Experimental "Solutions" to Select Stopping Problems

Richard Voepel

Rutgers University

April 19, 2018

Outline

1. Introduction
2. Coin Flipping
3. Shepp's Urn

Definition by Example

Everyone knows what a stopping problem is, even if they don't know that they know!

Definition by Example

Everyone knows what a stopping problem is, even if they don't know that they know!

- How long do you search for an employee, and when do you simply settle and hire an individual for your open job position?

Definition by Example

Everyone knows what a stopping problem is, even if they don't know that they know!

- How long do you search for an employee, and when do you simply settle and hire an individual for your open job position?
- How long do you keep your house on the market, and when do you simply accept an offer and sell?

Definition by Example

Everyone knows what a stopping problem is, even if they don't know that they know!

- How long do you search for an employee, and when do you simply settle and hire an individual for your open job position?
- How long do you keep your house on the market, and when do you simply accept an offer and sell?
- For holders of American options, how long do you keep the option, and when should you simply sell for the predetermined price?

Definition by Definition

Definition (Discrete Time Stopping Problems)

A (discrete time) stopping problem is composed of the following parts:

- A sequence $\{X_i\}$ of random variables, whose joint distribution is known.
- A sequence $\{r_i\}$ of reward functions, where the function r_i accepts as arguments the observed values of X_1 through X_i, i.e. $r_i = r_i(x_1, \ldots, x_i)$.

Given these parts, the "goal" of a stopping problem is to decide, after observing the value of the random variable X_i, whether to accept the reward function r_i (to stop) or to observe the value of the next random variable X_{i+1} (to go).

Remarks

- Almost all texts add to the goal that the so called "stopping rule" maximize the expected reward.
- Some texts allow for a payout on 0 observations. This is equivalent to X_0 being a constant random variable.
- Some texts allow for unbounded observations, or even infinite observations. We will mostly ignore these cases.
- The stopping rule can depend on anything thus far observed, any rewards that have been passed up, and can even be a random variable itself!

Our Two Problems

Problem (Coin Flipping)

Flip a coin any positive number of times. When you decide to stop, you earn $\frac{h}{h+t}$ *units of money, where* h *is the number of heads and* t *is the number of tails thus far observed. When should you stop flipping?*

Problem (Shepp's Urn)

An urn is filled with some number p *of* \$1 *bills, and some number* m *of* \$1 *"anti-bills". How many times (including* 0*) should you draw an object from this urn?*

First Considerations

Let's begin thinking about our coin flipping problem:

- Since we have to flip at least once, what is the first decision in our stopping rule?

First Considerations

Let's begin thinking about our coin flipping problem:

- Since we have to flip at least once, what is the first decision in our stopping rule?
- If the first flip is a heads, it is obvious that we should stop, as we will receive the maximum possible reward.
- If the first flip is a tails, it is obvious that we should go, since stopping would pay 0 units.

Second Considerations

What should we do when the sequence of flips turns up TH?

- If we stop, we will earn 0.5 units – no bad at all. But can we do better?

Second Considerations

What should we do when the sequence of flips turns up TH?

- If we stop, we will earn 0.5 units – no bad at all. But can we do better?
- By the law of large numbers, we could do better (perhaps only slightly so) by continuing. But how long would we have to go? Is it reasonable to assume we have nothing better to do with our lives than flip coins?

Medina and Zeilberger

These and other questions related to the coin flipping problem have been addressed in the paper *An Experimental Mathematics Perspective on the Old, and Still Open, Question of When to Stop?* by Medina and Zeilberger (see reference slide for more details).

Medina and Zeilberger

These and other questions related to the coin flipping problem have been addressed in the paper *An Experimental Mathematics Perspective on the Old, and Still Open, Question of When to Stop?* by Medina and Zeilberger (see reference slide for more details).

Let's recreate a portion of their arguments here, in preparation for our second problem of Shepp's Urn.

Is There a "Best" Stopping Rule?

The very first question we should ask is if we can specify a stopping rule which maximizes the expected reward for this problem.

Is There a "Best" Stopping Rule?

The very first question we should ask is if we can specify a stopping rule which maximizes the expected reward for this problem.

This is essentially an open question; there are infinitely many cases that have not been resolved, and some of them sound quite innocuous.

Is There a "Best" Stopping Rule?

The very first question we should ask is if we can specify a stopping rule which maximizes the expected reward for this problem.

This is essentially an open question; there are infinitely many cases that have not been resolved, and some of them sound quite innocuous.

What we can say for sure is that such a rule exists, and in fact can be encoded by a "stopping boundary" – a criterion that establishes threshold values for the number of heads and tails that we must either stop or go on in order to maximize our expected reward.

The Stopping Boundary

Theorem (Chow and Robbins)

There exists a sequence of naturals $\{\beta_n\}$ such that if the number of observed heads minus the number of observed tails after n coin tosses is greater than or equal to β_n, then you should stop. Else, you should go.

Proof: See the final remark in the paper *On Optimal Stopping Rules for s_n/n* by Chow and Robbins.

The Stopping Boundary

Theorem (Dvoretzky)

The stopping boundary sequence for the coin flipping problem satisfies $\beta_n/\sqrt{n} \in \Theta(1)$.

Theorem (Shepp)

The stopping boundary sequence for the coin flipping problem satisfies $\lim_{n\to\infty} \beta_n/\sqrt{n} \approx 0.83992$. *In particular, the limit exists.*

Proofs: See the respective papers on the reference slide.

Beyond Asymptotics

These results are nice, and there is some solid mathematics behind them, but they say nothing about specific cases we want to consider.

Beyond Asymptotics

These results are nice, and there is some solid mathematics behind them, but they say nothing about specific cases we want to consider.

In particular, what is the value of β_8? If you have observed the sequence TTTHHHHH of flips, should you stop or go?

Beyond Asymptotics

These results are nice, and there is some solid mathematics behind them, but they say nothing about specific cases we want to consider.

In particular, what is the value of β_8? If you have observed the sequence TTTHHHHH of flips, should you stop or go?

Should we even bother asking this question? Is it worth our time to compute β_8 when we are working in a model that implicitly assumes we are willing and able to flip coins for eternity?

A New Model

Rather than assume that we can flip coins as many times as we desire, let us add the assumption that there is a finite (but potentially large) number of coin flips available to us, say N.

A New Model

Rather than assume that we can flip coins as many times as we desire, let us add the assumption that there is a finite (but potentially large) number of coin flips available to us, say N.

If you happen to use up all of your flips you **must** stop, and as a consolation prize your reward will be $\max(1/2, h/N)$, where h is the number of observed heads in your N flips.

A New Model

Rather than assume that we can flip coins as many times as we desire, let us add the assumption that there is a finite (but potentially large) number of coin flips available to us, say N.

If you happen to use up all of your flips you **must** stop, and as a consolation prize your reward will be $\max(1/2, h/N)$, where h is the number of observed heads in your N flips.

Define $f_N(h, t)$ to be the expected reward of the bounded version of our coin flipping problem. Can we now compute this exactly?

Backwards Induction

Since our reward is now clearly defined for all $h + t = N$, we may use backwards induction to compute $f_N(h, t)$ for **all** $h + t \leq N$.

Backwards Induction

Since our reward is now clearly defined for all $h + t = N$, we may use backwards induction to compute $f_N(h, t)$ for **all** $h + t \leq N$.

Boundary: $f_N(h, N - h) = \max(1/2, h/N)$ for $0 \leq h \leq N$.
Induction:

$$f_N(h, t) = \max\left(\frac{f_N(h + 1, t) + f_N(h, t + 1)}{2}, \frac{h}{h + t}\right)$$

Results

Using this bounded time model, we can begin to say interesting things about our original problem:

Results

Using this bounded time model, we can begin to say interesting things about our original problem:

- $f_N(h,t)$ is increasing in N and bounded, so the limit exists (and happens to be equal to our expected reward for the original problem).

Results

Using this bounded time model, we can begin to say interesting things about our original problem:

- $f_N(h, t)$ is increasing in N and bounded, so the limit exists (and happens to be equal to our expected reward for the original problem).
- We may now classify not only when a particular sequence of flips is a stop or go, but how many flips you need to take in order to switch from stop to go.

Results

Using this bounded time model, we can begin to say interesting things about our original problem:

- $f_N(h,t)$ is increasing in N and bounded, so the limit exists (and happens to be equal to our expected reward for the original problem).
- We may now classify not only when a particular sequence of flips is a stop or go, but how many flips you need to take in order to switch from stop to go.
- For instance, $F(0,0) = \lim_{N\to\infty} f_N(0,0) \approx 0.79295$, and a sequence of 7 tails and 10 heads is optimal (you should stop) unless you are willing to flip your coin at least another 1404 times.

Further Statistics...

This approach provides a great deal of information about our stopping problem, but is ultimately limited by the fact that it only deals in expectations. What can we do if we want to study variance? Skewness or kurtosis? What if we need a complete description of the probability distribution?

Further Statistics...

This approach provides a great deal of information about our stopping problem, but is ultimately limited by the fact that it only deals in expectations. What can we do if we want to study variance? Skewness or kurtosis? What if we need a complete description of the probability distribution?

We simply repeat our backwards induction argument, but use probability generating functions for our boundary values rather than expected rewards!

...And New Stopping Rules

What if we aren't interested in trying to maximize our expected reward at all? Perhaps we are more interested in breaking some threshold reward, at which point we will stop, guaranteeing the payout.

...And New Stopping Rules

What if we aren't interested in trying to maximize our expected reward at all? Perhaps we are more interested in breaking some threshold reward, at which point we will stop, guaranteeing the payout.

Before we explore these questions further, let's switch to studying our second problem – Shepp's Urn.

A Whole New Billgame

Recall the stopping problem of Shepp's Urn:

Problem (Shepp's Urn)

An urn is filled with some number p of \$1 bills, and some number m of \$1 "anti-bills". How many times (including 0) should you draw an object from this urn?

A Whole New Billgame

Recall the stopping problem of Shepp's Urn:

Problem (Shepp's Urn)

An urn is filled with some number p of \$1 bills, and some number m of \$1 "anti-bills". How many times (including 0) should you draw an object from this urn?

Unlike the coin flipping problem, we are already of bounded length; we can skip straight to a recursive definition of the expected reward for any given urn.

Forwards Recursion

Consider the urn with p bills and 0 anti-bills. Clearly any stopping rule that maximizes expected reward will draw all p bills from the urn before stopping, so the expected reward is $V(0, p) = p$.

Forwards Recursion

Consider the urn with p bills and 0 anti-bills. Clearly any stopping rule that maximizes expected reward will draw all p bills from the urn before stopping, so the expected reward is $V(0, p) = p$.

Consider the urn with 0 bills and m anti-bills. Clearly any stopping rule that maximizes expected reward will simply not draw anything from the urn, so the expected reward is $V(m, 0) = 0$.

Forwards Recursion

Consider the urn with p bills and m anti-bills. How can we define its expected reward under a strategy that maximizes that value?

Forwards Recursion

Consider the urn with p bills and m anti-bills. How can we define its expected reward under a strategy that maximizes that value?

Boundary: $V(m, 0) = 0$ for all $m \geq 0$.
$V(0, p) = p$ for all $p \geq 0$.

Recursion:

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

$$V(m, p) = \max\left(0, E(m, p)\right)$$

Tabular Data

Table 1

$V(m, p)$

p (plus) \ m (minus)	0	1	2	3	4	5	6	7	8	9
9	9	8.10	7.20	6.31	5.43	4.58	3.75	2.95	2.21	1.53
8	8	7.11	6.22	5.35	4.49	3.66	2.86	2.11	1.43	0.84
7	7	6.13	5.25	4.39	3.56	2.76	2.01	1.34	0.75	0.30
6	6	5.14	4.29	3.45	2.66	1.91	1.23	0.66	0.23	0
5	5	4.17	3.33	2.54	1.79	1.12	0.55	0.15	0	0
4	4	3.20	2.40	1.66	1.00	0.44	0.07	0	0	0
3	3	2.25	1.50	0.85	0.34	0	0	0	0	0
2	2	1.33	0.67	0.20	0	0	0	0	0	0
1	1	0.50	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Been There...

This is a well studied problem; if there is a question concerning the expected reward of Shepp's Urn under the expectation maximizing strategy, chances are someone has already asked (and answered) it.

Been There...

This is a well studied problem; if there is a question concerning the expected reward of Shepp's Urn under the expectation maximizing strategy, chances are someone has already asked (and answered) it.

However, just like in the coin flipping case, utilizing such a strategy makes implicit assumptions about what we are willing and able to do. Do I want to play a given urn with positive expected reward hundreds or thousands of times in order to leverage that expectation? Will I even have the opportunity to play that urn multiple times?

...Haven't Done That Yet

Let us try and write down not just the expected reward for the maximizing strategy, let's write down the complete probability distribution, as encoded by a probability generating function.

...Haven't Done That Yet

Let us try and write down not just the expected reward for the maximizing strategy, let's write down the complete probability distribution, as encoded by a probability generating function.

Recall that a a PGF of some discrete random variable X taking values in $\mathbb{Z}$ is given by $G(z) = \sum_{x \in R(X)} p(x) z^x$ – we take the probability of each outcome x and multiply it by a formal object z^x.

...Haven't Done That Yet

Let us try and write down not just the expected reward for the maximizing strategy, let's write down the complete probability distribution, as encoded by a probability generating function.

Recall that a a PGF of some discrete random variable X taking values in $\mathbb{Z}$ is given by $G(z) = \sum_{x \in R(X)} p(x) z^x$ – we take the probability of each outcome x and multiply it by a formal object z^x.

This results in a formal laurent polynomial, which we can then manipulate in order to recover information about any moment of the corresponding random variable!

...Haven't Done That Yet

In this new language, what do our boundary cases look like?

An urn with p bills and 0 anti-bills would be emptied by a player maximizing the expected reward – there is only one outcome for the random variable corresponding to this urn, which has probability 1. Thus, $G(z) = U(0, p; z) = z^p$.

...Haven't Done That Yet

In this new language, what do our boundary cases look like?

An urn with p bills and 0 anti-bills would be emptied by a player maximizing the expected reward – there is only one outcome for the random variable corresponding to this urn, which has probability 1. Thus, $G(z) = U(0, p; z) = z^p$.

An urn with 0 bills and m anti-bills would be immediately passed over by a player maximizing the expected reward – there is only one outcome for the random variable corresponding to this urn, which has probability 1. Thus, $G(z) = U(m, 0; z) = z^0 = 1$.

Recursion for PGFs

Consider the urn with p bills and m anti-bills. How can we define its PGF under a strategy that maximizes the expected reward?

Recursion for PGFs

Consider the urn with p bills and m anti-bills. How can we define its PGF under a strategy that maximizes the expected reward?

Boundary: $U(m, 0; z) = 1$ for all $m \geq 0$.
$U(0, p; z) = z^p$ for all $p \geq 0$.

Recursion:

$$W(m, p; z) = \frac{m}{m+p}\left[U(m-1, p; z) \cdot z^{-1}\right] + \frac{p}{m+p}\left[U(m, p-1; z) \cdot z^{1}\right]$$

$$U(m, p; z) = \begin{cases} W(m, p; z) & \text{if } \left.\frac{\partial}{\partial z} W(m, p; z)\right|_{z=1} > 0 \\ 1 & \text{otherwise} \end{cases}$$

First Results

We saw previously on our table of data that $V(6,4) \approx .07 > 0$. According to our strategy, we will draw at least once from this urn – but what is the actual probability of making any money?

First Results

We saw previously on our table of data that $V(6,4) \approx .07 > 0$. According to our strategy, we will draw at least once from this urn – but what is the actual probability of making any money?

$$U(6,4;z) \approx .40000z + .37619 + .11429z^{-1} + .10952z^{-2}$$

While we do have a respectable 40% chance of making \$1, we also have a greater than 20% chance of losing **at least** \$1.

First Results

We saw previously on our table of data that $V(6,4) \approx .07 > 0$. According to our strategy, we will draw at least once from this urn – but what is the actual probability of making any money?

$$U(6,4;z) \approx .40000z + .37619 + .11429z^{-1} + .10952z^{-2}$$

While we do have a respectable 40% chance of making \$1, we also have a greater than 20% chance of losing **at least** \$1.

Now, \$1 doesn't seem like much, but what if the bills came in \$100 denominations, or \$10000? Could you afford the risk of (potentially) massive debt on the worse-than-coin-flip chance of making good money?

A Different Goal

If maximizing our expected reward can sometimes leave us with moderate probabilities of losing money, how can we go about constructing a stopping rule that guarantees those probabilities remain small?

A Different Goal

If maximizing our expected reward can sometimes leave us with moderate probabilities of losing money, how can we go about constructing a stopping rule that guarantees those probabilities remain small?

Let's consider again our recursive definition of $U(m, p; z)$:

$$W(m, p; z) = \frac{m}{m+p}\left[U(m-1, p; z) \cdot z^{-1}\right] + \frac{p}{m+p}\left[U(m, p-1; z) \cdot z^{1}\right]$$

$$U(m, p; z) = \begin{cases} W(m, p; z) & \text{if } \left.\frac{\partial}{\partial z} W(m, p; z)\right|_{z=1} > 0 \\ 1 & \text{otherwise} \end{cases}$$

A Different Goal

The piece of this recursive definition that is responsible for encoding the expected reward is the following boolean functional on the PGFs $W(m, p; z)$:

$$\left.\frac{\partial}{\partial z} W(m, p; z)\right|_{z=1} > 0$$

A Different Goal

The piece of this recursive definition that is responsible for encoding the expected reward is the following boolean functional on the PGFs $W(m, p; z)$:

$$\frac{\partial}{\partial z} W(m, p; z)\Big|_{z=1} > 0$$

If we want a different goal besides maximizing the expected reward, we need only change this conditional into some other (arbitrary) boolean functional on PGFs! We now have complete control over a wide and varied family of stopping rules.

Risk Aversion

We will focus on the two parameter family of risk averse functionals given below:

$$F(W; d, q) = q - \sum_{i=-\infty}^{-d} \text{coeff}(W, z, i) > 0$$

This functional reads "the probability of receiving a reward of $-d$ dollars or less is less than q".

Risk Aversion

We will focus on the two parameter family of risk averse functionals given below:

$$F(W; d, q) = q - \sum_{i=-\infty}^{-d} \text{coeff}(W, z, i) > 0$$

This functional reads "the probability of receiving a reward of $-d$ dollars or less is less than q".

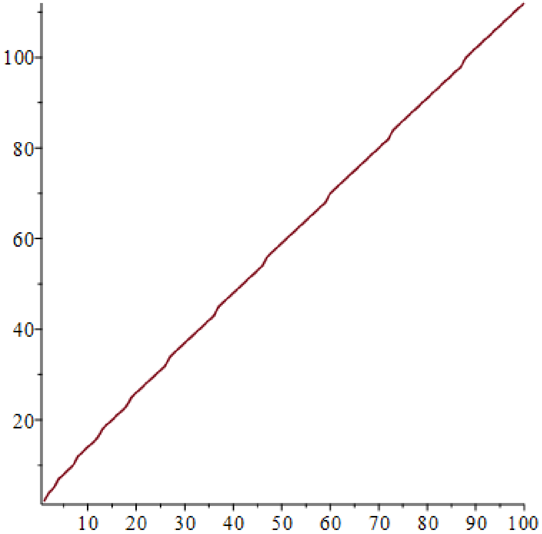
When $d = 1$ we have a well defined notion of a stopping boundary, and so we will primarily focus on this case.

Stopping Boundaries

For the expectation maximizing strategy, our stopping boundary β_p behaves quite well:

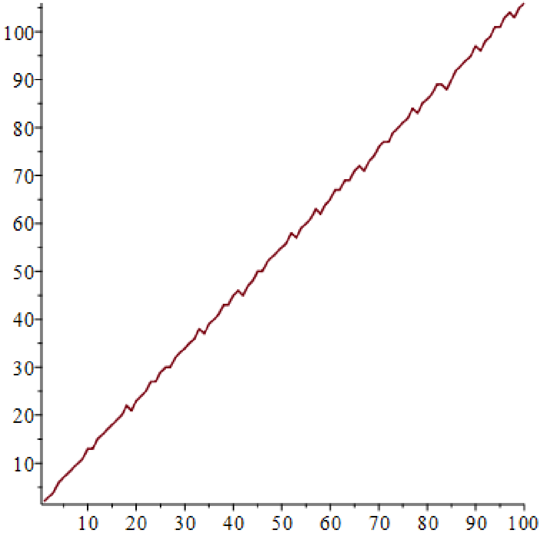
$$\beta_p \approx p + \alpha\sqrt{2p}$$

$$\alpha \approx 0.83992$$



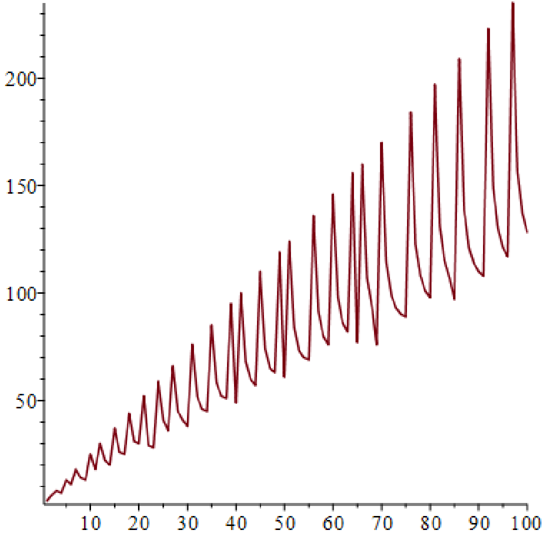
Stopping Boundaries

For the strategy which is heavily risk averse, the stopping boundary exhibits some strange phenomena...



Stopping Boundaries

...which is readily apparent if we relax the risk aversion. By adding more bills to the urn, we suddenly no longer want to draw!



Comparing Strategies

How does risk aversion compare to expectation maximization?

- For an urn with 20 bills and 20 anti-bills, Shepp's strategy has an expected reward of $2.30. The chance that this strategy will result in at least $2 is 61%.

Comparing Strategies

How does risk aversion compare to expectation maximization?

- For an urn with 20 bills and 20 anti-bills, Shepp's strategy has an expected reward of \$2.30. The chance that this strategy will result in at least \$2 is 61%.
- The same urn under the strategy given by $F(W; 2, 1/16) > 0$ results in an 81% chance of earning at least \$2.

Comparing Strategies

How does risk aversion compare to expectation maximization?

- For an urn with 20 bills and 20 anti-bills, Shepp's strategy has an expected reward of \$2.30. The chance that this strategy will result in at least \$2 is 61%.
- The same urn under the strategy given by $F(W; 2, 1/16) > 0$ results in an 81% chance of earning at least \$2.
- For an urn with 50 bills and 50 anti-bills, Shepp's strategy has an expected reward of \$3.70, with a 60% chance of earning at least \$3.

Comparing Strategies

How does risk aversion compare to expectation maximization?

- For an urn with 20 bills and 20 anti-bills, Shepp's strategy has an expected reward of \$2.30. The chance that this strategy will result in at least \$2 is 61%.
- The same urn under the strategy given by $F(W; 2, 1/16) > 0$ results in an 81% chance of earning at least \$2.
- For an urn with 50 bills and 50 anti-bills, Shepp's strategy has an expected reward of \$3.70, with a 60% chance of earning at least \$3.
- The same urn under the strategy given by $F(W; 1, 1/8) > 0$ results in a 71% chance of earning at least \$3.

I Want to Believe

This direction of study, which is nearly purely experimental in nature, has nevertheless produced some serious questions that merit serious answers:

- For Shepp's strategy, now that we have the PGFs of every urn (in principle, anyway), what kind of random variables are we looking at? Evidence suggests that the standardized moments of certain infinite families of urns converge to novel random variables.

I Want to Believe

This direction of study, which is nearly purely experimental in nature, has nevertheless produced some serious questions that merit serious answers:

- For Shepp's strategy, now that we have the PGFs of every urn (in principle, anyway), what kind of random variables are we looking at? Evidence suggests that the standardized moments of certain infinite families of urns converge to novel random variables.
- We know that we can increase our chances of making certain threshold rewards, but how can we best choose the parameters in our risk averse strategies? Are there other functionals worth considering?

I REALLY Want to Believe

And we shouldn't stop here. The techniques and ideas contained in Medina and Zeilberger's paper can basically apply to any discrete time stopping problem.

I REALLY Want to Believe

And we shouldn't stop here. The techniques and ideas contained in Medina and Zeilberger's paper can basically apply to any discrete time stopping problem.

How many problems have been declared "solved" once an expectation maximizing strategy has been found? How many of those solutions are currently being implemented in the real world? How much more efficient could we be with a more careful, experiment oriented analysis?

References

[1] W.M. Boyce.
On a Simple Optimal Stopping Problem
Discrete Math., 5:297–312, 1973.

[2] Y.S. Chow and Herbert Robbins.
On Optimal Stopping Rules for s_n/n
Ill. J. Math., 9:444–454, 1965.

[3] Aryeh Dvoretzky.
Existence and Properties of Certain Optimal Stopping Rules.
Proc. Fifth Berkeley Symp. Math. Statist. Prob., 1:441–452, Univ. of California Press, 1967.

[4] L.A. Medina and D. Zeilberger.
An Experimental Mathematics Perspective on the Old, and still Open, Question of When To Stop?
Contemporary Math., 517:265–274, 2010.

[5] Larry A. Shepp.
Explicit Solutions to Some Problems of Optimal Stopping.
The Annals of Mathematical Statistics, 40:993-1010, 1969.

[6] Julian D.A. Wiseman.
The Expected Value of $s_n/n \approx 0.79295350640$
http://www.jdawiseman.com/papers/easymath/coin-stopping.html