View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Tamkang University Institutional Repository

SIAM J. DISCRETE MATH.
Vol. 19, No. 1, pp. 149–164
© 2005 Society for Industrial and Applied Mathematics

A RANDOM VERSION OF SHEPP'S URN SCHEME*

ROBERT W. CHEN†, ALAN ZAME†, CHIEN-TAI LIN‡, AND HSIU-FEN WU‡

Abstract. In this paper, we consider the following random version of Shepp's urn scheme: A player is given an urn with n balls. p of these balls have value $+1$ and $n-p$ have value -1. The player is allowed to draw balls randomly, without replacement, until he or she wants to stop. The player knows n, the total number of balls, but knows only that p, the number of balls of value $+1$, is a number selected randomly from the set $\{0, 1, 2, \ldots, n\}$. The player wishes to maximize the expected value of the sum of the balls drawn. We first derive the player's optimal drawing policy and an algorithm to compute the player's expected value at the stopping time when he or she uses the optimal drawing policy. Since the optimal drawing policy is rather intricate and the computation of the player's optimal expected value is quite cumbersome, we present a very simple drawing policy, which is asymptotically optimal. We also show that this random urn scheme is equivalent to a random coin tossing problem.

Key words. urn scheme, optimal drawing policy, random coin tossing process, stopping time, the "k" in the hole policy

AMS subject classifications. Primary, 60G40; Secondary, 60K99

DOI. 10.1137/S0895480102418099

1. Introduction.

In [8], Shepp considered the following optimal stopping problem: A player is given an urn with n balls. p of these balls have value $+1$ and $n-p$ have value -1. The player knows n and p. The player's goal is to maximize the expected value of the sum of the balls drawn. The player may draw as long as he or she wishes, without replacement. Shepp was interested in knowing for what n and p there is a drawing policy for which $V(n,p)$, the expected value of the game if there are n balls, p of which are $+1$, is positive. He showed that for a given n there is an integer $\gamma(n)$ such that $V(n,p) > 0$ if and only if $p \geq \gamma(n)$. More precisely, he showed that there exists a $\beta(p)$ for which $V(n,p) > 0$ if and only if $0 \leq n-p \leq \beta(p)$.

In [1], Boyce was interested in the following bond-selling problem: A corporation must repay 10 million dollars in bank loans in three months, and it wishes to sell bonds to repay the loan. However, the company's economists predict that in three months bond prices will be lower (interest rates higher). Should the corporation issue the bonds now, wait a month or two, or wait the full three months? For this bond-selling problem, Boyce introduced a random version of Shepp's urn scheme, which can be stated as follows: A player is given an urn with n balls. p of these balls have value $+1$ and $n-p$ have value -1. The player is allowed to draw balls randomly, without replacement, until he or she wants to stop. The player knows n, the total number of balls, but only knows the distribution of p, the number of balls of value $+1$. The player wishes to maximize the expected value of the sum of his draws. Boyce briefly studied this problem and proposed a procedure to compute the player's expected value at the stopping time when he or she uses an optimal drawing policy.

*Received by the editors November 18, 2002; accepted for publication (in revised form) June 26, 2004; published electronically June 16, 2005.

http://www.siam.org/journals/sidma/19-1/41809.html

†Department of Mathematics, University of Miami, Coral Gables, FL 33124 (chen@math.miami.edu, zame@math.miami.edu). The work of the first author was partially supported by the Department of Mathematics, Tamkang University of Taiwan, and was partially done during his visit.

‡Department of Mathematics, Tamkang University, Taipei, Taiwan (chien@math.tku.edu.tw, hfwu@math.tku.edu.tw).

In this paper, we study this random version of Shepp's urn scheme for the case when the distribution of p is uniform over the set $\{0, 1, 2, \ldots, n\}$. In section 2, we derive the player's optimal drawing policy and an algorithm to compute the player's expected value at the stopping time when he or she uses the optimal drawing policy. It will be seen that the optimal drawing policy is very intricate. Also the computation of the player's optimal expected value at the stopping time is quite cumbersome, especially as n gets large. In section 3, we present a very simple drawing policy and show that this simple drawing policy is not only asymptotically optimal, but also performs very well even when n is small. Our data reveal that for $n = 10, 20, \ldots, 1{,}000$, the difference between the expected value at the stopping time under an optimal drawing policy and the expected value at the stopping time under this simple drawing policy is less than 1. In section 4, we will show that this random urn problem can be stated as a random coin tossing problem.

2. An optimal drawing policy. In order to compute the expected value of the game, we consider what the remaining value of the game would be, conditioned on the outcome of the first k draws. Suppose there are n balls and k balls have been drawn, j of which have value +1. Let $E(n, k, j)$ denote the remaining expected value of the game from this point on. The following lemma gives the critical recursion for E. Its proof is a straightforward application of Bayes' law, but we include a proof for the sake of completeness.

LEMMA 1. *If* $0 \le j \le k \le n-1$, *then*

$$E(n,k,j) = \max\left\{0, \frac{2j-k}{k+2} + \frac{j+1}{k+2}E(n,k+1,j+1) + \frac{k-j+1}{k+2}E(n,k+1,j)\right\}.$$

Proof. Having drawn k balls, with j "+1 balls," the player has to decide whether to play any further. The player may draw another ball; suppose that the conditional probability that it is +1 is $\alpha(n, k, j) = \alpha$ and that it is -1 is $\beta = 1 - \alpha$. The expected value of the remainder of the game, if another ball is drawn, is then

$$\alpha - \beta + \alpha E(n, k+1, j+1) + \beta E(n, k+1, j).$$

Thus, the player should draw another ball if this is positive and stop otherwise. The recursion relation will then follow if we can show that $\alpha = \frac{j+1}{k+2}$. To do this, let $X_i = 1$ if the ith draw is a +1, or let $X_i = 0$ otherwise for all $i = 1, 2, \ldots, k+1$. Let $S_k = \sum_{i=1}^{k} X_i$. Then it is easy to see that

$$\alpha = P(X_{k+1} = 1 | S_k = j) = P([S_k = j] \cap [X_{k+1} = 1])/P(S_k = j).$$

For each $i = 0, 1, 2, \ldots, n$, let A_i denote the event $[p = i]$. Since the distribution of p is uniform over the set $\{0, 1, 2, \ldots, n\}$, $P(A_i) = \frac{1}{n+1}$ for all $i = 0, 1, 2, \ldots, n$.

$$\begin{aligned} P(S_k = j) &= \sum_{i=1}^{n} P([S_k = j] \cap A_i) = \sum_{i=1}^{n} P([S_k = j]|A_i)P(A_i) \\ &= \frac{1}{n+1}\sum_{i=j}^{n-k+j} \frac{\binom{i}{j}\binom{n-i}{k-j}}{\binom{n}{k}} = \frac{\binom{n+1}{k+1}}{(n+1)\binom{n}{k}} = \frac{1}{k+1} \end{aligned}$$

since $P([S_k = j]|A_i) = 0$ if $i < j$ or $i > n - k + j$.

Similarly,

$$P([S_k = j] \cap [X_{k+1} = 1]) = \sum_{i=1}^{n} P([S_k = j] \cap [X_{k+1} = 1] \cap A_i)$$

$$= \sum_{i=1}^{n} P([S_k = j] \cap [X_{k+1} = 1] | A_i) P(A_i) = \frac{1}{n+1} \sum_{i=j+1}^{n-k+j} \frac{(i-j)\binom{i}{j}\binom{n-i}{k-j}}{(n-k)\binom{n}{k}}$$

$$= \frac{(j+1)}{(n+1)(n-k)\binom{n}{k}} \sum_{i=j+1}^{n-k+j} \binom{i}{j+1}\binom{n-i}{k-j} = \frac{(j+1)}{(k+1)(k+2)}$$

since $P([S_k = j] \cap [X_{k+1} = 1] | A_i) = 0$ if $i < j + 1$ or $i > n - k + j$.

Therefore,

$$\alpha = \left\{ \frac{(j+1)}{(k+1)(k+2)} \right\} \Big/ \left\{ \frac{1}{(k+1)} \right\} = \frac{(j+1)}{(k+2)}.$$

This completes the proof of Lemma 1. □

It is clear that $E(n, n, j) = 0$ for all $j = 0, 1, 2, \ldots, n$ since there are no balls left. It is also clear that if the player draws k balls, j of which have value $+1$, the player should stop drawing unless $E(n, k, j) > 0$. Therefore, the optimal drawing policy can be stated as follows: At the beginning, the player will draw a ball if and only if $E(n, 0, 0) > 0$. Suppose that the player has drawn k balls, j of which have value $+1$; the player will continue to draw if and only if $E(n, k, j) > 0$.

Boyce briefly studied this problem in [1]. He produces a procedure for computing $E(n, 0, 0)$. The procedure requires computing all of $E(n, k, j)$ for $0 \le j \le k < n$ in order to get $E(n, 0, 0)$. It is clear that the computation is very cumbersome, and for each new n, we have to compute all new $E(n, k, j)$ for all $0 \le j \le k < n$ to determine the new optimal drawing policy.

Table 1 gives partial values of $E(120, k, j)$ for $0 \le j \le k \le 25$.

The optimal drawing policy can also be stated as follows: For each given n, we create a table such as Table 1. We start from the position in which $k = 0$ and $j = 0$ and move one step down or one step to the southeast according to when a "−1" ball is drawn or a "+1" ball is drawn. We will stop drawing if and only if we reach a zero. However, even when n is moderate, it takes too much time to construct such a table.

THEOREM 1. *$E(n, 0, 0)$ is a strictly increasing function of n.*

Proof. It is sufficient to show that $E(n+1, k, j) \ge E(n, k, j)$ and $E(n+1, k, k) > E(n, k, k)$ for all $0 \le j \le k \le n$. Since $E(n+1, n, j) = \max\{0, \frac{2j-n}{n+2}\}$ and $E(n, n, j) = 0$ for all $0 \le j \le n$, and since $E(n+1, k, j) = \max\{0, \frac{2j-k}{k+2} + \frac{j+1}{k+2} E(n, k+1, j+1) + \frac{k-j+1}{k+2} E(n, k+1, j)\}$ and $E(n, k, j) = \max\{0, \frac{2j-k}{k+2} + \frac{j+1}{k+2} E(n, k+1, j+1) + \frac{k-j+1}{k+2} E(n, k+1, j)\}$ for all $0 \le j \le k \le n-1$ and $n \ge 1$, by mathematical induction we can conclude that $E(n+1, k, k) > E(n, k, k)$ for all $0 \le k \le n$ and $n \ge 1$. Therefore, $E(n, 0, 0)$ is a strictly increasing function of n. □

THEOREM 2. $E(n, 0, 0) \le \frac{n}{4} + o(n)$.

TABLE 1
$E(120, k, j)$.

k	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$
0	27.49	-	-	-	-	-	-	-	-
1	6.98	48.00	-	-	-	-	-	-	-
2	1.28	19.38	61.80	-	-	-	-	-	-
3	0	7.13	31.63	71.20	-	-	-	-	-
4	0	2.02	15.30	42.19	77.70	-	-	-	-
5	0	0.23	6.62	23.98	50.79	82.28	-	-	-
6	0	0	2.29	12.72	32.17	57.64	85.55	-	-
7	0	0	0.44	6.03	19.41	39.42	63.05	87.91	-
8	0	0	0	2.33	10.91	26.01	45.63	67.32	89.81
9	0	0	0	0.56	5.49	16.32	32.13	50.84	70.68
10	0	0	0	0	2.28	9.55	21.80	37.61	55.18
11	0	0	0	0	0.61	5.02	14.07	27.04	42.39
12	0	0	0	0	0	2.20	8.47	18.73	31.86
13	0	0	0	0	0	0.64	4.60	12.35	23.27
14	0	0	0	0	0	0.01	2.11	7.60	16.37
15	0	0	0	0	0	0	0.67	4.24	10.97
16	0	0	0	0	0	0	0.04	2.00	6.87
17	0	0	0	0	0	0	0	0.68	3.91
18	0	0	0	0	0	0	0	0.07	1.89
19	0	0	0	0	0	0	0	0	0.67
20	0	0	0	0	0	0	0	0	0.09
21	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0

TABLE 1
Continued.

k	$p=j$	$j=10$	$j=11$	$j=12$	$j=13$	$j=14$	$j=15$	$j=16$	$j=17$
0	-	-	-	-	-	-	-	-	-
1	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-	-
9	90.82	-	-	-	-	-	-	-	-
10	73.35	91.67	-	-	-	-	-	-	-
11	58.77	75.47	92.23	-	-	-	-	-	-
12	46.51	61.75	77.15	95.57	-	-	-	-	-
13	36.19	50.04	64.22	78.47	92.73	-	-	-	-
14	27.53	40.02	53.05	66.26	79.50	92.75	-	-	-
15	20.35	31.44	43.37	55.61	67.94	80.29	92.65	-	-
16	14.50	24.15	34.96	46.29	57.79	69.33	80.89	92.44	-
17	9.84	18.03	27.68	38.10	48.82	59.64	70.47	81.32	92.16
18	6.25	12.97	21.43	30.92	40.88	51.01	61.20	71.40	81.60
19	3.61	8.90	16.12	24.64	33.83	43.32	52.91	62.52	72.14
20	1.78	5.72	11.70	19.19	27.60	36.45	45.47	54.55	63.64
21	0.65	3.44	8.09	14.53	22.11	30.31	38.78	47.36	55.96
22	0.10	1.68	5.25	10.61	17.32	24.84	32.76	40.86	49.00
23	0	0.63	3.10	7.39	13.18	19.98	27.35	34.97	42.69
24	0	0.10	1.57	4.83	9.68	15.72	22.49	29.64	36.95
25	0	0	0.60	2.88	6.79	12.02	18.17	24.82	31.72

Table 1
Continued.

k	$j=18$	$j=19$	$j=20$	$j=21$	$j=22$	$j=23$	$j=24$	$j=25$
0	-	-	-	-	-	-	-	-
1	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-
9	-	-	-	-	-	-	-	-
10	-	-	-	-	-	-	-	-
11	-	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-	-
13	-	-	-	-	-	-	-	-
14	-	-	-	-	-	-	-	-
15	-	-	-	-	-	-	-	-
16	-	-	-	-	-	-	-	-
17	-	-	-	-	-	-	-	-
18	91.80	-	-	-	-	-	-	-
19	81.76	91.38	-	-	-	-	-	-
20	72.73	81.82	90.91	-	-	-	-	-
21	64.57	73.17	81.78	90.39	-	-	-	-
22	57.17	65.33	73.50	81.67	89.83	-	-	-
23	50.44	58.20	65.96	73.72	81.48	89.24	-	-
24	44.31	51.69	59.08	66.46	73.85	81.23	88.62	-
25	38.72	45.74	52.77	59.81	66.85	73.89	80.93	87.96

Proof. Following Boyce [2], where the player starts with an urn with n balls and a known p of "+1" balls, let $V(n,p)$ be the player's expected score at the stopping time when he or she uses the optimal drawing policy. It is easy to see that $E(n,0,0) \leq \frac{1}{n+1}\sum_{j=0}^{n} V(n,j)$. If $n=2m$, then by a theorem of Shepp, $V(2m,j) \leq V(2m,m)$ for all $j=0,1,2,\ldots,m$ and $V(2m,j) \leq 2j-2m+V(2m,m)$ for all $j=m+1,m+2,\ldots,2m$. Since $V(2m,m) \approx \sqrt{m}$, it is easy to see that $E(n,0,0) \leq \frac{n}{4}+o(n)$. The proof for the case when n is odd is similar.

In theory, we can compute the expected score at the stopping time under the optimal drawing policy and describe the optimal drawing policy for each given positive integer n. However, even when n is just moderately large, the computation is very cumbersome and it is very difficult to describe the optimal drawing policy precisely. In section 3, we will present a simple drawing policy, which is not only asymptotically optimal, but also performs very well even when n is small. Our data reveal that for $n = 10, 20, \ldots, 1{,}000$, $E(n,0,0)-W(n,k_n) < 1$, where $W(n,k_n)$ is the expected value at the stopping time when the player uses the simple drawing policy, which will be introduced in section 3. ∎

3. A simple drawing policy. One natural approach to determine when to stop is to play until we are a certain amount "in the hole." Here we continue drawing until the number of "−1" balls drawn is k more than the number of "+1" balls drawn. We will call this strategy "the k in the hole drawing policy." Let $W(n,k)$ be the expected value of the game following "the k in the hole drawing policy" when the urn originally contains n balls. One would expect the optimal choice for k to depend on n. We will show that it does. We will also show how to compute this optimal k very quickly. Most important, we will show that for any given n "the k in the hole drawing policy" is asymptotically optimal if we choose the best k.

THEOREM 3. *For each integer $1 \le k \le n$, if $n-k$ is even,*

$$W(n,k) = \frac{1}{n+1}\left\{\frac{(n-k+2)(n-k)}{4} - \sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\},$$

and if $n-k$ is odd,

$$W(n,k) = \frac{1}{n+1}\left\{\frac{(n-k+1)^2}{4} - \sum_{j=(n-k+1)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}.$$

Proof. We give the proof when $n-k$ is even; when $n-k$ is odd the proof is similar. For each $j = 0, 1, 2, \ldots, n$, let $W(n,k,j)$ be the expected value at the stopping time when the player uses "the k in the hole drawing policy" and the urn originally contains $n-j$ balls of value -1 and j balls of value of $+1$. It is easy to see that

$$W(n,k) = \frac{1}{n+1}\sum_{j=0}^{n} W(n,k,j).$$

It is also clear that $W(n,k,j) = 2j-n$ if $j > n-k$ and $W(n,k,j) = -k$ if $j < (n-k)/2$. If $(n-k)/2 \le j \le n-k$, then by the reflection principle [5, p. 72], $W(n,k,j) = 2j-n$ with probability $1-\binom{n}{k+j}/\binom{n}{j}$ and $W(n,k,j) = -k$ with probability $\binom{n}{k+j}/\binom{n}{j}$. Therefore,

$$\begin{aligned} W(n,k) &= \frac{1}{n+1}\sum_{j=0}^{n} W(n,k,j) \\ &= \frac{1}{n+1}\left\{\sum_{j=0}^{(n-k-2)/2}(-k) + \sum_{j=n-k+1}^{n}(2j-n)\right. \\ &\qquad \left. + \sum_{j=(n-k)/2}^{n-k}\left\{(2j-n)\left[1-\frac{\binom{n}{k+j}}{\binom{n}{j}}\right] + (-k)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}\right\} \\ &= \frac{1}{n+1}\left\{-\frac{k(n-k)}{2} + \sum_{j=(n-k)/2}^{n}(2j-n) - \sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\} \\ &= \frac{1}{n+1}\left\{\frac{(n-k+2)(n-k)}{4} - \sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}. \end{aligned}$$

This completes the proof of Theorem 3. □

For each positive integer n, let

$$\begin{aligned} k_n &= \min\{k \mid 1 \le k \le n,\ W(n,k) = \max\{W(n,j) \mid 1 \le j \le n\}\}, \\ \underline{k}_n &= \max\{k \mid 1 \le k \le n,\ (n+1)(n-k)^2 \ge 8n(n-1)k^3\}, \\ \overline{k}_n &= \min\{k \mid 1 \le k \le n,\ (n+1)(n+k)^2 \le 2(n-k)^2k^3\}. \end{aligned}$$

TABLE 2
$W(n,k)$, $E(n,0,0)$, $\underline{k}_n$, k_n, *and* $\overline{k}_n$.

n	$W(n,1)$	$W(n,2)$	$W(n,3)$	$W(n,4)$	$W(n,5)$	$E(n,0,0)$	$\underline{k}_n$,	k_n	$\overline{k}_n$
10	**1.65**	1.59	1.33	1.03	0.77	**1.65**	1	1	3
20	3.57	**3.82**	3.61	3.29	2.95	**3.82**	2	2	3
30	5.50	**6.08**	5.96	5.66	5.31	**6.08**	2	2	3
40	7.43	**8.35**	8.33	8.06	7.71	**8.37**	2	2	4
50	9.36	10.62	**10.70**	10.47	10.14	**10.70**	2	3	4
60	11.29	12.89	**13.07**	12.89	12.57	**13.08**	2	3	4
70	13.22	15.16	**15.45**	15.31	15.01	**15.46**	3	3	4
80	15.15	17.43	**17.83**	17.73	17.45	**17.85**	3	3	4
90	17.08	19.71	**20.21**	20.15	19.90	**20.25**	3	3	4
100	19.01	21.98	**22.59**	22.58	22.34	**22.66**	3	3	4
200	38.32	44.71	46.41	**46.86**	46.85	**46.98**	3	4	5
300	57.64	67.45	70.24	71.15	**71.37**	**71.53**	4	5	6
400	76.95	90.19	94.07	95.44	95.90	**96.17**	4	6	6
500	96.27	112.94	117.91	119.74	120.43	**120.89**	4	6	7
600	115.58	135.68	141.74	144.04	144.96	**145.65**	5	7	7
700	134.90	158.42	165.57	168.33	169.50	**170.44**	5	7	8
800	154.21	181.16	189.40	192.63	194.03	**195.25**	5	7	8
900	173.53	203.90	213.23	216.93	218.56	**220.07**	5	8	8
1,000	192.84	226.64	237.07	241.23	243.10	**244.90**	5	8	9

TABLE 2
Continued.

n	$W(n,6)$	$W(n,7)$	$W(n,8)$	$W(n,9)$	$W(n,10)$	$E(n,0,0)$	$\underline{k}_n$,	k_n	$\overline{k}_n$
10	0.53	0.34	0.18	0.08	0	**1.65**	1	1	3
20	2.61	2.28	1.97	1.68	1.41	**3.82**	2	2	3
30	4.94	4.57	4.21	3.86	3.52	**6.08**	2	2	3
40	7.33	6.95	6.56	6.18	5.81	**8.37**	2	2	4
50	9.76	9.37	8.97	8.57	8.18	**10.70**	2	3	4
60	12.20	11.80	11.40	10.99	10.58	**13.08**	2	3	4
70	14.64	14.25	13.84	13.43	13.02	**15.46**	3	3	4
80	17.09	16.70	16.29	15.88	15.46	**17.85**	3	3	4
90	19.55	19.16	18.75	18.34	17.91	**20.25**	3	3	4
100	22.01	21.62	21.22	20.80	20.37	**22.66**	3	3	4
200	46.63	46.32	45.95	45.55	45.13	**46.98**	3	4	5
300	71.29	71.05	70.73	70.36	69.96	**71.53**	4	5	6
400	**95.95**	95.80	95.53	95.20	94.82	**96.17**	4	6	6
500	**120.62**	120.54	120.33	120.04	119.69	**120.89**	4	6	7
600	145.28	**145.30**	145.14	144.88	144.55	**145.65**	5	7	7
700	169.95	**170.05**	169.94	169.72	169.43	**170.44**	5	7	8
800	194.42	**194.80**	194.75	194.57	194.30	**195.25**	5	7	8
900	219.29	219.55	**219.56**	219.41	219.17	**220.07**	5	8	8
1,000	243.96	244.30	**244.37**	244.26	244.05	**244.90**	5	8	9

It is easy to see that $\underline{k}_n \le \overline{k}_n$ for each positive integer n. In Theorem 8, we will prove that $\underline{k}_n \le k_n \le \overline{k}_n$ for each positive integer n.

Table 2 provides some numerical values of $W(n,k)$, $E(n,0,0)$, k_n, $\underline{k}_n$, and $\overline{k}_n$ for various n and k.

Table 2 provides numerical evidence of the following: (I) For each k, $W(n,k)$ is increasing in n. (II) For each n, $W(n,k)$ first increases and then decreases in k (for small n, $W(n,k)$ is decreasing in k). (III) $\underline{k}_n \le k_n \le \overline{k}_n$. (IV) $W(n,k_n) \approx \frac{n}{4}$.

Theorem 4 proves (I). Theorem 5 proves (IV). Theorem 7 gives a partial answer to (II). Theorem 8 proves (III).

THEOREM 4. *For each fixed k $(1 \le k \le n)$, $W(n,k)$ is increasing in n.*

Proof. We will prove the case when $n-k$ is even since the proof for the case when $n-k$ is odd is similar. By Theorem 3

$$W(n+1,k)=\frac{1}{n+2}\left\{\frac{(n-k+2)^2}{4}-\sum_{j=(n-k+2)/2}^{n-k+1}(2j+k-n-1)\frac{\binom{n+1}{k+j}}{\binom{n+1}{j}}\right\}$$

$$=\frac{1}{n+2}\left\{\frac{(n-k+2)^2}{4}-\sum_{j=(n-k)/2}^{n-k}(2j+k+1-n)\frac{\binom{n+1}{k+j+1}}{\binom{n+1}{j+1}}\right\}$$

and

$$W(n,k)=\frac{1}{n+1}\left\{\frac{(n-k+2)(n-k)}{4}-\sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}.$$

Therefore,

$$W(n+1,k)-W(n,k)=\frac{1}{(n+1)(n+2)}\left\{\frac{(n+2)^2-k^2}{4}\right.$$

$$\left.-\sum_{j=(n-k)/2}^{n-k}\left\{(n+1)(2j+k+1-n)\frac{\binom{n+1}{k+j+1}}{\binom{n+1}{j+1}}-(n+2)(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}\right\}.$$

To show that $W(n+1,k)-W(n,k)\ \geq\ 0$, it is sufficient to show that

$$\sum_{j=(n-k)/2}^{n-k}\left\{(n+1)(2j+k+1-n)\frac{\binom{n+1}{k+j+1}}{\binom{n+1}{j+1}}-(n+2)(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}$$
$$\leq\frac{(n+2)^2-k^2}{4}.$$

Notice that

$$\sum_{j=(n-k)/2}^{n-k}\left\{(n+1)(2j+k+1-n)\frac{\binom{n+1}{k+j+1}}{\binom{n+1}{j+1}}-(n+2)(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}$$

$$=\sum_{j=(n-k)/2}^{n-k}\left\{\frac{(n+1)(2j+k+1-n)(j+1)-(n+2)(2j+k-n)(k+j+1)}{k+j+1}\right\}\frac{\binom{n}{k+j}}{\binom{n}{j}}.$$

For fixed $1\leq k\leq n$, let

$$g(j)=(n+1)(2j+k+1-n)(j+1)-(n+2)(2j+k-n)(k+j+1).$$

After simplification,

$$g(j)=-2j^2-(2nk+5k+1-2n)j+(n^2k+2nk+2n+1-nk^2-2k^2-k).$$

Since $k \geq 1$, $g(j)$ is decreasing for $j \geq (n-k)/2$ and $g(j) \leq 0$ if

$$j \geq \frac{1}{4}\{\, 2n - 2nk - 5k - 1 + \sqrt{4(k^2+1)n^2 + 12(k^2+1)n + (9k^2+2k+9)}\,\}.$$

Therefore, there are at most

$$\frac{1}{4}\{-2nk \;-\; 3k \;+\; 3 \;+\; \sqrt{4(k^2+1)n^2 + 12(k^2+1)n + (9k^2+2k+9)}\}$$

terms of $g(j)$ which are nonnegative and

$$\sum_{j=(n-k)/2}^{n-k} \left\{ \frac{(n+1)(2j+k+1-n)(j+1) - (n+2)(2j+k-n)(k+j+1)}{k+j+1} \right\} \frac{\binom{n}{k+j}}{\binom{n}{j}}$$
$$\leq \frac{(n+1)(n+2-k)\{-2nk-3k+3+\sqrt{4(k^2+1)n^2+12(k^2+1)n+(9k^2+2k+9)}\}}{4(n+2+k)}.$$

To show that $W(n+1,k) - W(n,k) \;\geq\; 0$, now it is sufficient to show that

$$\frac{(n+1)(n+2-k)\{-2nk-3k+3+\sqrt{4(k^2+1)n^2+12(k^2+1)n+(9k^2+2k+9)}\}}{4(n+2+k)}$$
$$\leq \frac{(n+2+k)(n+2-k)}{4},$$

which is equivalent to showing that

$$(n+1)\{-2nk-3k+3+\sqrt{4(k^2+1)n^2+12(k^2+1)n+(9k^2+2k+9)}\} \leq (n+2+k)^2.$$

To show that

$$(n+1)\{-2nk-3k+3+\sqrt{4(k^2+1)n^2+12(k^2+1)n+(9k^2+2k+9)}\} \leq (n+2+k)^2,$$

it is sufficient to show that

$$(n+1)^2\{4(k^2+1)n^2+12(k^2+1)n+(9k^2+2k+9)\} \leq \{(n+1)(2nk+3k-3)+(n+2+k)^2\}^2.$$

After simplification,

$$\{(n+1)(2nk+3k-3)+(n+2+k)^2\}^2 - (n+1)^2\{4(k^2+1)n^2+12(k^2+1)n+(9k^2+2k+9)\}$$
$$= (4k-3)n^4 + (8k^2+18k-6)n^3 + (4k^3+42k^2+30k-10)n^2 + (14k^3+70k^2+24k-16)n$$
$$+ (k^4+14k^3+42k^2+12k-8).$$

Since $1 \leq k \leq n$,

$$(4k-3)n^4 + (8k^2+18k-6)n^3 + (4k^3+42k^2+30k-10)n^2 + (14k^3+70k^2+24k-16)n$$
$$+ (k^4+14k^3+42k^2+12k-8) \geq 0.$$

This completes the proof of Theorem 4. □

THEOREM 5. $W(n,k_n) \;\approx\; \frac{n}{4}$.

Proof. Notice that for each fixed k,

$$\frac{1}{n^2}\sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}} \approx \int_{1/2}^{1}(2x-1)(1-x)^k x^{-k}dx$$

and

$$\frac{1}{n^2}\sum_{j=(n-k+1)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}} \approx \int_{1/2}^{1}(2x-1)(1-x)^k x^{-k}dx.$$

Let $t=(1-x)/x$; then

$$\int_{1/2}^{1}(2x-1)(1-x)^k x^{-k}dx = \int_0^1 (1-t)t^k(1+t)^{-3}dt.$$

By the mean value theorem,

$$\int_0^1 (1-t)t^k(1+t)^{-3}dt = a_k\int_0^1(1-t)t^k dt = \frac{a_k}{(k+1)(k+2)}$$

for some constant $\frac{1}{8} < a_k < 1$. Therefore, $\frac{4W(n,k)}{n} \approx (1-\frac{k}{n})^2 - \frac{b_k}{(k+1)(k+2)}$ as $n\to\infty$ for any fixed positive integer k, where b_k is a constant between $\frac{1}{2}$ and 4. Since k is arbitrary and $\frac{4W(n,k_n)}{n} \geq \frac{4W(n,k)}{n}$ for all $1\leq k\leq n$, $\frac{4W(n,k_n)}{n}\to 1$ as $n\to\infty$ and this completes the proof of Theorem 5. □

Combining Theorems 2 and 5, we have the following theorem.

THEOREM 6. *The k_n in the hole drawing policy is asymptotically optimal. Even though the computation for $W(n,k)$ is much simpler and faster than that for $E(n,0,0)$, we still have to compute $W(n,k)$ for all k, $1\leq k\leq n$ to identify k_n. The next result enables us to reduce the amount of required computation somewhat.*

THEOREM 7. *For $1\leq k\leq n-4$, if $W(n,k)\geq W(n,k+2)$, then $W(n,k)\geq W(n,k+2j)$ for all $1\leq j\leq (n-k)/2$.*

Proof. We will give the proof for the case when $n-k$ is even since the proof for the case when $n-k$ is odd is similar. By a direct computation, Theorem 7 holds for $1\leq n<10$, so we will assume that $n\geq 10$ in the proof below. It is also easy to verify that Theorem 7 holds if $k=n-4$; we will assume $n-k\geq 6$. For each $1\leq k\leq n$, let $u(n,k)=\frac{(n-k)(n-k+2)}{4}$ and $v(n,k)=\sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}$. Also for $1\leq k\leq n-2$, let $u'(n,k)=u(n,k)-u(n,k+2)$ and $v'(n,k)=v(n,k)-v(n,k+2)$. It is easy to see that $u'(n,k)=n-k$ for all $1\leq k\leq n-2$ and $n-k$ is even. Since $\frac{\binom{n}{j+1+k}}{\binom{n}{j+1}}\geq\frac{\binom{n}{j+k+2}}{\binom{n}{j}}$ if $j\geq\frac{n-k}{2}-1$, $v'(n,k)\geq 0$. It is also clear that $(n+1)W'(n,k)=(n+1)\{W(n,k)-W(n,k+2)\}=u'(n,k)-v'(n,k)$. Now we will prove that for fixed n, there exists a k'_n such that $v'(n,k)\geq u'(n,k)$ if $k\leq k'_n$ and $v'(n,k)<u'(n,k)$ if $k'_n<k<n-2$. Since $u'(n,k)$ is a linear function in k, it is sufficient to prove that $v'(n,k)$ is convex in k, i.e., to prove that $v'(n,k)+v'(n,k+4)\geq 2v'(n,k+2)$. It is equivalent to proving that $v(n,k)-3v(n,k+2)+3v(n,k+4)-v(n,k+6)\geq 0$. After

simplification,

$$v(n,k)-3v(n,k+2)+3v(n,k+4)-v(n,k+6)$$

$$=\sum_{(n-k)/2}^{n-k}(2j+k-n)\left\{\frac{\binom{n}{j+k}}{\binom{n}{j}}-3\frac{\binom{n}{j+k+1}}{\binom{n}{j-1}}+3\frac{\binom{n}{j+k+2}}{\binom{n}{j-2}}-\frac{\binom{n}{j+k+3}}{\binom{n}{j-3}}\right\}.$$

It is easy to see that, to show $v(n,k)-3v(n,k+2)+3v(n,k+4)-v(n,k+6)\geq 0$, it is sufficient to show that for $n-k$ even, $n\geq 10,\ n-k\geq 6$, and $\frac{(n-k)}{2}\leq j\leq n-k-3$,

$$\frac{\binom{n}{j+k}}{\binom{n}{j}}-3\frac{\binom{n}{j+k+1}}{\binom{n}{j-1}}+3\frac{\binom{n}{j+k+2}}{\binom{n}{j-2}}-\frac{\binom{n}{j+k+3}}{\binom{n}{j-3}}\geq 0.$$

Since for $j=n-k-2,\ n-k-1,\ n-k$,

$$\frac{\binom{n}{j+k}}{\binom{n}{j}}-3\frac{\binom{n}{j+k+1}}{\binom{n}{j-1}}+3\frac{\binom{n}{j+k+2}}{\binom{n}{j-2}}-\frac{\binom{n}{j+k+3}}{\binom{n}{j-3}}\geq 0,$$

it is sufficient to show that for $n-k$ even, $n\geq 10,\ n-k\geq 6$, and $\frac{(n-k)}{2}\leq j\leq n-k-3$,

$$\frac{\binom{n}{j+k}}{\binom{n}{j}}-3\frac{\binom{n}{j+k+1}}{\binom{n}{j-1}}+3\frac{\binom{n}{j+k+2}}{\binom{n}{j-2}}-\frac{\binom{n}{j+k+3}}{\binom{n}{j-3}}=\frac{(j-3)!(n-j)!}{(j+k+3)!(n-j-k)!}h(j,k,n)\geq 0,$$

where

$$h(j,k,n)=j(j-1)(j-2)(j+k+1)(j+k+2)(j+k+3)$$

$$-3(j-1)(j-2)(n-j+1)(j+k+2)(j+k+3)(n-k-j)$$

$$+3(j-2)(n-j+1)(n-j+2)(j+k+3)(n-k-j)(n-k-j-1)$$

$$-(n-j+1)(n-j+2)(n-j+3)(n-k-j)(n-k-j-1)(n-k-j-2).$$

Let $n-k=2y,\ j=y+x,\ k=z,n=2y+z$; then

$$h(j,k,n)=h(y+x,z,2y+z)=8(8x^3+x)y^3+\{12(8x^3-4x^2+x)z+12(8x^3-12x^2+x)\}y^2$$

$$+\{12(4x^3-4x^2+x)z^2+48(2x^3-4x^2+x)z+4(14x^3-36x^2+13x)\}y$$

$$+\{4(2x^3-3x^2+x)z^3+12(2x^3-5x^2+2x)z^2+4(7x^3-21x^2+11x)z+12(x^3-3x^2+24x)\}.$$

Since $x=0,1,2,\ldots,y-3$, $y=3,4,\ldots$, $z=2,4,\ldots$, it is easy to check that $h(y+x,z,2y+z)\geq 0$. The proof of Theorem 7 now is complete. □

We can use Theorem 7 to identify k_n by comparing $W(n,k)$ and $W(n,k+2)$. Once we find k_1 and k_2 such that $W(n,2k_1-1)\geq W(n,2k_1+1)$ and $W(n,2k_2)\geq W(n,2k_2+2)$, then $k_n=2k_1-1$ if $W(n,2k_1-1)\geq W(n,2k_2)$ and $k_n=2k_2$ if $W(n,2k_1-1)<W(n,2k_2)$. The problem here is that we still don't know how many values of $W(n,k)$ we will have to compute. Fortunately the next theorem gives a lower

bound and an upper bound for k_n, which helps us to reduce the amount of required computation to identify k_n.

THEOREM 8. *Let k_n, $\underline{k}_n$, and $\overline{k}_n$ be as defined above. Then for all $n \geq 1$, $\underline{k}_n \leq k_n \leq \overline{k}_n$.*

Proof. We will give the proof for $k_n \leq \overline{k}_n$ since the proof for $\underline{k}_n \leq k_n$ is similar. We also give the proof only for the case when $n-k$ is even since the proof for the case when $n-k$ is odd is also similar. We will assume that $n \geq 1{,}000$ since Table 2 above reveals that Theorem 8 holds for $n \leq 1{,}000$. By Theorem 7, if $v'(n,k) < u'(n,k) = n-k$ and $v'(n,k-1) < u'(n,k-1) = n-k+1$, then $k_n \leq k$. Now

$$v'(n,k) = v(n,k) - v(n,k+2) = \frac{(n-k)}{\binom{n}{k}} + \sum_{j=k+1}^{(n+k)/2} (n+k-2j) \left\{ \frac{\binom{n}{j-k}}{\binom{n}{j}} - \frac{\binom{n}{j-k-1}}{\binom{n}{j+1}} \right\}$$

$$= \frac{(n-k)}{\binom{n}{k}} + (n+1) \sum_{j=k+1}^{(n+k)/2} (n+k-2j)^2 \frac{j!\,(n-j-1)!}{(j-k)!\,(n-j+k+1)!}$$

$$< \frac{(n-k)}{\binom{n}{k}} + (n+1) \sum_{j=k+1}^{(n+k)/2} \frac{(n+k-2j)^2}{(n-j)(n+k+1-j)} \left(\frac{j}{n+k-j} \right)^k$$

$$< \frac{(n-k)}{\binom{n}{k}} + \frac{4(n+1)}{(n-k)(n+k)} \sum_{j=k+1}^{(n+k)/2} (n+k-2j)^2 \left(\frac{j}{n+k-j} \right)^k$$

$$\approx \frac{(n-k)}{\binom{n}{k}} + \frac{4(n+1)(n+k)^2}{(n-k)} \int_{1/2}^{1} (2x-1)^2 (1-x)^k x^{-k} dx.$$

For $2 \leq k \leq n-2$ and n large, $\frac{(n-k)}{\binom{n}{k}}$ is negligible. Also notice that

$$\int_{1/2}^{1} (2x-1)^2 (1-x)^k x^{-k} dx = \lim_{m\to\infty} \sum_{j=1}^{2mk} \frac{1}{4mk} \frac{j^2}{4m^2k^2} \left(\frac{2mk-j}{2mk+j} \right)^k$$

$$\leq \lim_{m\to\infty} \sum_{j=1}^{2mk} \frac{1}{16k^3} \frac{j^2}{m^3} e^{-j/m} \leq \lim_{m\to\infty} \frac{1}{16k^3} \frac{1}{m^3} \frac{(e^{2/m} + e^{1/m})}{(e^{1/m}-1)^3} = \frac{1}{8k^3}.$$

Therefore, $v'(n,k) = v(n,k) - v(n,k+2) < \frac{(n+1)(n+k)^2}{2(n-k)k^3}$. Now if $(n+1)(n+k)^2 \leq 2(n-k)^2k^3$, then $v'(n,k) < \frac{(n+1)(n+k)^2}{2(n-k)k^3} \leq (n-k)$ and $k_n \leq k$ and this completes the proof of Theorem 8. ∎

Table 3 provides some numerical values of $\underline{k}_n, k_n, k_n^*, \overline{k}_n, W(n,k_n)$, and $\frac{(n-k_n+1)^2}{4(n+1)}$ for $n = 100, 200, 300, \ldots, 3{,}000$, where $k_n^* =$ the integer part of $(\frac{n}{2})^{1/3}$.

By Theorems 7 and 8, we can identify the optimal k_n very quickly. From the proof of Theorem 5, $W(n,k) = \frac{1}{n+1}\{\frac{(n-k+2)(n-k)}{4} - \frac{n^2c_k}{(k+1)(k+2)}\}$ if $n-k$ is even and $W(n,k) = \frac{1}{n+1}\{\frac{(n-k+1)^2}{4} - \frac{n^2c_k}{(k+1)(k+2)}\}$ if $n-k$ is odd, where c_k is a constant between $\frac{1}{2}$ and 1. Therefore, the optimal $k_n = d_n n^{1/3}$, where d_n is a constant less than 1.

TABLE 3
$\underline{k}_n, k_n, k_n^*, \overline{k}_n, W(n, k_n)$ *and* $\frac{(n-k_n+1)^2}{4(n+1)}$.

n	$\underline{k}_n$	k_n	k_n^*	$\overline{k}_n$	$W(n, k_n)$	$\frac{(n-k_n+1)^2}{4(n+1)}$
100	2	3	3	4	22.59	23.77
200	2	4	4	5	46.86	48.27
300	3	5	5	6	71.37	72.77
400	3	6	6	6	95.95	97.27
500	3	6	6	7	120.62	122.27
600	4	7	7	7	145.30	146.77
700	4	7	7	8	170.05	171.77
800	4	7	7	8	194.80	196.77
900	4	8	8	8	219.56	221.27
1,000	4	8	8	9	244.37	246.27
1,100	5	8	8	9	269.18	271.26
1,200	5	8	8	9	293.99	296.26
1,300	5	9	8	9	318.81	320.77
1,400	5	9	8	9	343.65	345.76
1,500	5	9	9	10	368.50	370.76
1,600	5	9	9	10	393.35	395.76
1,700	5	9	9	10	418.20	420.76
1,800	6	10	9	10	443.06	445.26
1,900	6	10	9	10	467.93	470.26
2,000	6	10	10	11	492.81	495.26
2,100	6	10	10	11	517.69	520.26
2,200	6	10	10	11	542.56	545.26
2,300	6	10	10	11	567.44	570.26
2,400	6	11	10	11	592.33	594.76
2,500	6	11	10	11	617.22	619.76
2,600	6	11	10	11	642.12	644.76
2,700	6	11	11	12	667.02	669.76
2,800	7	11	11	12	691.91	694.76
2,900	7	11	11	12	716.81	719.76
3,000	7	11	11	12	741.71	744.76

By Theorem 8, the constant d_n is approximately between $\frac{1}{2}$ and $(\frac{1}{2})^{1/3}$. From Table 3, it seems that k_n = the integer part of $\{\frac{1}{2} + (\frac{n}{2})^{1/3}\}$. However, we do not have a proof yet. We can start with k = the integer part of $\{\frac{1}{2} + (\frac{n}{2})^{1/3}\}$ and compare $W(n,k), W(n,k+2)$ and $W(n,k+1), W(n,k+3)$. Then we either increase k by 1 or decrease k by 1. By this procedure, we can identify k_n very quickly. For example, even when $n = 100{,}000$, we need at most 14 comparisons to identify the optimal k_n. Table 3 also confirms Theorem 8, which implies that $k_n \to \infty$ as $n \to \infty$ even though $k_n \to \infty$ very slowly.

THEOREM 9. *$k_n \to \infty$ as $n \to \infty$.*

Although we are not able to prove that k_n is nondecreasing in n, we have the following weaker theorem, which is interesting and useful to identify k_n.

THEOREM 10. *For all $n \geq 1$, $|k_{n+1} - k_n| \leq 1$.*

Proof. By Theorem 7, it is sufficient to prove $v(n+1,k-1) - v(n+1,k) \geq v(n,k) - v(n,k+1) \geq v(n+1,k+1) - v(n+1,k+2)$. We will give only the proof for $v(n,k) - v(n,k+1) \geq v(n+1,k+1) - v(n+1,k+2)$ since the proof for $v(n+1,k-1) - v(n+1,k) \geq v(n,k) - v(n,k+1)$ is similar. We will assume that

$n+k$ is even since the proof for the case when $n+k$ is odd is similar. Notice that

$$
\begin{aligned}
&v(n,k)-v(n,k+1)\\
&=\sum_{j=k}^{(n+k)/2-1}(n+k-2j)\frac{\binom{n}{j-k}}{\binom{n}{j}}-\sum_{j=k}^{(n+k)/2-1}(n+k-1-2j)\frac{\binom{n}{j-k}}{\binom{n}{j+1}}\\
&=\sum_{j=k}^{(n+k)/2-1}\frac{\binom{n}{j-k}}{\binom{n}{j}}\left\{(n+k-2j)-(n+k-1-2j)\frac{j+1}{n-j}\right\}
\end{aligned}
$$

and

$$
\begin{aligned}
&v(n+1,k+1)-v(n+1,k+2)\\
&=\sum_{j=k+1}^{(n+k)/2}(n+k+2-2j)\frac{\binom{n+1}{j-k-1}}{\binom{n+1}{j}}-\sum_{j=k+1}^{(n+k)/2}(n+k+1-2j)\frac{\binom{n+1}{j-k-1}}{\binom{n+1}{j+1}}\\
&=\sum_{j=k}^{(n+k)/2-1}(n+k-2j)\frac{\binom{n+1}{j-k}}{\binom{n+1}{j+1}}-\sum_{j=k}^{(n+k)/2-1}(n+k-1-2j)\frac{\binom{n+1}{j-k}}{\binom{n+1}{j+2}}\\
&=\sum_{j=k}^{(n+k)/2-1}\frac{\binom{n+1}{j-k}}{\binom{n+1}{j+1}}\left\{(n+k-2j)-(n+k-1-2j)\frac{j+2}{n-j}\right\}\\
&=\sum_{j=k}^{(n+k)/2-1}\frac{\binom{n}{j-k}}{\binom{n}{j}}\left\{(n+k-2j)-(n+k-1-2j)\frac{j+2}{n-j}\right\}\frac{j+1}{n+k+1-j}.
\end{aligned}
$$

Since $1\le j\le \frac{n+k}{2}-1$, $\frac{j+1}{n+k+1-j}<1$. Therefore, $v(n,k)-v(n,k+1)\ge v(n+1,k+1)-v(n+1,k+2)$ and this completes the proof of Theorem 10. □

From our computation, we notice that k_n is nondecreasing in n. If this statement is true, we can further reduce the computation for identifying k_n. However, we do not have a proof for this statement either. It is worthwhile to point out that both "the k_n^* in the hole drawing policy" and "the $\overline{k}_n$ in the hole drawing policy" are also asymptotically optimal. However, we do not know how big the difference between $W(n,k_n)$ and $W(n,k_n^*)$ will be. If the difference between $W(n,k_n)$ and $W(n,k_n^*)$ is bounded, then we can just use "the k_n^* in the hole drawing policy."

4. A random coin tossing problem. In this section, we will show that our urn problem is in fact equivalent to the following coin tossing problem, which can be described as follows: A player is given a coin and is allowed to toss the coin at most n times, but can stop any time he or she wishes. The player gets a $+1$ each time a head is tossed and a -1 each time a tail is tossed. The player does not know the probability "p" of getting a head on each toss but knows that p has a uniform distribution over the interval $[0,1]$.

For each positive integer n and integers $0 \le j \le k \le n$, let $G(n,k,j)$ be the player's additional (conditional) expected value at the stopping time when he or she uses an optimal stopping rule for the remaining game given that the player has tossed the coin k times and j of which are heads.

LEMMA 2. *If* $0 \le j \le k \le n-1$, *then*

$$G(n,k,j) = \max\left\{0, \frac{2j-k}{k+2} + \frac{j+1}{k+2}G(n,k+1,j+1) + \frac{k-j+1}{k+2}G(n,k+1,j)\right\}.$$

By mathematical induction, it is easy to show that for fixed n and k, $G(n,k,j)$ is increasing in j. It makes sense to define j_{nk} to be the smallest j such that $G(n,k,j) > 0$. The optimal stopping rule can then be stated as follows: If the player did not stop earlier and has tossed the coin k times, j of which are heads, then the player should continue to toss as long as $j \ge j_{nk}$ unless $k = n$. It is clear that $E(n,k,j) = G(n,k,j)$ for all $0 \le j \le k \le n$ and $n \ge 1$, so this random coin tossing problem is equivalent to the random version of Shepp's urn scheme problem.

For each nonnegative integer k, "the k in the hole stopping rule" says the player will continue to toss the coin if the number of tails tossed is still less than k + the number of heads tossed. Let $H(n,k)$ be the expected value of the game when the player uses "the k in the hole stopping rule."

THEOREM 11. *For all* $1 \le k \le n$, $H(n,k) = W(n,k)$.

Proof. It is sufficient to show that if $n-k$ is even,

$$H(n,k) = \frac{1}{n+1}\left\{\frac{(n-k+2)(n-k)}{4} - \sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\},$$

and if $n-k$ is odd,

$$H(n,k) = \frac{1}{n+1}\left\{\frac{(n-k+1)^2}{4} - \sum_{j=(n-k+1)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}.$$

We give the proof for the case when $n-k$ is even; when $n-k$ is odd the proof is similar. For each $j = 0,1,2,\ldots,n$, let $H(n,k,j)$ be the value at the stopping time when the player uses "the k in the hole stopping rule" assuming that there are j heads in n tosses. Then it is clear that $H(n,k,j) = 2j-n$ if $j > n-k$,

$$H(n,k,j) = (2j-n)\left\{1 - \binom{n}{k+j}\bigg/\binom{n}{k}\right\} - k\binom{n}{k+j}\bigg/\binom{n}{k}$$

if $(n-k)/2 \le j \le n-k$ (by the reflection principle), and $H(n,k,j) = -k$ if $j < (n-k)/2$. For each $j = 0,1,2,\ldots,n$, let $P(j)$ be the probability of getting j heads in n tosses. Given that the probability of getting a head in a toss is p, $P(j) = \binom{n}{j}p^j(1-p)^{n-j}$. Hence

$$\begin{aligned} H(n,k) &= \int_0^1 \sum_{j=0}^{n} H(n,k,j)P(j)dp \\ &= \frac{1}{n+1}\left\{\frac{(n-k+2)(n-k)}{4} - \sum_{j=(n-k)/2}^{n-k}(2j+k-n)\frac{\binom{n}{k+j}}{\binom{n}{j}}\right\}. \end{aligned}$$

Therefore, $H(n,k) = W(n,k)$. □

Acknowledgments. We would like to sincerely thank the referee for invaluable suggestions and comments which improved the presentation of this paper significantly. We also want to thank Professor Donald Berry for mentioning this Bayesian (random) version of Shepp's urn scheme to us. R. Chen would like to express his sincere thanks to the Department of Mathematics of Tamkang University of Taiwan, especially to Professor S. J. Hu.

REFERENCES

[1] W. M. Boyce, *Stopping rules for selling bonds*, RAND J. Econ., 1 (1970), pp. 27–53.
[2] W. M. Boyce, *On a simple optimal stopping problem*, Discrete Math., 5 (1973), pp. 297–312.
[3] R. W. Chen and F. K. Hwang, *On the values of an* (m, p) *urn*, Congr. Numer., 41 (1984), pp. 75–84.
[4] R. W. Chen, A. M. Odlyzko, L. A. Shepp, and A. Zame, *An optimal acceptance policy for an urn scheme*, SIAM J. Discrete Math., 11 (1998), pp. 183–195.
[5] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd Edition, John Wiley & Sons, New York, 1967.
[6] H. W. Gould, *Combinatorial Identities: A Standardized Set of Tables Listing* 500 *Binomial Coefficient Summations*, Henry W. Gould, Morgantown, WV., 1972.
[7] N. L. Johnson and S. Kotz, *Urn Models and Their Applications*, John Wiley & Sons, New York, 1977.
[8] L. A. Shepp, *Explicit solutions to some problems of optimal stopping*, Ann. Math. Statist., 40 (1969), pp. 993–1010.

Copyright of SIAM Journal on Discrete Mathematics is the property of Society for Industrial and Applied Mathematics. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.