Statistica Sinica **9**(1999), 879-891

ON THE PROBABILITY OF CORRECT SELECTION IN THE LEVIN-ROBBINS SEQUENTIAL ELIMINATION PROCEDURE

Cheng-Shiun Leu*† and Bruce Levin†

**New York State Psychiatric Institute and †Columbia University*

Abstract: We prove the general validity of a formula conjectured by Levin and Robbins (1981) to give a lower bound for the probability of correct selection in a sequential elimination procedure designed to identify the best of c binomial populations. The formula is elementary to calculate and simplifies the design of the selection procedure.

Key words and phrases: Lower bound formula, probability of correct selection, sequential elimination procedure.

1. Introduction

Suppose we have $c \geq 2$ coins, and for any coin i $(1 \leq i \leq c)$, let p_i be the probability of coming up heads on a single toss. We wish to select a coin with the highest such probability. Levin and Robbins (1981) introduced the following sequential elimination procedure to accomplish that goal. Begin by tossing the coins vector-at-a-time. For $n = 1, 2, \ldots$, let $\mathbf{X}^{(n)} = (X_1^{(n)}, \ldots, X_c^{(n)})$ be the vector that reports the number of heads observed for each coin after n tosses, and let $\mathbf{X}^{[n]} = (X_1^{[n]}, \ldots, X_c^{[n]})$ be the ordered $\mathbf{X}^{(n)}$ vector with $X_1^{[n]} \geq X_2^{[n]} \geq \cdots \geq X_c^{[n]}$. Let r be a positive integer chosen in advance of all tosses. Define C to be the set of coins under consideration at any given time, with the convention that $c = \#(C)$. Define $N_r^{(C)}$ to be the *time of first elimination in a c coin game with coins C*,

$$N_r^{(C)} = \inf\{n \geq 1 : \max_{i,j \in C}\{X_i^{(n)} - X_j^{(n)}\} = X_1^{[n]} - X_c^{[n]} = r\}.$$

If $N_r^{(C)} = n$, we drop from further consideration, after toss n, any and all coins i satisfying $X_i^{(n)} = X_c^{[n]}$, i.e. all coins that have fallen r heads behind the leader. If more than one coin remains the procedure continues, starting from the current tallies of the remaining subset of coins $C' \subset C$, and iterates with $N_r^{(C')}$ until $c - 1$ coins have been eliminated. Thereupon we declare the remaining coin as "best". Define $P_r^{(C)}[i]$ to be the probability that coin i is selected as best by this procedure. (We also call this an "*e-game*" [e- for elimination] as if in a race that coin i wins with probability $P_r^{(C)}[i]$.) The following inequalities were conjectured by Levin and Robbins:

Conjecture. For any set of coins C with probabilities $\{p_1, \ldots, p_c\}$, let $w_i = p_i/(1-p_i)$, and suppose, without loss of generality, that $p_1 \geq p_2 \geq \cdots \geq p_c$. Then for any positive integer r,

$$P_r^{(C)}[1] \geq w_1^r / \sum_{i=1}^{c} w_i^r \quad \textit{and} \quad P_r^{(C)}[c] \leq w_c^r / \sum_{i=1}^{c} w_i^r. \tag{1}$$

Inequality (1) allows determination of r to achieve any desired probability of correct selection $P^\dagger$, given any specified odds $w_1, \ldots, w_c$. In particular, if the odds ratio $w_1/w_2 \geq \delta$, it suffices to choose $r \geq \{\log[(c-1)P^\dagger/(1-P^\dagger)]\}/\log\delta$. In the context of medical trials, it is obviously desirable to eliminate the inferior treatments while maintaining a high probability of correctly selecting the best treatment.

The conjectured inequalities in (1) arose because of certain properties Levin and Robbins (1981) proved for a related sequential procedure without elimination of inferior coins. The stopping rule there was

$$M_r^{(C)} = \inf\{n \geq 1 : X_1^{[n]} - X_2^{[n]} = r\},$$

with the obvious selection of the coin corresponding to $X_1^{[n]}$ at time $M_r^{(C)} = n$ as best. Let $P_r^*[i]$ denote the probability of selecting coin i as best with stopping rule $M_r^{(C)}$. Levin and Robbins proved that (1) holds for $P_r^*[i]$ and that for $i < j$,

$$P_r^*[i]/P_r^*[j] \geq (w_i/w_j)^r. \tag{2}$$

They conjectured that inequalities (2) and (1) should hold true for the e-game as well, on the basis of a rigorous proof for $r = 1$ together with simulation studies for "least favorable" parameter configurations of the form $p_1 > p_2 = \cdots = p_c$.

It turns out that the conjecture concerning inequality (2) is not generally true. Zybert and Levin (1987) proved that the conjecture does hold for $c = 3$ coins for least favorable configurations, but, surprisingly, there exist $p_1 > p_2 > p_3$ for which it does not hold. The violations of the inequalities in (2) for the e-game are numerically small, and the expressions $(w_i/w_j)^r$ are quite good approximations to $P_r^*[i]/P_r^*[j]$. In fact, Zybert and Levin showed that, notwithstanding the failure of (2) in the e-game, (1) with $c = 3$ coins *does* hold for any r and any $p_1 \geq p_2 \geq p_3$. This curious situation raises the following questions: if (2) is not the general reason that (1) is true for $c = 3$ in the e-game, does (1) even hold for $c > 3$, or is the case $c = 3$ somehow special? If (1) does hold for any c, what is the fundamental reason?

In this paper we prove that conjecture (1) does hold generally for any number of coins $c \geq 2$ and any positive integer r. The idea of the proof is to demonstrate a somewhat stronger result than (1): the conditional probability that coin 1 (respectively coin c) wins given that the sample path $(X_1^{(n)}, \ldots, X_c^{(n)})$ passes through any symmetric subset of configurations (those left invariant under permutations of the labels of the coins) still obeys the lower (respectively upper) bound inequalities. Inequality (1) follows simply by choosing the subset containing the single configuration $(0, \ldots, 0)$. It is of interest to note that the proof makes no use of the notion of least favorable configurations. This is similar to other proofs utilizing fundamental symmetries (see Levin (1984)).

2. Definitions and Notations

In the Levin-Robbins elimination procedure, it is important to consider differences of the form $X_j^{(n)} - X_c^{[n]}$ $(j = 1, \ldots, c)$, and we define the vector of such differences as the *configuration* of $\mathbf{X}^{(n)}$. With a slight abuse of notation, we write the configuration of $\mathbf{X}^{(n)}$ as $\mathbf{X}^{(n)} - X_c^{[n]} = (X_1^{(n)} - X_c^{[n]}, \ldots, X_c^{(n)} - X_c^{[n]})$. The *ordered configuration* of $\mathbf{X}^{(n)}$ is the configuration of $\mathbf{X}^{[n]}$, i.e. $\mathbf{X}^{[n]} - X_c^{[n]}$. For example, with $c = 3$, the configuration (120) indicates coin 1 has one fewer head than coin 2, and one more than coin 3. The ordered configuration is (210).

In a c coin game, we define for $s = 0, \ldots, r$ the sets

$$B_s^{(C)} = \{\mathbf{b} = (b_1 b_2 \cdots b_c) \mid s = b_1 \geq b_2 \geq \cdots \geq b_c = 0\}.$$

$B_r^{(C)}$ is the set of all possible ordered configurations of $\mathbf{X}^{(n)}$ at the time of first elimination, and the union of $B_s^{(C)}$ for $s < r$ comprise all ordered configurations prior to that time. Before the time of first elimination, the coin tallies pass through a sequence of configurations which are certain permutations of certain ordered configurations in $\bigcup_{s=0}^{r-1} B_s^{(C)}$. This observation suggests that we analyze $P_r^{(C)}[1]$ in terms of conditional probabilities given that $\mathbf{X}^{[n]} - X_c^{[n]}$ passes through any one of the ordered configurations $\mathbf{b} \in B_s^{(C)}$ for some $0 < s < r$. For example, with $r = 3$, $c = 3$, and $\mathbf{b} = (210) \in B_2^{(C)}$, we might consider the conditional probability of selecting coin 1 given that one of the configurations $\{(210), (201), (120), (102), (021), (012)\}$ is reached before the time of first elimination. Also, define $S_{\mathbf{v}}^{(C)} = \{$ all distinct permutations of $\mathbf{v}\}$ for any configuration $\mathbf{v}$. The above set is $S_{(210)}^{(C)}$, while $S_{(110)}^{(C)} = \{(110), (101), (011)\}$.

Now let $N_s^{(C)}$ be the first time any configuration with maximum component $s > 0$ is reached by $\mathbf{X}^{(n)}$ in a c coin game. In symbols,

$$N_s^{(C)} = \inf\{n \geq 1 : X_1^{[n]} - X_c^{[n]} = s\}.$$

Also, for $\mathbf{b} \in B_s^{(C)}$ and $\mathbf{d} \in S_{\mathbf{b}}$, let $N_s^{(C)}(\mathbf{d})$ be the first time a given configuration $\mathbf{d}$ is reached by $\mathbf{X}^{(n)}$ (if ever; if $\mathbf{d}$ is never reached, set $N_s^{(C)}(\mathbf{d}) = \infty$):

$$N_s^{(C)}(\mathbf{d}) = \inf\{n \geq 1 : \mathbf{X}^{(n)} - X_c^{[n]} = \mathbf{d}\}.$$

We define the stopping times $N_0^{(C)}$ and $N_0^{(C)}(\mathbf{0})$ to be identically zero. The event $[N_s^{(C)} = N_s^{(C)}(\mathbf{d})]$ for $0 < s < r$ thus consists of all sequences of tosses in which $\mathbf{X}^{(n)}$ reaches configuration $\mathbf{d}$ before any other configuration $\mathbf{d}'$ with $\max_i d_i' = s$. For $s = 0$, $[N_0^{(C)} = N_0^{(C)}(\mathbf{0})]$ is the entire sample space. We note that events of the form $[N_s^{(C)} = N_s^{(C)}(\mathbf{d})]$ and $[N_s^{(C)} = N_s^{(C)}(\mathbf{d}) < N_r^{(C)}]$ are equivalent for $0 \leq s < r$, i.e., stopping time $N_s^{(C)}$ must occur prior to the time of first elimination $N_r^{(C)}$. For any $\mathbf{b}$ in $B_s^{(C)}$, we define $W_{\mathbf{d}}^{(C)}$ to be the conditional probability that permutation $\mathbf{d}$ of $\mathbf{b}$ is the first to be reached, given that some permutation of $\mathbf{b}$ is reached at time $N_s^{(C)}$,

$$W_{\mathbf{d}}^{(C)} = P[N_s^{(C)} = N_s^{(C)}(\mathbf{d})] \Bigg/ \sum_{\mathbf{d}' \in S_{\mathbf{b}}^{(C)}} P[N_s^{(C)} = N_s^{(C)}(\mathbf{d}')]$$

For any configuration $\mathbf{b}$ and integers i and j in $\{1, \ldots, c\}$, let $\mathbf{b}_{ij}$ be the configuration that interchanges components i and j of $\mathbf{b}$. Define $w_{ij} = w_i/w_j = [p_i/(1-p_i)]/[p_j/(1-p_j)]$, the odds ratio for coin i and j.

3. Proof of the conjecture

The following lemma and its corollaries apply Wald's change of measure argument to events that occur at time $N_s^{(C)}$.

Lemma 1. *For any number of coins c, any $\mathbf{b} \in B_s^{(C)}$ and any $i, j \in \{1, \ldots, c\}$*

$$P[N_s^{(C)} = N_s^{(C)}(\mathbf{b})] = w_{ij}^{b_i - b_j} P[N_s^{(C)} = N_s^{(C)}(\mathbf{b_{ij}})].$$

Proof. By definition, $P[N_s^{(C)} = N_s^{(C)}(\mathbf{b})] = \sum_{n \geq 1} \sum_\alpha P^{(n)}(\alpha)$, where $P^{(n)}(\alpha) = \prod_{k=1}^c w_k^{X_k^{(n)}(\alpha)}(1-p_k)^n$ is the product-binomial probability function of the first n tosses, and α represents any sequence of binary outcome vectors in the event $[N_s^{(C)} = N_s^{(C)}(\mathbf{b}) = n]$. For any such α,

$$\begin{aligned} &P^{(n)}(\alpha) \\ &= w_i^{[X_i^{(n)}(\alpha) - X_j^{(n)}(\alpha)]} w_j^{[X_j^{(n)}(\alpha) - X_i^{(n)}(\alpha)]} (1-p_i)^n (1-p_j)^n w_j^{X_i^{(n)}(\alpha)} w_i^{X_j^{(n)}(\alpha)} \\ &\quad \cdot \prod_{k \neq i,j} w_k^{X_k^{(n)}(\alpha)} (1-p_k)^n \\ &= w_{ij}^{[X_i^{(n)}(\alpha) - X_j^{(n)}(\alpha)]} P^{(n)}(\alpha_{ij}), \end{aligned}$$

where α_{ij} is a sequence of outcomes formed by transposing the outcomes of coins i and j in the sequence $\alpha \in [N_s^{(C)} = N_s^{(C)}(\mathbf{b}) = n]$. Note that, on this event, $X_i^{(n)}(\alpha) - X_j^{(n)}(\alpha) = b_i - b_j$, and there is a one-one correspondence between sequences in $[N_s^{(C)} = N_s^{(C)}(\mathbf{b}) = n]$ and those in $[N_s^{(C)} = N_s^{(C)}(\mathbf{b_{ij}}) = n]$. Therefore

$$\begin{aligned}
P[N_s^{(C)} &= N_s^{(C)}(\mathbf{b}) = n]/w_{ij}^{b_i-b_j} \\
&= \sum_{n\geq 1} \sum_{\alpha\in[N_s^{(C)}=N_s^{(C)}(\mathbf{b})=n]} P^{(n)}(\alpha_{ij}) \\
&= \sum_{n\geq 1} \sum_{\alpha'\in[N_s^{(C)}=N_s^{(C)}(\mathbf{b_{ij}})=n]} P^{(n)}(\alpha') = P[N_s^{(C)} = N_s^{(C)}(\mathbf{b_{ij}})].
\end{aligned}$$

In particular, Lemma 1 holds for configurations that occur at time of first elimination, which we state as

Corollary 1. *For any number of coins c, any configuration $\mathbf{a} \in B_r^{(C)}$ at first elimination, and $i, j \in \{1, \ldots, c\}$,*

$$P[N_r^{(C)} = N_r^{(C)}(\mathbf{a})] = w_{ij}^{a_i-a_j} P[N_r^{(C)} = N_r^{(C)}(\mathbf{a_{ij}})].$$

Because any permutation $\mathbf{d}$ of $\mathbf{b}$ can be represented as a sequence of transpositions, by using Lemma 1, we can express $P[N_s^{(C)} = N_s^{(C)}(\mathbf{d})]$ as $P[N_s^{(C)} = N_s^{(C)}(\mathbf{b})]$ times a *permutation constant*, defined as the product of the odds ratios of the coins used in the transpositions raised to the power of the corresponding differences between the number of heads. The permutation constant is uniquely defined, because if λ and λ' are two such permutation constants corresponding to different sequences of transpositions leading to the same permutation $\mathbf{d}$ of $\mathbf{b}$, then $P[N_s^{(C)} = N_s^{(C)}(\mathbf{d})] = \lambda P[N_s^{(C)} = N_s^{(C)}(\mathbf{b})] = \lambda' P[N_s^{(C)} = N_s^{(C)}(\mathbf{b})]$ implies $\lambda = \lambda'$. Therefore, a unique constant $K_{\mathbf{b}}^{(C)}$ can be defined such that

$$W_{\mathbf{b_{ij}}}^{(C)} = P[N_s^{(C)} = N_s^{(C)}(\mathbf{b_{ij}})] \Big/ \sum_{\mathbf{d}'\in S_{\mathbf{b}}^{(C)}} P[N_s^{(C)} = N_s^{(C)}(\mathbf{d}')] = w_{ij}^{-(b_i-b_j)} \Big/ K_{\mathbf{b}}^{(C)},$$

where $K_{\mathbf{b}}^{(C)}$ is the summation of all appropriate permutation constants. For example, in the illustration with $(r = 3, c = 3)$ mentioned above,

$$\begin{aligned}
&P[N_2^{(C)} = N_2^{(C)}((120))] = w_{12}^{1-2} \times P[N_2^{(C)} = N_2^{(C)}((210))], \\
&P[N_2^{(C)} = N_2^{(C)}((012))] = w_{13}^{0-2} \times P[N_2^{(C)} = N_2^{(C)}((210))], \\
&P[N_2^{(C)} = N_2^{(C)}((201))] = w_{23}^{0-1} \times P[N_2^{(C)} = N_2^{(C)}((210))], \ \ldots \\
&P[N_2^{(C)} = N_2^{(C)}((102))] = w_{23}^{0-1} w_{13}^{1-2} \times P[N_2^{(C)} = N_2^{(C)}((210))],
\end{aligned}$$

and therefore

$$W^{(C)}_{(120)} = P[N^{(C)}_2 = N^{(C)}_2((120))] \Big/ \sum_{\mathbf{d}'\in S^{(C)}_{(210)}} P[N^{(C)}_2 = N^{(C)}_2(\mathbf{d}')]$$
$$= \frac{w^{-1}_{12} \times P[N^{(C)}_2 = N^{(C)}_2((210))]}{(1 + w^{-1}_{12} + w^{-2}_{13} + w^{-1}_{23} + w^{-1}_{23}w^{-2}_{12} + w^{-1}_{23}w^{-1}_{13}) \times P[N^{(C)}_2 = N^{(C)}_2((210))]}$$
$$= w^{-1}_{12}/(1 + w^{-1}_{12} + w^{-2}_{13} + w^{-1}_{23} + w^{-1}_{23}w^{-2}_{12} + w^{-1}_{23}w^{-1}_{13}) = w^{-1}_{12}/K^{(C)}_{(210)}.$$

That is, $K^{(C)}_{(210)}$, in this case, equals the unique sum of permutation constants $1 + w^{-1}_{12} + w^{-2}_{13} + w^{-1}_{23} + w^{-1}_{23}w^{-2}_{12} + w^{-1}_{23}w^{-1}_{13}$. In the last term above, if instead of $(210) \to (201) \to (102)$ we use $(210) \to (120) \to (102)$, the permutation constant $w^{-1}_{12}w^{-2}_{23}$ equals $w^{-1}_{23}w^{-1}_{13}$. We can apply the same argument to arbitrary sequences of transpositions, and get the following from Lemma 1.

Corollary 2. $P[N^{(C)}_s = N^{(C)}_s(\mathbf{d})] = K^{(C)}_b W^{(C)}_{\mathbf{d}} P[N^{(C)}_s = N^{(C)}_s(\mathbf{b})]$.

Now let $\mathbf{a} \in B^{(C)}_r$ be an ordered configuration at first elimination, and let $\mathbf{u} \in S_{\mathbf{a}}$. The same change of measure argument as given in Lemma 1 shows that $P[N^{(C)}_r = N^{(C)}_r(\mathbf{u})$ and $N^{(C)}_s = N^{(C)}_s(\mathbf{d})] = w^{u_i-u_j}_{ij} \times P[N^{(C)}_r = N^{(C)}_r(\mathbf{u_{ij}})$ and $N^{(C)}_s = N^{(C)}_s(\mathbf{d_{ij}})]$. Then Lemma 1 implies.

Corollary 3. *For any* $\mathbf{a} \in B^{(C)}_r$, $\mathbf{u} \in S^{(C)}_{\mathbf{a}}$, $\mathbf{b} \in B^{(C)}_s$ *for* $0 \le s < r$, *and* $\mathbf{d} \in S^{(C)}_{\mathbf{b}}$, $P[N^{(C)}_r = N^{(C)}_r(\mathbf{u})|N^{(C)}_s = N^{(C)}_s(\mathbf{d})] = w^{(u_i-u_j)-(d_i-d_j)}_{ij} \cdot P[N^{(C)}_r = N^{(C)}_r(\mathbf{u_{ij}})|N^{(C)}_s = N^{(C)}_s(\mathbf{d_{ij}})]$.

Since $w^{u_i-u_j}_{ij} = W^{(C)}_{\mathbf{u}}/W^{(C)}_{\mathbf{u}_{ij}}$, and $w^{d_i-d_j}_{ij} = W^{(C)}_{\mathbf{d}}/W^{(C)}_{\mathbf{d}_{ij}}$, if now $\mathbf{u}'$ is the *same* permutation of $\mathbf{u}$ as $\mathbf{d}'$ is of $\mathbf{d}$, Corollary 3 could be written generally as follows.

Corollary 4. $(W^{(C)}_{\mathbf{d}}/W^{(C)}_{\mathbf{u}})P[N^{(C)}_r = N^{(C)}_r(\mathbf{u})|N^{(C)}_s = N^{(C)}_s(\mathbf{d})] = (W^{(C)}_{\mathbf{d}'}/W^{(C)}_{\mathbf{u}'})$ $P[N^{(C)}_r = N^{(C)}_r(\mathbf{u}')|N^{(C)}_s = N^{(C)}_s(\mathbf{d}')]$.

We can now prove the main result. The case $s = 0$ yields the conjectured inequalities (1).

Theorem 1. *For any set of* $c \ge 2$ *coins* C *with corresponding probabilities* $p_1 \ge \cdots \ge p_c$, *any integers* s *and* r *such that* $0 \le s \le r$, *and any* $\mathbf{b} \in B^{(C)}_s$,

$$P^{(C)}_r[1|N^{(C)}_s = N^{(C)}_s(\mathbf{d}) \textit{ for some } \mathbf{d} \in S^{(C)}_{\mathbf{b}}]$$
$$= \sum_{\mathbf{d}\in S^{(C)}_{\mathbf{b}}} W^{(C)}_{\mathbf{d}} P^{(C)}_r[1|N^{(C)}_s = N^{(C)}_s(\mathbf{d})] \ge w^r_1 / \sum_{i=1}^{c} w^r_i$$

and

$$P_r^{(C)}[c|N_s^{(C)} = N_s^{(C)}(\mathbf{d}) \text{ for some } \mathbf{d} \in S_{\mathbf{b}}^{(C)}]$$
$$= \sum_{\mathbf{d}\in S_{\mathbf{b}}^{(C)}} W_{\mathbf{d}}^{(C)} P_r^{(C)}[c|N_s^{(C)} = N_s^{(C)}(\mathbf{d})] \le w_c^r / \sum_{i=1}^{c} w_i^r.$$

Proof. We use mathematical induction on the number of coins c. Let $c = 2$ and consider any $C = \{i, j\}$ for some coin j competing with a better coin i $(i < j)$. For $s = 0$, the classical gambler's ruin problem (Feller (1957)) states $P_r^{(C)}[i|N_0^{(C)} = N_0^{(C)}(\mathbf{0})] = P_r^{(C)}[i] = w_i^r/(w_i^r + w_j^r)$ and $P_r^{(C)}[j] = w_j^r/(w_i^r + w_j^r)$. For $0 < s \le r$, the only $\mathbf{b} \in B_s^{(C)}$ is of the form $(s0)$, and it can be seen that any sample path leading to a correct selection *must* pass through $(s0)$ or $(0s)$ for any $s = 0, \ldots, r-1$. Therefore $P_r^{(C)}[i|N_s^{(C)} = N_s^{(C)}(\mathbf{d})$ for $\mathbf{d} = (s0)$ or $(0s)] = P_r^{(C)}[i] = w_i^r/(w_i^r + w_j^r)$, and $P_r^{(C)}[j] = 1 - P_r^{(C)}[i] = w_j^r/(w_i^r + w_j^r)$, the classical gambler's ruin probability, and conclude Theorem 1 is true when $c = 2$. Assume, then, that for any $k \in \{2, \ldots c-1\}$, any subset K of C with corresponding probabilities $p_{i_1} \ge p_{i_2} \ge \cdots \ge p_{i_k}$, any $s = 0, \ldots, r$, and any $\mathbf{b} \in B_s^{(K)}$, $\sum_{\mathbf{d}\in S_{\mathbf{b}}^{(K)}} W_{\mathbf{d}}^{(K)}\ P_r^{(K)}[i_1|N_s^{(K)} = N_s^{(K)}(\mathbf{d})] \ge w_{i_1}^r/\sum_{j=1}^k w_{i_j}^r$ and $\sum_{\mathbf{d}\in S_{\mathbf{b}}^{(K)}} W_{\mathbf{d}}^{(K)} P_r^{(C)}[i_k|N_s^{(K)} = N_s^{(K)}(\mathbf{d})] \le w_{i_k}^r/\sum_{j=1}^k w_{i_j}^r$. We prove the theorem for c coins with probabilities $p_1 \ge \cdots \ge p_c$. First consider the case $s < r$. In order to be precise in the enumeration of permutations of $\mathbf{b}$ in $S_{\mathbf{b}}^{(C)}$, we define $Perm(c)$ to be the set of all $c!$ permutation functions $\sigma : (12\cdots c) \to \sigma((12\cdots c))$ on c items. For any configuration $\mathbf{v}$, define $\nu(\mathbf{v})$ to be the number of permutations $\sigma \in Perm(c)$ such that $\sigma(\mathbf{v}) = \mathbf{v}$. Thus there are $c!/\nu(\mathbf{b})$ distinct permutations of $\mathbf{b} \in S_{\mathbf{b}}^{(C)}$. Therefore, enumerating the permutations $\sigma \in Perm(c)$ as $\sigma_1, \ldots, \sigma_{c!}$, we have

$$\sum_{\mathbf{d}\in S_{\mathbf{b}}^{(C)}} W_{\mathbf{d}}^{(C)} P_r^{(C)}[1|N_s^{(C)} = N_s^{(C)}(\mathbf{d})]$$
$$= \sum_{n=1}^{c!} W_{\sigma_n(\mathbf{b})}^{(C)} \cdot P_r^{(C)}[1|N_s^{(C)} = N_s^{(C)}(\sigma_n(\mathbf{b}))] \cdot 1/\nu(\mathbf{b}). \tag{3}$$

Now $P_r^{(C)}[1|N_s^{(C)} = N_s^{(C)}(\sigma_n(\mathbf{b}))] = \sum_{\mathbf{a}\in B_r^{(C)}} \sum_{m=1}^{c!} P[N_r^{(C)} = N_r^{(C)}(\sigma_m(\mathbf{a}))| N_s^{(C)} = N_s^{(C)}(\sigma_n(\mathbf{b}))] \times P_r^{(C)}[1|N_r^{(C)} = N_r^{(C)}(\sigma_m(\mathbf{a}))] \times 1/\nu(\mathbf{a})$ where we have used the Markovian property $P_r^{(C)}[1|N_r^{(C)} = N_r^{(C)}(\sigma_m(\mathbf{a})),\ N_s^{(C)} = N_s^{(C)}(\sigma_n(\mathbf{b}))] = P_r^{(C)}[1|\ N_r^{(C)} = N_r^{(C)}(\sigma_m(\mathbf{a}))]$. This follows from the fact that the probability of ultimately selecting coin 1 (or any other coin) in the e-game given that a specific configuration occurs at the time of first elimination does not depend on the path leading up to that configuration prior to the time of first elimination. Thus (3) becomes

$$\begin{aligned}&\sum_{n=1}^{c!}\sum_{\mathbf{a}\in B_r^{(C)}}\sum_{m=1}^{c!} W^{(C)}_{\sigma_n(\mathbf{b})} P[N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))|N_s^{(C)}=N_s^{(C)}(\sigma_n(\mathbf{b}))]\\ &\quad\times P_r^{(C)}[1|N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))]\cdot 1/\nu(\mathbf{a})\cdot 1/\nu(\mathbf{b})\\ &=\sum_{\mathbf{a}\in B_r^{(C)}}\sum_{n=1}^{c!}\sum_{m=1}^{c!} W^{(C)}_{\sigma_n(\mathbf{b})} P[N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))|N_s^{(C)}=N_s^{(C)}(\sigma_n(\mathbf{b}))]\\ &\quad\times P_r^{(C)}[1|N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))]\cdot 1/\nu(\mathbf{a})\cdot 1/\nu(\mathbf{b}) \qquad (4)\end{aligned}$$

Since $S_{c!}$ is a group under permutation multiplication, for any given σ_n and σ_m in $S_{c!}$, there exists one and only one $\sigma_k \in S_{c!}$ such that $\sigma_n^{-1}\sigma_m = \sigma_k$. Using Corollary 4, we have $(W^{(C)}_{\sigma_m(\mathbf{a})}/W^{(C)}_{\sigma_k(\mathbf{a})}) \times P[N_r^{(C)} = N_r^{(C)}(\sigma_k(\mathbf{a}))|N_s^{(C)} = N_s^{(C)}(\mathbf{b})] = K_b^{(C)} W^{(C)}_{\sigma_n(\mathbf{b})} \times P[N_r^{(C)} = N_r^{(C)}(\sigma_m(\mathbf{a}))|N_s^{(C)} = N_s^{(C)}(\sigma_n(\mathbf{b}))]$, and (4) becomes

$$\begin{aligned}&\sum_{\mathbf{a}\in B_r^{(C)}}\sum_{k=1}^{c!}\sum_{m=1}^{c!}\frac{1}{K_b^{(C)}}\cdot\frac{W^{(C)}_{\sigma_m(\mathbf{a})}}{W^{(C)}_{\sigma_k(\mathbf{a})}} P[N_r^{(C)}=N_r^{(C)}(\sigma_k(\mathbf{a}))|N_s^{(C)}=N_s^{(C)}(\mathbf{b})]\\ &\quad\times P_r^{(C)}[1|N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))]\cdot 1/\nu(\mathbf{a})\cdot 1/\nu(\mathbf{b})\\ &=\sum_{\mathbf{a}\in B_r^{(C)}}\sum_{k=1}^{c!}\frac{1}{K_b^{(C)}}\cdot {W^{(C)}_{\sigma_k(\mathbf{a})}}^{-1} P[N_r^{(C)}=N_r^{(C)}(\sigma_k(\mathbf{a}))|N_s^{(C)}=N_s^{(C)}(\mathbf{b})]\cdot 1/\nu(\mathbf{b})\\ &\quad\cdot\ \{\sum_{m=1}^{c!} W^{(C)}_{\sigma_m(\mathbf{a})} P_r^{(C)}[1|N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))]\cdot 1/\nu(\mathbf{a})\}. \qquad (5)\end{aligned}$$

It should be clear that expression (5) for $P_r^{(C)}[1|N_s^{(C)} = N_s^{(C)}(\mathbf{d})$ for some $\mathbf{d} \in S_{\mathbf{b}}^{(C)}]$ can be extended to $P_r^{(C)}[i|N_s^{(C)} = N_s^{(C)}(\mathbf{d})$ for some $\mathbf{d} \in S_{\mathbf{b}}^{(C)}]$ for any other coin i simply by replacing 1 with i in the term in braces in (5). In particular, the leading terms multiplying the expression in braces are the same for any such i. We now argue that it suffices to show that the lower bound inequality holds in the remaining case $s = r$, i.e., for any $\mathbf{a} \in B_r^{(C)}$,

$$\sum_{m=1}^{c!} W^{(C)}_{\sigma_m(\mathbf{a})} P_r^{(C)}[1|N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))]\cdot 1/\nu(\mathbf{a}) \ge w_1^r/\sum_{i=1}^{c} w_i^r. \qquad (6)$$

Because

$$\sum_{i=1}^{c}\sum_{m=1}^{c!} W^{(C)}_{\sigma_m(\mathbf{a})} P_r^{(C)}[i|N_r^{(C)}=N_r^{(C)}(\sigma_m(\mathbf{a}))]\cdot 1/\nu(\mathbf{a})$$

$$= P[\text{ some coin } i \text{ wins}|N_r^{(C)} = N_r^{(C)}(\mathbf{u}) \text{ for some } \mathbf{u} \in S_{\mathbf{a}}^{(C)}] = 1$$

(proof omitted), we have that (6) is equivalent to (7) below:

$$\frac{\sum_{m=1}^{c!} W_{\sigma_m(\mathbf{a})}^{(C)} P_r^{(C)}[1|N_r^{(C)} = N_r^{(C)}(\sigma_m(\mathbf{a})) = N_r^{(C)}] \cdot \frac{1}{\nu(\mathbf{a})}}{\sum_{i=2}^{c} \sum_{m=1}^{c!} W_{\sigma_m(\mathbf{a})}^{(C)} P_r^{(C)}[i|N_r^{(C)} = N_r^{(C)}(\sigma_m(\mathbf{a})) = N_r^{(C)}] \cdot \frac{1}{\nu(\mathbf{a})}} \geq \frac{w_1^r}{\sum_{j=2}^{c} w_j^r}. \quad (7)$$

Moving the denominator of the left hand side of (7) to the right hand side, multiplying by all the leading terms of (5) before the braces, and summing over $\sum_{\mathbf{a}\in B_r^{(C)}} \sum_{k=1}^{c!}$ implies that

$$\begin{aligned} &P_r^{(C)}[1|N_s^{(C)} = N_s^{(C)}(\mathbf{d}) \text{ for some } \mathbf{d} \in S_{\mathbf{b}}^{(C)}] \\ &\geq (w_1^r / \sum_{j=2}^{c} w_j^r) \sum_{i=2}^{c} P_r^{(C)}[i|N_s^{(C)} = N_s^{(C)}(\mathbf{d}) \text{ for some } \mathbf{d} \in S_{\mathbf{b}}^{(C)}], \end{aligned}$$

which concludes the proof that (6) suffices.

To prove (6), for any $\mathbf{u} \in S_{\mathbf{a}}^{(C)}$, by definition, $W_{\mathbf{u}}^{(C)} = P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]/ \sum_{\mathbf{u}'\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}')]$, and therefore the left hand side (lhs) of (6) is

$$\begin{aligned} &\sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}} W_{\mathbf{u}}^{(C)} \times P_r^{(C)}[1|N_r^{(C)} = N_r^{(C)}(\mathbf{u})] \\ &= \sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}} \frac{P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]}{\sum_{\mathbf{u}'\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}')]} \cdot P_r^{(C)}[1|N_r^{(C)} = N_r^{(C)}(\mathbf{u})]. \quad (8) \end{aligned}$$

We must now consider the circumstances after the first elimination. Given that a sequence of tosses reaches a configuration in $S_{\mathbf{a}}^{(C)}$, there are $\binom{c}{n_0(\mathbf{a})}$ different subsets of coins that may still remain in the game, where $n_0(\mathbf{a})$ is defined as the number of 0s appearing in $\mathbf{a}$. Therefore, $S_{\mathbf{a}}^{(C)}$ can be divided into $\binom{c}{n_0(\mathbf{a})}$ disjoint subsets that contain all the configurations in $S_{\mathbf{a}}^{(C)}$ with the same subset of coins remaining in the game. Enumerate the subsets as $F_1, \ldots, F_{\binom{c}{n_0(\mathbf{a})}}$. Then (8) becomes

$$\begin{aligned} &\sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}} \sum_{\mathbf{u}\in F_i} P[N_r^{(C)}{=}N_r^{(C)}(\mathbf{u})] P_r^{(C)}[1|N_r^{(C)}{=}N_r^{(C)}(\mathbf{u})] / \sum_{\mathbf{u}'\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)}{=}N_r^{(C)}(\mathbf{u}')] \\ &= \sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}} \Big\{ [\sum_{\mathbf{u}'\in F_i} P[N_r^{(C)}{=}N_r^{(C)}(\mathbf{u}')]] \cdot \sum_{\mathbf{u}\in F_i} \frac{P[N_r^{(C)}{=}N_r^{(C)}(\mathbf{u})]}{\sum_{\mathbf{u}'\in F_i} P[N_r^{(C)}{=}N_r^{(C)}(\mathbf{u}')]} \\ &\quad \times P_r^{(C)}[1|N_r^{(C)}{=}N_r^{(C)}(\mathbf{u})] \Big\} / \sum_{\mathbf{u}'\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)}{=}N_r^{(C)}(\mathbf{u}')]. \quad (9) \end{aligned}$$

We wish to re-express the term $P[N_r^{(C)}=N_r^{(C)}(\mathbf{u})]/\sum_{\mathbf{u}'\in F_i} P[N_r^{(C)}=N_r^{(C)}(\mathbf{u}')]$ as a weight in the *reduced* game that continues after the first elimination. To do this, for every $1 \le i \le \binom{c}{n_0(\mathbf{a})}$, choose the configuration $\mathbf{u_i^*}$ that is ordered apart from the zero components fixed in F_i. For $\mathbf{u} \in F_i$, represent probability $P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]$ as $P[N_r^{(C)} = N_r^{(C)}(\mathbf{u_i^*})]$ times the appropriate permutation constant. The weight that results after cancelling $P[N_r^{(C)} = N_r^{(C)}(\mathbf{u_i^*})]$ from numerator and denominator is identical to the weight $W^{(C_i')}_{\mathbf{d}'(\mathbf{u})}$, say, where $\mathbf{d}'(\mathbf{u})$ is the reduced configuration from which the $c - n_0(\mathbf{a})$ coin game continues with coins $C_i' = \{j_1, \ldots, j_{c-n_0(\mathbf{a})}\}$, namely $(u_{j_1} - a_0, \ldots, u_{j_{c-n_0(\mathbf{a})}} - a_0)$, $u_{j_k} \ne 0$, where $a_0 = \min_{1\le j\le c, u_j\ne 0} u_j$. For example, with $c = 4$, $r = 2$, one of the subsets F_i is $\{(2011), (1021), (1012)\}$. Choose $\mathbf{u_i^*} = (2011)$ and $\mathbf{u} = (1021)$. Then the term

$$
\begin{aligned}
&P[N_2^{(C)} = N_2^{(C)}(\mathbf{u})]/\sum_{\mathbf{u}'\in F_i} P[N_2^{(C)} = N_2^{(C)}(\mathbf{u}')] \\
&= \frac{P[N_2^{(C)} = N_2^{(C)}((1021))]}{P[N_2^{(C)}=N_2^{(C)}((2011))] + P[N_2^{(C)}=N_2^{(C)}((1021))] + P[N_2^{(C)}=N_2^{(C)}((1012))]} \\
&= P[N_2^{(C)} = N_2^{(C)}((2011))]w_{13}^{-1}/(P[N_2^{(C)} = N_2^{(C)}((2011))](1 + w_{13}^{-1} + w_{14}^{-1})) \\
&= w_{13}^{-1}/(1 + w_{13}^{-1} + w_{14}^{-1}) = W^{(C_i')}_{(010)} = W^{(C_i')}_{\mathbf{d}'(\mathbf{u})},
\end{aligned}
$$

where $C_i' = \{1, 3, 4\}$ and $\mathbf{d}'(\mathbf{u}) = (010)$ is the reduced configuration in the game that continues. That is, $W^{(C_i')}_{(010)}$ is identical to the conditional probability that $\mathbf{d}'(\mathbf{u}) = (010)$ would be the first configuration to be reached among all permutations of (010), given some such permutation is reached at time $N_1^{(C')}$ in a three coin game with coins $C_i' = \{1, 3, 4\}$. In general, then, writing $P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]/\sum_{\mathbf{u}'\in F_i} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}')] = W^{(C_i')}_{\mathbf{d}'(\mathbf{u})}$, left hand side of (6) = (8) = (9) becomes

$$
\sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}} \Big\{ [\sum_{\mathbf{u}'\in F_i} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}')]] \sum_{\mathbf{d}'(\mathbf{u})\in F_i'} W^{(C)}_{\mathbf{d}'(\mathbf{u})} P_r^{(C')}[1|N_{s'}^{(C')} = N_{s'}^{(C')}(\mathbf{d}'(\mathbf{u}))] \Big\}
$$
$$
/ \sum_{\mathbf{u}'\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}')]. \tag{10}
$$

where $s' = \max_l (d'(\mathbf{u}))_l$ and $F_i' = \bigcup_{\mathbf{u}\in F_i}\{\mathbf{d}'(\mathbf{u})\} = S^{(C')}_{\mathbf{d}'(\mathbf{a})}$ comprises the set of all permutations of reduced configurations from F_i. By the inductive hypotheses,

$$
\sum_{\mathbf{d}'(\mathbf{u})\in F_i'} W^{(C')}_{\mathbf{d}'(\mathbf{u})} P_r^{(C')}[1|N_{s'}^{(C')} = N_{s'}^{(C')}(\mathbf{d}'(\mathbf{u}))] \ge \frac{w_1^r}{\sum_{l=1}^{c-n_0(\mathbf{a})} w_{j_l}^r} = \frac{1}{\sum_{j\in C_i'} w_{1j}^{-r}}. \tag{11}
$$

Consider next the leading factor in (10), which we rewrite without primes as $1/\sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]$. Define

$$E_{\mathbf{a}}^{(C)} = \{\mathbf{u} \in S_{\mathbf{a}}^{(C)} | u_1 = r\}$$

and

$$n_r(\mathbf{a}) = \sum_{j=1}^{c} I_{\{a_j = r\}} = \#\{a_j = r\}.$$

We can generate each $\mathbf{u} \in S_{\mathbf{a}}^{(C)}$ as a $\mathbf{u} \in E_{\mathbf{a}}^{(C)}$ followed by a transposition of component 1 with j for $j = 1, \ldots, c$. Then we have

$$\sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})] = \sum_{\mathbf{u}\in E_{\mathbf{a}}^{(C)}} \sum_{j=1}^{c} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}_{1j})] \frac{1}{n_r(\mathbf{a})}$$

$$= P[N_r^{(C)} = N_r^{(C)}(\mathbf{a})] \sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}} \sum_{\mathbf{u}\in F_i \cap E_{\mathbf{a}}^{(C)}} \sum_{j=1}^{c} K_a^{(C)} W_{\mathbf{u}}^{(C)} \frac{1}{n_r(\mathbf{a})} w_{1j}^{u_j - u_1}, \tag{12}$$

by Corollaries 1 and 2. We want to put the remaining terms of (10) in similar form. Thus (10) is

$$\sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}} \sum_{\mathbf{u}\in F_i} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})] \cdot \{*\} \sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]$$

$$= \frac{\sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}} \sum_{\mathbf{u}\in F_i \cap E_{\mathbf{a}}^{(C)}} \sum_{j\in C_i'} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}_{1j})] \cdot \frac{1}{n_r(\mathbf{a})} \cdot \{*\}}{\sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]}, \tag{13}$$

where $\{*\}$ is the left hand side of (11). Since for any $\mathbf{u} \in F_i \cap E_{\mathbf{a}}^{(C)}$ and $j \in C_i'$, $P[N_r^{(C)} = N_r^{(C)}(\mathbf{u}_{1j})] = K_a^{(C)} W_{\mathbf{u}}^{(C)} w_{1j}^{u_j - u_1} \times P[N_r^{(C)} = N_r^{(C)}(\mathbf{a})]$ we have that the left side of (6) is (10) and is

$$\frac{P[N_r^{(C)} = N_r^{(C)}(\mathbf{a})] \sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}} \sum_{\mathbf{u}\in F_i \cap E_{\mathbf{a}}^{(C)}} K_a^{(C)} W_{\mathbf{u}}^{(C)} \frac{1}{n_r(\mathbf{a})} \sum_{j\in C_i'} w_{1j}^{u_j - u_1} \cdot \{*\}}{\sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}} P[N_r^{(C)} = N_r^{(C)}(\mathbf{u})]}. \tag{14}$$

Combining (11) and (14), the left size of (6) is at least

$$\frac{P[N_r^{(C)}=N_r^{(C)}(\mathbf{a})]\sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}}\sum_{\mathbf{u}\in F_i\cap E_{\mathbf{a}}^{(C)}}K_a^{(C)}W_{\mathbf{u}}^{(C)}\frac{1}{n_r(\mathbf{a})}(\sum_{j\in C_i'}w_{1j}^{u_j-u_1})/(\sum_{j\in C_i'}w_{1j}^{-r})}{\sum_{\mathbf{u}\in S_{\mathbf{a}}^{(C)}}P[N_r^{(C)}=N_r^{(C)}(\mathbf{u})]}$$

$$=\frac{\sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}}\sum_{\mathbf{u}\in F_i\cap E_{\mathbf{a}}^{(C)}}K_a^{(C)}W_{\mathbf{u}}^{(C)}\frac{1}{n_r(\mathbf{a})}\cdot\frac{\sum_{j\in C_i'}w_{1j}^{u_j-u_1}}{\sum_{j\in C_i'}w_{1j}^{-r}}}{\sum_{i=1}^{\binom{c}{n_0(\mathbf{a})}}\sum_{\mathbf{u}\in F_i\cap E_{\mathbf{a}}^{(C)}}K_a^{(C)}W_{\mathbf{u}}^{(C)}\frac{1}{n_r(\mathbf{a})}\sum_{j=1}^{c}w_{1j}^{u_j-u_1}}$$

using (12). It suffices, then, to show that

$$\sum_{j\in C_i'}w_{1j}^{u_j-u_1}\Big/\sum_{j\in C_i'}w_{1j}^{-r}\geq\sum_{j=1}^{c}w_{1j}^{u_j-u_1}\Big/\sum_{j=1}^{c}w_{1j}^{-r}$$

for any i and C_i', because then lhs(6) $\geq 1/\sum_{j=1}^{c}w_{1j}^{-r} = w_1^r/\sum_{j=1}^{c}w_j^r$. But for $\mathbf{u}\in F_i\cap E_{\mathbf{a}}^{(C)}$, $u_1=r$, so $u_j-u_1\geq -r$, whence $\{\sum_{j\in C_i'}w_{1j}^{u_j-u_1}/\sum_{j\in C_i'}w_{1j}^{-r}\}\geq 1$. Note also that $j\in C_i'$ means $u_j>0$ while $j\notin C_i'$ implies $u_j=0$, because C_i' are the coins not eliminated when configuration $\mathbf{u}$ is reached at first elimination.

Therefore $\sum_{j\notin C_i'}w_{1j}^{u_j-u_1}=\sum_{j\notin C_i'}w_{1j}^{-r}$; thus

$$\frac{\sum_{j\in C_i'}w_{1j}^{u_j-u_1}}{\sum_{j\in C_i'}w_{1j}^{-r}}\geq\frac{\sum_{j\in C_i'}w_{1j}^{u_j-u_1}+\sum_{j\notin C_i'}w_{1j}^{u_j-u_1}}{\sum_{j\in C_i'}w_{1j}^{-r}+\sum_{j\notin C_i'}w_{1j}^{-r}}=\frac{\sum_{j=1}^{c}w_{1j}^{u_j-u_1}}{\sum_{j=1}^{c}w_{1j}^{-r}}.$$

This concludes the proof of the lower bound for $P_r^{(C)}[1]$. For the upper bound inequality, the development is entirely analogous, and is omitted for brevity.

4. Discussion

It would be natural to attempt to prove the lower bound inequality in (1) by proving only (6), instead of the stronger Theorem 1. As it sounds, (6) states that the conditional probability of correct selection given that the sample path reaches one of the symmetrical subsets of configurations $S_{\mathbf{a}}$ at the time of first elimination satisfies the lower bound formula, and since all paths must lead to some such subset, the inequality follows. The stronger theorem is required, though, because the inductive step at (11) takes place in a reduced game where the current configuration is no longer on an elimination boundary. This is the appropriate generalization of the observation made for the two coin game, where $w_1^r/(w_1^r + w_2^r)$ equalled the weighted average of gambler's ruin probabilities starting at any configuration of the form $(s0)$ or $(0s)$ inside the elimination boundary.

References

Feller, W. (1957). *An Introduction to Probability Theory and Its Application.* Wiley, New York.

Levin, B. (1984). On a sequential selection procedure of Bechhofer, Kiefer, and Sobel. *Statist. Probab. Lett.* **2**, 91-94.

Levin, B. and Robbins, H. (1981). Selecting the highest probability in binomial or multinomial trials. *Proc. Natl. Acad. Sci. USA* **78**, 4663-4666.

Zybert, P. and Levin, B. (1987). Selecting the highest of three binomial probabilities. *Proc. Natl. Acad. Sci. USA* **84**, 8180-8184.

600 West 168th Street, Columbia University, New York, NY 10032, U.S.A.

E-mail: Bruce.Levin@columbia.edu

(Received July 1997; accepted September 1998)