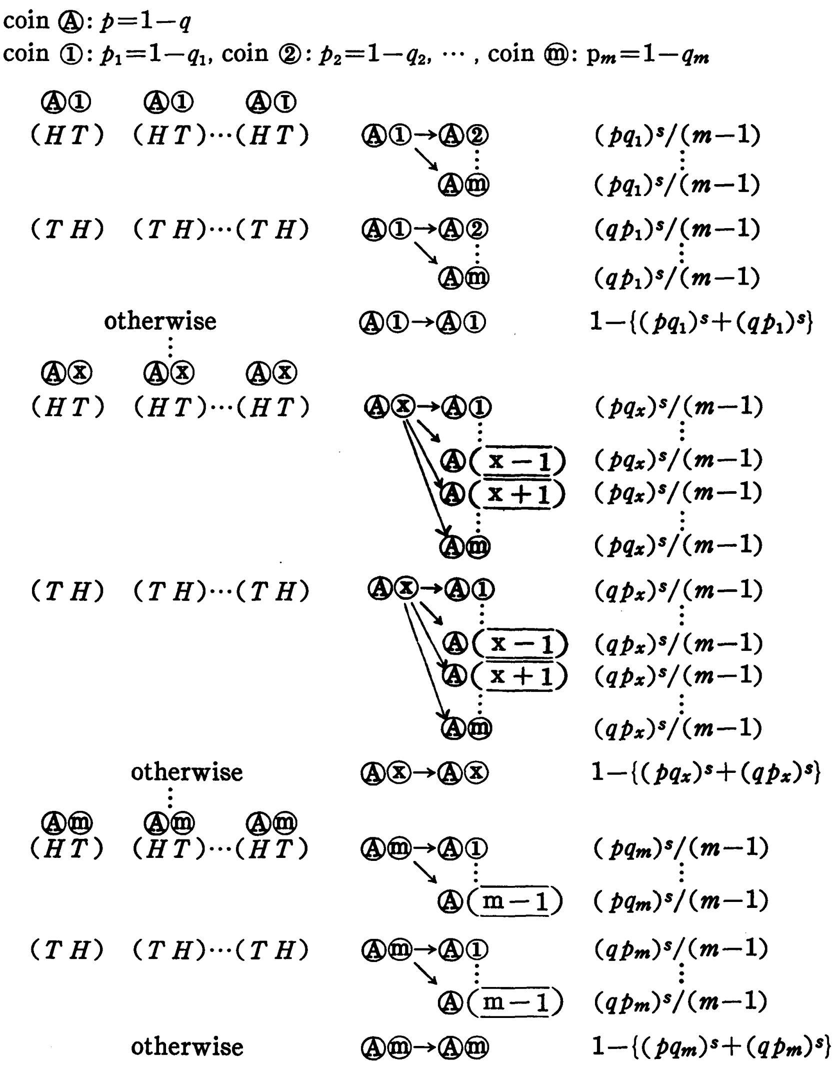


**Fig. 2.** Coin transition in the test subblock.

Let N be the coin transition probability matrix in which $P_{ij}$ is the transition probability from the current favorite coin ⓘ to the new favorite coin ⓙ $(i, j=1, 2, \cdots, m)$.

$$N=\begin{pmatrix} P_{11}\ P_{12}\cdots\cdots\cdots\cdots\ P_{1m} \\ \vdots \qquad\qquad\qquad \vdots \\ P_{i1}\ P_{i2}\cdots\cdots P_{ij}\cdots\cdots P_{im} \\ \vdots \qquad\qquad\qquad \vdots \\ P_{m1}\ P_{m2}\cdots\cdots\cdots\ P_{mm} \end{pmatrix} \tag{3.1}$$

where

$$P_{ij}=\frac{1}{m-1}\{(pq_i)^s+(qp_i)^s\} \quad (i\neq j,\ i, j=1, 2, \cdots, m)$$

and

$$P_{ii}=1-\{(pq_i)^s+(qp_i)^s\} \quad (i=1, 2, \cdots, m)$$

And let $P_i$ be the stationary probability of coin ⓘ being the favorite ($i=1, 2, \cdots, m$). In order to calculate $P_i$, we at first consider the following.

$$(3.2)\qquad \begin{cases} P_i=\sum_{j=1}^{m} P_j P_{ji} & (i=1, 2, \cdots, m) \\ \sum_{i=1}^{m} P_i=1 \end{cases}$$

From (3. 1) and (3. 2), we obtain

$$P_1=\frac{\prod_{k\neq 1}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}$$

$$P_2=\frac{\prod_{k\neq 2}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}$$

$$\vdots$$

$$(3.3)\qquad P_t=\frac{\prod_{k\neq t}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}$$

$$\vdots$$

$$P_m=\frac{\prod_{k\neq m}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(pq_k)^{s_i}+(qp_k)^{s_i}\}}$$

That is, the stationary probabilities of coin ①, coin ②, $\cdots$, coin (m−1) and coin ⓜ are $P_1$, $P_2, \cdots, P_{m-1}$ and $P_m$ in (3. 3) respectively.

Here we shall assume $p_1>p_2>\cdots>p_m$. And let $P_1{}^i, P_2{}^i, \cdots, P_{m-1}{}^i$, and $P_m{}^i$ be the probabilities of coin ①, coin ②, $\cdots$, coin (m−1) and coin ⓜ being the favorite coins in the test block $T_i$ respectively. Clearly $P_1{}^i, P_2{}^i, \cdots, P_{m-1}{}^i$, and $P_m{}^i$ depend on $p, p_1, p_2, \cdots, p_m, m_i$ and $s_i$, and approach $P_1, P_2, \cdots, P_{m-1}$, and $P_m$ respectively as $m_i\to\infty$.

Now we must consider the following $m$ cases.

Case x. If coin Ⓐ coincides with coin ⓧ (i.e. $p=p_x$), we hope to hold $\sum P_x{}^i=\infty$ and $\sum P_t{}^i<\infty$ ($t\neq x$, $t=1, 2, \cdots, m$, $x=1, 2, \cdots, m$)

If one of the above $m$ cases is realized, from the Borel zero-one law we may conclude that with probability one only a finite number of test block $T_i$ will result in an incorrect choice of coins as the block tests have been made independent, by letting coin ① be the favorite at the beginning of each test block.

So lastly we must indicate each of the above $m$ cases being realized.

Case x.

When coin Ⓐ coincides with coin ⓧ (i.e. $p=p_x$), from (3. 3) we obtain

$$P_1{}^i=\frac{\prod_{k\neq 1}^{m}\{(p_xq_k)^{s_i}+(q_xp_k)^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(p_xq_k)^{s_i}+(q_xp_k)^{s_i}\}}$$

$$\vdots$$

(3. 4)
$$P_t{}^i=\frac{\prod_{k\neq t}^{m}\{(p_xq_k)^{s_i}+(q_xp_k)^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(p_xq_k)^{s_i}+(q_xp_k)^{s_i}\}}$$

$$\vdots$$

$$P_m{}^i=\frac{\prod_{k\neq m}^{m}\{(p_xq_k)^{s_i}+(q_xp_k)^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(p_xq_k)^{s_i}+(q_xp_k)^{s_i}\}}$$

And further we can transform the above equations into the following

$$P_1{}^i=\frac{\prod_{k\neq 1}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}$$

$$\vdots$$

(3. 5)
$$P_t{}^i=\frac{\prod_{k\neq t}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}$$

$$\vdots$$

$$P_m{}^i=\frac{\prod_{k\neq m}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\sum_{j=1}^{m}\prod_{k\neq j}^{m}\{\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}$$

where

$$\alpha_0=p_1/p_1=1,\ \alpha_1=p_2/p_1,\ \cdots,\ \alpha_t=p_t/p_1,\ \cdots,\ \alpha_{m-1}=p_m/p_1$$

and

$$\beta_0=q_1/q_1=1,\ \beta_1=q_2/q_1,\cdots,\ \beta_t=q_t/q_1,\cdots,\ \beta_{m-1}=q_m/q_1.$$

Here, of course, $P_x^i=1-(P_1^i+P_2^i+\cdots+P_{x-1}^i+P_{x+1}^i+\cdots+P^i_m)$ holds. So that, from the above thing, if it is assured that $\sum P_1^i<\infty,\cdots,\ \sum P_{x-1}^i<\infty,\ \sum P_{x+1}^i<\infty,\cdots$, and $\sum P_m^i<\infty$, it follows that $\sum P_x^i=\infty$. For that reason, it is enough to show that $\sum P_1^i<\infty,\cdots$, $\sum P_{x-1}^i<\infty,\ \sum P_{x+1}^i<\infty,\cdots$, and $\sum P_m^i<\infty$. From (3. 5) we obtain

$$P_1^i=\frac{1}{\dfrac{\prod_{k\neq x}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq 1}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}+\dfrac{\sum_{j=1,\,\neq x}^{m}\prod_{k\neq j}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq 1}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}}$$

$$<2(\alpha_{x-1}/\alpha_0)^{s_i}$$

$$\vdots$$

$$P_{x-1}^i=\frac{1}{\dfrac{\prod_{k\neq x}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq x-1}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}+\dfrac{\sum_{j=1,\,\neq x}^{m}\prod_{k\neq j}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq x-1}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}}$$

$$<2(\alpha_{x-1}/\alpha_{x-2})^{s_i}$$

$$P_{x+1}^i=\frac{1}{\dfrac{\prod_{k\neq x}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq x+1}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}+\dfrac{\sum_{j=1,\,\neq x}^{m}\prod_{k\neq j}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq x+1}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}}$$

$$<2(\beta_{x-1}/\beta_x)^{s_i}$$

$$\vdots$$

$$P_m^i=\frac{1}{\dfrac{\prod_{k\neq x}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq m}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}+\dfrac{\sum_{j=1,\,\neq x}^{m}\prod_{k\neq j}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}{\prod_{k\neq m}^{m}\{(\alpha_{k-1}/\alpha_{x-1})^{s_i}+(\beta_{k-1}/\beta_{x-1})^{s_i}\}}}$$

$$<2(\beta_{x-1}/\beta_{m-1})^{s_i}.$$

As it is found from (3. 5) that $\alpha_{x-1}/\alpha_0<1,\ \alpha_{x-1}/\alpha_1<1,\cdots,\ \alpha_{x-1}/\alpha_{x-2}<1,\ \beta_{x-1}/\beta_x<1,\cdots$, and $\beta_{x-1}/\beta_{m-1}<1$, we have the following equations

$$\sum P_1^i < 2\sum(\alpha_{x-1}/\alpha_0)^{s_i}$$
$$\vdots$$
$$\sum P_{x-1}^i < 2\sum(\alpha_{x-1}/\alpha_{x-2})^{s_i}$$
$$\sum P_{x+1}^i < 2\sum(\beta_{x-1}/\beta_x)^{s_i}$$
$$\vdots$$
$$\sum P_m^i < 2\sum(\beta_{x-1}/\beta_{m-1})^{s_i}, \text{ e.g., } s_i = i.$$

NIIGATA UNIVERSITY

## References

[1] COVER, T. M.: *A note on the two-armed bandit problem with finite memory.* Information and Control, Vol. 12, (1968), 371–377.
[2] COVER, T. M.: *Hypothesis testing with finite statistics.* Ann. Math. Stat., Vol. 40, (1969), 828–835.
[3] COVER, T. M. and HELLMAN, M. E.: *The two-armed bandit problem with time-invariant finite memory.* IEEE Trans. Information Theory, Vol. IT-16, (1970), 185–195.
[4] FU, K. S.: *Stochastic automata as models of learning systems.* Computer and Information Sciences-, Academic Press, New York, (1967), 177–191.
[5] FU, K. S.: Sequential methods in pattern recognition and machine learning. Academic Press, New York and London, (1968).
[6] FU, K. S. and Mendel. J. M.: Adaptive, learning, and pattern recognition systems theory and applications. Academic Press, New York and London, (1970).
[7] HELLMAN, M. E. and COVER, T. M.: *Learning with finite memory.* Ann. Math. Stat., Vol. 41, (1970), 765–782.
[8] HELLMAN, M. E. and COVER, T. M.: *On memory saved by randomization.* Ann. Math. Stat., Vol. 42, (1971). 1075–1078.
[9] ISBELL, J. R.: *On a problem of Robbins.* Ann. Math. Stat., Vol. 30, (1959), 606–610.
[10] ROBBINS, H.: *A sequential decision problem with a finite memory.* Proc. Nat. Acad. Sci. Vol. 42, (1956), 920–923.
[11] SMITH, C. V. and PYKE, R.: *The Robbins-Isbell two-armed bandit problem with finite memory.* Ann. Math. Stat., Vol. 36, (1965), 1375–1386.
[12] SAMUELS, S, M.: *Randomized rules for the two-armed bandit problem with finite memory.* Ann. Math. Stat., Vol. 39, (1968), 2103–2107.
[13] TANAKA, K. and INADA, K.: *Some extension of the two-armed bandit problem with finite memory.* Sci. Rep. Niigata Univ., Ser. A, No. 10, (1973), 5–13.
[14] TSETLIN, M. L.: *On the behavior of finite automata in random media.* Autom. and Remote Control, Vol. 22, No. 10, (1961), 1210–1219.
[15] VARSHAVSKII, V. I. and VORONTSOVA, I. P.: *On the behavior of stochastic automata with variable structure.* Autom. and Remote Control, Vol. 24, No. 3, (1963), 327–333.