EXPERIMENTAL RESULTS AND GENERALIZATIONS FOR THE STOPPING PROBLEM OF SHEPP'S URN

RICHARD VOEPEL

Abstract. In this paper we revisit the stopping problem of Shepp's Urn, extending the body of knowledge concerning moments of urn random variables by adapting a methodology encountered in Medina and Zeilberger's paper [7]. We then offer generalizations of these random variables by altering the strategy used to decide when to stop, and provide an experimental analysis of the results.

1. Preamble

In [8] L.A. Shepp posed the following stopping problem:

Problem 1.1 (Shepp's Urn)**.** An urn contains a known number of balls with value -1, say m, and a known number of balls with value $+1$, say p. If you are allowed to draw from this urn without replacement as many times (including 0) as you desire, which m and p have drawing strategies where the expected total value of the balls drawn is positive, what is the maximum expected value obtainable, and which strategy attains it?

Such a problem is of interest due to its links to the 'ESP problem' posed by Breiman [1], its relationship to a larger class of 'random urn problems' studied in mathematical finance [3], and more generally as an intriguing question by itself.

After introducing the problem, Shepp concluded that an optimal strategy for playing these urns is to draw so long as the maximum expected value (often simply called the value $V(m,p)$) is positive, and that these values could be computed by a simple recursive formula. Shepp also proved some basic monotonicity results for the values of related urns, as well as an asymptotic formula for $\beta(p)$, the largest m such that $V(m,p) \geq 0$. Boyce then elaborated on these results in [2], improving the monotonicity bounds, providing an exact algorithm for computing $\beta(p)$, and tabulating data on this 'stopping boundary' for p up to 100.

Work has continued on Shepp's Urn in recent years, providing generalizations to the original problem by introducing new parameters (a fixed starting fortune for the player) [4], reducing the amount of information available (the number of positive balls is uniformly random) [6], or giving extra layers of decision making to the player (deciding before hand whether or not to accept or reject the next ball drawn) [5]. However, little work seems to have been done on extending knowledge of the original problem purely in the context of the higher moments of these urns. There is a great deal of information on expectation, but what about variance, skewness, or kurtosis? What about a complete description of the probability distribution?

Author addresses: Richard Z. Voepel, Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA
http://www.math.rutgers.edu/∼rzv2/ : rzv2@math.rutgers.edu .

In this paper we attempt to address these questions by adapting a methodology encountered in Medina and Zeilberger's paper [7] to construct probability generating functions for the random variables associated with playing Shepp's strategy on the urns with m minus balls and p plus balls, denoted by $U(m,p)$. By having access to these probability generating functions we may (in principle) compute any moment for any urn random variable, giving insight beyond what expectation can reasonably provide for the question of whether or not to play an urn. We then generalize our study of Shepp's Urn by choosing new stopping strategies, based not on the expected value, but rather on an arbitrary functional on the probability generating function. Interpreting the value of each ball monetarily, we choose to study strategies that are shortsightedly risk averse in the following sense: if the probability of losing d or more dollars while playing an urn is greater than q, or if the probability of gaining at least 1 dollar is 0, then stop playing.

2. Beyond Expectation

For Shepp's Urn, there is a very simple recursive formula for the computation of $V(m,p)$. First, note that $V(m,0)=0$, since a player would never want to draw, and $V(0,p)=p$, since a player would always want to draw. Confronted with an urn $U(m,p)$, the player can compute the expected value of drawing at least one ball:

$$\mathbb{E}(U(m,p)) = \frac{m}{m+p}\left(-1+V(m-1,p)\right) + \frac{p}{m+p}(1+V(m,p-1))$$

The expected value of an urn is the probability that you draw a plus or minus ball times *the value of the resulting urn*, plus or minus one accordingly. In other words, we assume that the future actions of the player are in line with the strategy we are computing, in this case the strategy which optimizes expectation. Thus, we know the value of $V(m,p)$; it is either the expected value of $U(m,p)$ or 0, whichever is larger.

$$V(m,p) = \max(\mathbb{E}(U(m,p)), 0) \tag{2.1}$$

We may write equation (2.1) in a slightly different form to emphasize the dependence of Shepp's strategy on the expectation of $U(m,p)$:

$$V(m,p) = \begin{cases} \mathbb{E}(U(m,p)) & \text{if } \mathbb{E}(U(m,p)) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

This recursive formula loses some information, however. Instead of tracking just the value of the urns $V(m,p)$, we may track the probability generating functions in the exact same recursive manner. Let us define $g(x;m,p)$ to be the Laurent polynomial such that coefficient of x^k is the probability that, under play governed by Shepp's strategy, the player earns the value k. By identical reasoning to the base cases for $V(m,p)$, we know that $g(x;m,0)=1$, since the player would never want to draw, and $g(x;0,p)=x^p$, since the player would always want to draw. Confronted with an urn $U(m,p)$, the player can compute an auxiliary probability generating function $h(x;m,p)$ associated with the strategy "draw one ball, and then obey Shepp's strategy":

$$h(x;m,p) = \frac{m}{m+p}(x^{-1}\cdot g(x;m-1,p)) + \frac{p}{m+p}(x^{1}\cdot g(x;m,p-1))$$

This probability generating function encodes exactly what was promised in the same manner as above; after one draw, the strategy we are attempting to compute dictates our possible outcomes. Thus, we know what $g(x;m,p)$ must be:

$$g(x;m,p) = \begin{cases} h(x;m,p) & \text{if } \left.\frac{\mathrm{d}}{\mathrm{d}x}h(x;m,p)\right|_{x=1} > 0 \\ 1 & \text{otherwise} \end{cases} \tag{2.3}$$

If the expected value of the random variable dictated by our auxiliary strategy is positive, we have in fact found the probability generating function for $U(m,p)$ under Shepp's strategy, and otherwise we should not play the urn.

Armed with a complete description of the urn random variables associated with Shepp's strategy, we may compute any moments we desire. The reader may deem some computations more interesting than others, and may utilize the Maple package produced along with this paper in order to investigate urns of their choosing (see Section 5). We will focus on one case here: linear families of urns of the form $U(k_1 m, k_2 m)$.

3. Experimental Results on Higher Moments

In attempting to understand the higher moments of the urn random variables given by playing Shepp's strategy, our approach is rather straight forward; since we have a recursive formula for the probability generating functions $g(x;m,p)$ for these random variables, we may compute the moments by use of the rather basic fact that the r^{th} moment about the mean (or the central moment) m_r is a function of the so called *factorial moments about the mean*:

$$m_r = \sum_{k=1}^{r} S(r,k) \frac{d^k}{dx^k}\left(\frac{g(x;m,p)}{x^{\mu}}\right)$$

Here $S(r,k)$ are the Stirling numbers of the second kind, and μ is the expected value of $g(x;m,p)$. Furthermore, since central moments beyond the variance are typically proportional to a power of the variance, we will focus primarily on the *standardized moments* $\hat{m}_r = m_r/\sqrt{m_2^r}$.

So which urns deserve special attention? The answer is purely a matter of taste or convenience; for the purposes of this paper, we choose to focus on linear families of urns, that is to say urns of the form $U(k_1 m, k_2 m)$ for some choice of naturals $k_2 \geq k_1 \geq 1$. Other interesting families, particularly the family $U(\beta(p),p)$, will be amenable to this experimental analysis, but we will not address those cases here. Using the Maple code produced with this paper, we computed probability generating functions of urn random variables $U(m,p)$ for $m+p \leq 1450$, and have included a small selection of data in Table 1 for five linear families of urns, with $(k_1,k_2) = (1,3),(1,2),(2,3),(3,4)$ and $(1,1)$.

While the raw data obviously provides a more complete picture, it is almost reasonable just from the table below to guess that along linear families the standardized moments converge to some limit, at least for the lower moments; $\hat{m}_5$ and larger are very sensitive to the choice of (k_1,k_2), and take a great deal longer to settle down. This experimental analysis leaves us with an interesting, if difficult, open problem.

Table 1.

$U(k_1m, k_2m)$	(k_1, k_2)	$\hat{m_3}$	$\hat{m_4}$	$\hat{m_5}$
$U(1,3)$	$(1,3)$	1.154701	2.333333	3.849001
$U(120,360)$	$(1,3)$	1.830120	6.471945	23.982616
$U(480,1440)$	$(1,3)$	1.859808	6.676403	25.322562
$U(1,2)$	$(1,2)$	0.707107	1.500000	1.767767
$U(120,240)$	$(1,2)$	1.633647	5.725296	19.487670
$U(725,1450)$	$(1,2)$	1.727688	6.314074	23.161534
$U(2,3)$	$(2,3)$	0	1	0
$U(120,180)$	$(2,3)$	1.435910	4.795781	14.187846
$U(960,1440)$	$(2,3)$	1.680557	6.136094	21.834381
$U(3,4)$	$(3,4)$	0.537288	2.246558	3.206770
$U(120,160)$	$(3,4)$	1.269851	4.058409	10.548121
$U(1080,1440)$	$(3,4)$	1.648803	5.914477	20.400577
$U(1,1)$	$(1,1)$	0	1	0
$U(120,120)$	$(1,1)$	0.024085	1.547509	0.142672
$U(1450,1450)$	$(1,1)$	0.033575	1.540495	0.174673

Problem 3.1. Provide rigorous asymptotics for the sequence of standardized moments $\hat{m_r}$ of the urn random variables $U(k_1m, k_2m)$.

If these moments do converge, it would be clear that they do not define the standard Gaussian random variable, as the standardized moments do not follow the very well known formula $\hat{m_r} = (r-1)!!$ for even r, and 0 for odd r. While it is much less certain, we suspect that the limiting values for the standardized moments of linear families of urns correspond to 'novel' random variables, in the sense that no typically used random variable determined completely by its moments has these values.

Finally, still under the assumption that the moments do converge, a description of the random variables at hand would be complete if we could understand the limiting behavior of the variance of the linear families. Experimental evidence does little to inform us at this point as to what the dependence of this behavior is on (k_1, k_2), but at least for $(k_1, k_2) = (1, 1)$ we can conject that $\sigma^2 = o(\sqrt{m}\ln(m)^4)$ and $\sigma^2 = \omega(\sqrt{m}\ln(m)^2)$, using the standard Landau notation. This open problem may be easier than our earlier example.

Problem 3.2. Provide rigorous asymptotics for the variance of the urn random variables $U(m, m)$.

4. New Strategies and Experimental Results

In our development of the recursive formula for the probability generating functions of Shepp's strategy on the urns $U(m,p)$, we took some pains to highlight the exact places where the strategy actually entered the computation. Looking again at equation (2.3),

$$g(x;m,p) = \begin{cases} h(x;m,p) & \text{if } \left.\dfrac{\mathrm{d}}{\mathrm{d}x}h(x;m,p)\right|_{x=1} > 0 \\ 1 & \text{otherwise} \end{cases}$$

we see that Shepp's strategy truly only appears during the evaluation of our conditional, and perhaps in the base cases where either m or p is 0. This focus on expectation can be slightly misguided in practice. Consider the scenario in which you are not able to play a given urn $U(m,p)$ many times in sequence, or the scenario in which you must actually cover any negative results of playing that urn by paying money. Should we still follow Shepp's strategy of maximizing expectation, or is there a more prudent strategy to obey? While we make no final judgments on that question, we can quite easily explore natural families of strategies without altering our computational framework.

The act of taking the expected value of $h(x;m,p)$ is, at its heart, simply the application of some functional to the probability generating function. Were it not for the stated goal of finding an expectation maximizing strategy, that choice of functional is arbitrary. So the most natural extension of the Shepp Urn problem to consider is that of new strategies completely determined by some fixed functional on the space of probability generating functions. We will focus on the two parameter family of functionals given below:

$$F(g;d,q) = q - \sum_{i=-\infty}^{-d} \operatorname{coeff}(g,x,i) \tag{4.1}$$

Here, $\operatorname{coeff}(g,x,i)$ is the coefficient of the term x^i in the Laurent polynomial g. These functionals essentially compute "the probability q, offset by the probability of losing d or more dollars". This form is chosen so that when the conditional $\frac{d}{dx}h(x;m,p)|_{x=1} > 0$ is replaced by $F(g;d,q) > 0$ our new strategy reads "if the probability of losing d or more dollars is greater than q, do not play the urn". Leaving our choices of base cases for $g(x;m,p)$ unchanged from Shepp's case then completes the description of our strategy: "if the probability of losing d or more dollars while playing an urn is greater than q, or if the probability of gaining at least 1 dollar is 0, then stop playing.". We may now (in principle) compute probability generating functions for any urn under this new strategy in the same manner that we did for Shepp's strategy.

How do these shortsighted, risk averse strategies behave when compared to Shepp's strategy in the context of one time urns or when the player has limited resources to cover the potential costs of an urn? The most immediate comparison to be made is in the stopping boundary for each strategy. While Shepp's strategy has a well behaved boundary $\beta(p)$, the corresponding function $\beta_F(p)$ for these strategies exhibits erratic behavior, as can be seen in the figures below:

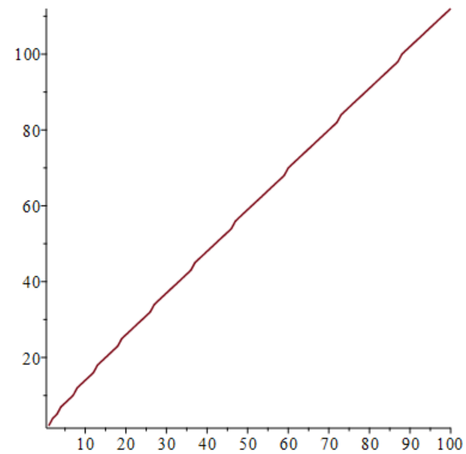

FIGURE 1. Shepp's Stopping Boundary

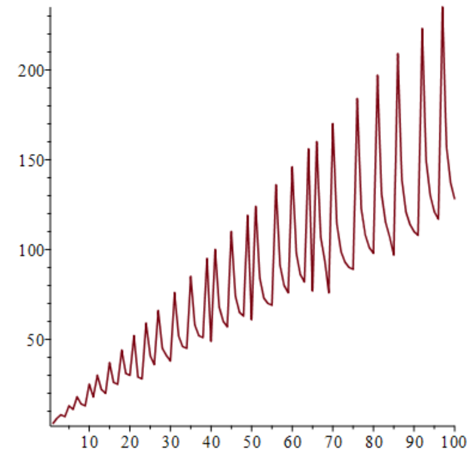

FIGURE 2. $F(g; 1, 1/2)$ Stopping Boundary

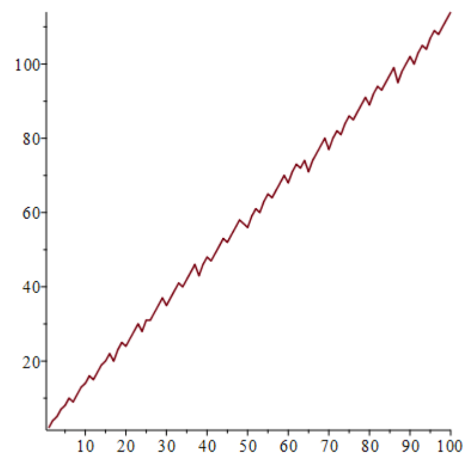

FIGURE 3. $F(g; 1, 1/4)$ Stopping Boundary

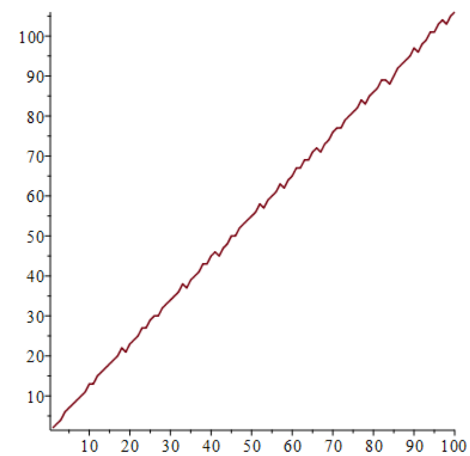

FIGURE 4. $F(g; 1, 1/8)$ Stopping Boundary

The striking feature is that for some urns and some levels of acceptable risk, *adding a plus ball* may cause a player to decline playing the resulting urn, which seems highly irrational. This is an emergent phenomenon of the recursive demands of the new strategies, and by our estimation should not be read into too much. We note that when the player is willing to accept small losses (i.e. when $d > 1$), the stopping boundary is not well defined as stated; the player will *always* draw at least once.

The primary result these strategies accomplish is a shift in the probabilities of specific payouts. For example, for the strategy given by $F(g; 2, 1/16) > 0$ on the urn $U(20,20)$, the player's probability of winning 2 or more units of value is approximately 81%, compared to an approximate 61% offered by Shepp's strategy (which has an expected value of approximately 2.3). Similarly, the strategy given by $F(g;1,1/8)>0$ gives an approximate 71% chance to win 3 or more units of value on the urn $U(50,50)$, compared to an approximate 60% offered by Shepp's strategy (which has an expected value of approximately 3.7). While there is no clear pattern as yet for which of the strategies $F(g;d,q)>0$ will optimize a particular chance of earning at least some value, generally speaking the probability mass functions behave in a manner the reader would expect; when compared to Shepp's strategy on an urn $U(m,p)$, probability mass is more concentrated, and it is concentrated around a point slightly lower than the value $V(m,p)$. While this does reduce the maximum possible earnings from any given play, it is clear that there are benefits to exploring alternate strategies in the cases where the player has a limited number of opportunities to play an urn. The reader is encouraged to use the Maple package produced with this paper to explore these strategies to their fullest.

5. Maple Code

Throughout this paper, several references have been made to Maple packages and data related to the problem at hand. These files may be downloaded from the following website: http://sites.math.rutgers.edu/~rzv2/Shepp

Instructions for how to use the Maple packages and data files are listed on that page. If you discover broken code or corrupted data, please use the author addresses listed in the footer of the first page of this paper to submit a small report.

References

1. L. Breiman, *Stopping rule problems*, in: E.F. Beckenbach, ed., Applied combinatorial mathematics (Wiley, New York, 1964)
2. W.M. Boyce, *On a Simple Optimal Stopping Problem*, Discrete Math. **5** (1973), 297–312.
3. W.M. Boyce, *Stopping rules for selling bonds*, Bell J. Econ. Mgmt. Sci. **1** (1970), 27–53.
4. R.W Chen; I. Grigorescu; M. Kang, *Optimal Stopping for Shepp's Urn with Risk Aversion*, Stochastics, **87**:4 (2015) 702–722.
5. R.W. Chen; A.M. Odlyzko; L.A. Shepp; A. Zame, *An Optimal Acceptance Policy for an Urn Scheme*, Siam J. Discrete Math, **11** (1998) 183–195.
6. R.W. Chen; C. Lin; H. Wu; A. Zame, *A Random Version of Shepp's Urn Scheme*, Siam J. Discrete Math, **19**:1 (2006) 149–164.
7. L.A. Medina; D. Zeilberger, *An Experimental Mathematics Perspective on the Old, and still Open, Question of When To Stop?* Contemporary Math. v. **517** (2010) 265-274 http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/stop.html
8. L.A. Shepp, *Explicit Solutions to Some Problems of Optimal Stopping*, Ann. Math. Statist., **40**:3 (1969) 993–1010.