Knowing When To Say When:
An Expanded Description of Stopping Problems and Their Solutions

A Senior Project

presented to

the Faculty of the Department of Mathematics

California Polytechnic State University, San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Science, Mathematics

by

Eric C. Bauer

April 2010

APPROVAL PAGE

TITLE: Knowing When To Say When: An
                               Expanded Description of Stopping
                               Problems and their Solutions

AUTHOR: Eric C. Bauer

DATE SUBMITTED: April 2010

Senior Project Advisor Name: Dr. Kent E. Morrison

Mathematics Department Chair Name: Dr. Don P. Rawlings

Contents

1 Overview

In a world of infinite possibilities and limitless opportunities, how is one to know precisely the right moment – the *perfect* moment – to stop with what you've got and rest, content in the knowledge that what you hold in your hand is most likely better than anything else coming down the road?

This general concept drives an area of mathematical interest called "optimal stopping theory". This topic has roots trailing back to 1875 and the British mathematician Arthur Cayley (Ferguson, 1989), but really began garnering interest in the mathematical community in the 1950's and 1960's. From there, the topic has spread far and wide, both as a teaching tool for mathematics undergraduate students and as a research vehicle for Mathematics PhDs.

It is a topic and group of problems that puzzle amateur and academic mathematicians alike. And – luckily for me – it's a topic that bends low to allow the mathematically-challenged to climb on board. Its problems are easy to state in non-mathematician terms but near-limitless in the depth of investigation they'll spawn/support. At the end of the day, they're *games* as much as problems and as such are intrinsically adept at piquing interest.

The intent of this paper is to offer a hand down to those who are curious about climbing on board an investigation of these topics, but may not possess the background or time to understand them at first glance. It's a long-standing belief of mine that part of the beauty of mathematics lies squarely in its inherent accessibility. Certainly not to disregard its complexity, nor to understate the years of dedication and focus required to comprehend the deeper structures of mathematics in general, my intent is simply to focus on the wonder of a few simple discoveries, and to illicit an imaginary "a-ha" moment from a fictional layman reader.

2 The Secretary Game Rebranded: *The Battle for Dessert Superiority*

One of the first mathematical games pondered in the topic of optimal stopping has been known by any one of a series of names – unfortunately, all of these names have a distinct 1950's flavor and as such tend to clash with the modern ear. Such names include: The Secretary Game, The Dowry Game, and The Beauty Contest Game. To keep with 21st Century sensibilities – not to mention my own personal gastronomic proclivities – let's phrase the question in the following way:

You're out to dinner with some friends. Your plates from the main course have just been cleared, and here comes the waiter back to your table: rolling the coveted dessert tray. Each shelf of the dessert tray is crowded with an array of silver serving domes, each with some manner of dessert hidden beneath. Your friends and you decide to play a little game:

- You will be presented with a series of desserts, one at a time;

- While considering a given dessert, you must choose: keep this dessert and eat it, or pass on it and continue reviewing the remaining desserts;

- Once you pass on a dessert, you can never go back to it again;

- You know the overall number of desserts that will be available to be offered to you (i.e., there is not an infinite number of desserts, and you do know the total number of desserts on the cart);

- There's no telling how good or bad a dessert may be – if distilled down to an objective, calculable "yumminess" scale, any given dessert you encounter may exist anywhere along the spectrum from "Make You Gag At The Sight" to "Nirvana";

- After you choose a dessert, all of the remaining desserts – and their respective "yumminess" factors – will be revealed (i.e., you will then know if the dessert you have chosen is the "Yummiest of All").

Your objective in this game is simple: pick the best dessert. If the dessert you end up with is the "Yummiest of All" you are of course the winner (and can claim dessert gloating rights over all other friends in attendance). If once all desserts are revealed you realize that some other dessert (or worse yet, several desserts!) is in fact superior to the one you have chosen, you will

suffer pangs of Orderer's Remorse and experience Acute Dessert Envy and so by definition will have lost the Battle for Dessert Superiority.

(This zero-sum game of dessert-ordering happiness may seem harsh, but rest easy knowing that at the end of the experience, everyone had a great time and thought the restaurant was nice even if the service was a bit spotty and the majority opinion held that we all really do need to do this again sometime soon, and we're not just saying it this time.)

2.1 Maximizing Your Odds of Winning

At first glance, it would be easy to think that success in this task is completely random. If you were to pick a dessert at random and hope for the best, your chances of winning would be $1/n$ (i.e., there are n desserts total, and you're hoping that the 1 you've selected is the best – each of the n desserts has a $1/n$ chance of being the best of all). So that's our baseline – we know we can't get worse odds than $1/n$; the question is: *can we do better?*

2.1.1 Odds Improve the Farther You Safely Go

On your path toward improving on these paltry $1/n$ odds, the first point to take note of is that built into the fabric of the game is one factor in your favor. The base nature of the game lends itself to giving you more information the more desserts you uncover. That is to say, if the first 4 desserts you uncover are horrible, and the 5th dessert is good, then you know that picking the 5th dessert will beat the first 4 desserts, but may not be better than the remaining (n-5) desserts. So keeping this dessert has a $1/(n$-5) chance of winning, which is still not ideal but is better than $1/n$. Point being: the later in the game you get, the better your odds of winning become. As such, we should consider strategies that – at the least – help us get as far into the game (i.e. uncover as many desserts) as possible before making a decision.

2.1.2 Definition of a "Candidate"

Say the first dessert you are offered is Boston Cream pie, and you pass on it. If the second dessert offered is a piece of dry toast, you can be certain of at least one thing: the toast is definitely NOT the best dessert on the tray. Everything that follows may be even worse than the toast, but then the Boston Cream pie would be the winner after all, so you've already lost. Point being: <u>if the dessert you're currently considering isn't the best dessert of all the ones you've been shown so far, then you may as well keep searching</u>.

This concept leads to the definition of the term "candidate", as follows:

> ***A dessert being considered for eating is considered a <u>candidate</u> only if it is the best dessert encountered thus far in the game.***

As shown in the Boston Cream pie/dry toast example above, in the win/lose game of dessert superiority, one only needs to consider keeping candidates. Choosing a dessert that is not a candidate would result in your immediate loss, leading not only to a substandard dessert experience but also the scorn and ridicule of your dinner party. Being unlucky is forgivable; being dumb is not.

2.1.3 A Strategy Emerges

If you are currently considering dessert number i of a total of n desserts, and no dessert prior to the ith dessert is superior to it (i.e., if the ith dessert is a candidate), then your chances of winning with the ith pick increase as i approaches n. This fact is essentially a combination of the points made above in §2.1.1 (the farther you can get the better) and §2.1.2 (only consider candidates).

To illustrate this point, consider the following scenario. You are considering 5 desserts, the first of which is a piece of pie (no ice cream, just imitation whipped topping – a so-so dessert to be certain):

- There's a $1/5 = 20\%$ chance that the pie in position 1 is the best of the 5.

- If the second dessert is a candidate (i.e., it is better than the pie in position 1), then we know that the pie is NOT the winner, and so there's a $1/4 = 25\%$ chance that the second dessert is the best of the remaining 4.

- If we pass on the second dessert and move to the third, and the third dessert happens to be a candidate (i.e. it is better than BOTH of the first two desserts), then we know that neither of the first two desserts were winners, and there's a $1/3 = 33\%$ chance that the third is best of the remaining 3.

- If the fourth dessert is a candidate, there's a $1/2 = 50\%$ chance that it's the best of the remaining 2.

- If the fifth dessert is a candidate, then we know we've won (i.e., a 1/1 = 100% chance that it's the highest of all).

Point being: the farther along you encounter a candidate, the better your chances of winning with that candidate.

Based on that observation, we will consider strategies where we look for a candidate as far into the series as we can. However, while the odds of winning with a candidate improve the further you get into the series, the odds of actually *finding* a candidate decrease the further you go. As such, we need to find a balance: get as far as you can into the series, but not too far.

Specifically, we want to find a number of desserts r to consider, rank, and discard, then pick the next candidate we encounter from that point forward.

The name of the game is determining what the ideal value for r would be, given a number of n total desserts.

2.1.4 A Discrete Example

Let's expand our case to $n = 6$, as shown in this figure:

Not knowing anything about the desserts or how they rank against each other, we can be certain of at least one thing: that one of them is better than the other 5. We don't know which of the 6 it is, but one of them is already a winner before we even consider the first dessert. At this point, each dessert has a 1 in 6 chance of being the best dessert of all.

Let us consider the following strategy: we observe the first dessert, note it, and discard it. We then search for the next candidate – that is to say, the next dessert that is better than the first dessert.

The first dessert has a 1/6 = 16.67% chance of being the best of all. With the strategy we're considering here, we would of course lose in this case.

If the second dessert is the best dessert of all (a scenario that also has a 1/6 chance of occurring), then we would surely choose it (it would be better than the first dessert, and so a candidate) and we would win.

Things become slightly less certain in the case where the third dessert is the best dessert of all (also a 1/6 chance of this case occurring). Here, it's not enough that this dessert is best, but we also have to ensure that we pick it. In what case do we pick it? Only if it's a candidate – i.e., if dessert 2 was WORSE than dessert 1. That is to say: if dessert 2 had been better than dessert 1 (but still not the best – in this scenario we still assume dessert 3 is the best), then we would have chosen it as our candidate and we would have lost. If dessert 2 was worse than dessert 1, however, it would not have been a candidate and we would have been free to carry on to dessert 3. There are two possible scenarios here – 1 better than 2 and 2 better than 1 – so there is a 1 in 2 (50%) chance of this occurring. A 1/6 chance of dessert 3 being best, times a 1/2 chance that dessert 1 is better than dessert 2, yields a total probability of (1/6)(1/2) = 1/12 = 8.33% to win in this scenario.

If we consider the case where dessert 4 is the best of the 6, then we can see a similar scenario unfolding – we need dessert 4 to be best of all (1/6 chance) AND we need dessert 1 to be the best of desserts 1 through 3 (a 1 in 3 chance). Total probability: (1/6)(1/3) = 1/18 = 5.56%.

With desserts 5 and 6 you can easily extend this logic. Each has a 1/6 chance of being the best of the 6, and in each case you need dessert 1 to be better than all the other desserts prior to the one you're considering (1 in 4 for dessert 5; 1 in 5 for dessert 6).

Adding each case's probability up, where the Probability of Win = the Probability of the dessert being best (which is always 1/6) multiplied by the Probability of selecting it as a candidate:

Stop with Dessert #	Case in Order to Win	Probability of Win
2	Dessert 2 best of all 6 (1/6)	(1/6)(1)= 0.1667
3	Dessert 3 best of all 6 (1/6) and Dessert 1 better than Dessert 2 (1/2)	(1/6)(1/2) = 1/12 = 0.0833
4	Dessert 4 best of all 6 (1/6) and Dessert 1 better than Desserts 2 and 3 (1/3)	(1/6)(1/3) = 1/18 = 0.0556
5	Dessert 5 best of all 6 (1/6) and Dessert 1 better than Desserts 2, 3, and 4 (1/4)	(1/6)(1/4) = 1/24 = 0.0417
6	Dessert 6 best of all 6 (1/6) and Dessert 1 better than Desserts 2, 3, 4, and 5 (1/5)	(1/6)(1/5) = 1/30 = 0.0333
	Total Probability of Win	0.3806

We immediately observe that the strategy of "observe and discard the first dessert, then select the next dessert better than the first" yields a chance of winning (38.06%) far superior to the chance of winning by just picking randomly (16.67%).

2.1.5 Next Steps

Let's now adjust this strategy a bit and see how our odds of winning may change. Instead of revealing just one dessert, let's reveal and evaluate the first two before acting:

Now in this case, note that the relative ranking of dessert 1 and dessert 2 doesn't matter one bit – all we care about is that one of them is better than the other, and that we're going to evaluate all other desserts against the BEST of the first two (for me personally, that's the Eskimo Pie – I'm more of an ice cream person in general and frankly don't see the point in eating pie unless it's Thanksgiving).

So in our adjusted strategy, we select the first candidate we encounter among the following 4 desserts. Bear in mind that there is a (1/6 + 1/6 = 1/3 =) 33.33% chance that one of the first two desserts IS the best dessert of the bunch, but let's put that aside for now.

If the third dessert is the best of all (again, a 1/6 chance) then – just as above – we win automatically. It's better than desserts 1 and 2, so it's a candidate, so we would pick it and win.

If the fourth dessert is the best of all (1/6 chance), then we would only win if we pick it. We pick it if it's a candidate, and it's a candidate only when/if dessert 3 is NOT a candidate (i.e. is worse than desserts 1 and 2). The chances that the best dessert of the first three is either dessert 1 or dessert 2 (i.e., dessert 3 is NOT the best of the first three) are 2 out of 3. Total odds: $(1/6)(2/3) = 1/9 = 11.11\%$.

If the fifth dessert is best of all (1/6 chance), then we win if the best dessert from desserts 1 through 4 is either dessert 1 or dessert 2 (2 out of 4). Again, if either dessert 3 or 4 was better than the best of desserts 1 and 2, we'd pick that one as a candidate and miss the chance to pick dessert 5. Total odds of this case: $(1/6)(2/4) = 1/12 = 8.33\%$.

Similarly, to win with the sixth dessert, you need the sixth dessert to be best of all (1/6 chance) and you need the best dessert from 1 through 5 to be either 1 or 2 – a 2 in 5 chance. Total odds: $(1/6)(2/5) = 1/15 = 6.67\%$.

Adding these cases with our modified strategy:

Stop with Dessert #	Case in Order to Win	Probability
3	Dessert 3 best of all 6 (1/6)	1/6 = 0.1667
4	Dessert 4 best of all 6 (1/6) and the best of Desserts 1 through 3 is either Dessert 1 or Dessert 2 (2/3)	(1/6)(2/3) = 1/9 = 0.1111
5	Dessert 5 best of all 6 (1/6) and the best of Desserts 1 through 4 is either Dessert 1 or Dessert 2 (2/4)	(1/6)(2/4) = 1/12 = 0.0833
6	Dessert 6 best of all 6 (1/6) and the best of Desserts 1 through 5 is either Dessert 1 or Dessert 2 (2/5)	(1/6)(2/5) = 1/15 = 0.0667
	Total Probability of Win	0.4278

Impressive! Using this modified, "view, evaluate, and discard TWO desserts then pick the next candidate" methodology, we net a whopping 42.78% chance of winning! We're now miles away from our "random pick" starting point of 16.67% odds of winning and much, much closer to even odds (50/50) than I think any reasonable observer would have initially guessed possible. Note also that the chances of losing right off the bat by discarding the first two desserts are only 33.33% – so the choice to pick the next candidate is clearly superior.

2.2 Generalizing n

So with $n = 6$, we know that reviewing 1 dessert then picking the next candidate yields a 38.06% chance of winning, and reviewing 2 desserts then picking the next candidate yields a 42.78% chance of winning. Not shown here, but easily calculated, your odds of winning drop back down to 39.17% when you review 3 desserts then pick the next candidate, and drop further to 30% when you review 4 and pick the next candidate.

So we're in the right ballpark of strategies – we just need to know when's the best time to stop for a given number of desserts n. When n is 6, we've shown that the ideal r is 2; what about other values of n?

Consider the generalized case depicted here:

Here we have a generalized number of n desserts. Let's consider the strategy where we review, evaluate, rank, and discard $(s − 1)$ desserts, then look for the first candidate starting with dessert s.

Moving along the same lines of logic as with our $n = 6$ cases above, we note:

- Winning with dessert s only requires that s is the best dessert of all, and there's a $1/n$ chance of this being the case.

- Winning with dessert $(s + 1)$ requires that dessert $(s + 1)$ is best dessert of all $(1/n$ chance$)$, and that of the first s desserts, the best dessert must be one of desserts 1 through $(s − 1)$, chances of which are $(s − 1)$ in s.

- Winning with dessert $(s + 2)$ requires that dessert $(s + 2)$ is best of all $(1/n)$ and that the best dessert of the first $(s + 1)$ desserts is found from dessert 1 to dessert $(s − 1)$, chances of which are $(s − 1)/(s + 1)$.

- …

- Winning with dessert n requires that dessert n is best of all $(1/n)$ and that the best dessert of the first $(n − 1)$ desserts is found from dessert 1 to dessert $(s − 1)$, chances of which are $(s − 1)/(n − 1)$.

In mathematical notation, where P(s, n) represents the probability of winning with n desserts by picking candidates starting with dessert s:

$$P(s,n) = \left(\frac{1}{n}\right)\left(\frac{s-1}{s-1}\right) + \left(\frac{1}{n}\right)\left(\frac{s-1}{s}\right) + \left(\frac{1}{n}\right)\left(\frac{s-1}{s+1}\right) + \left(\frac{1}{n}\right)\left(\frac{s-1}{s+2}\right) + \ldots + \left(\frac{1}{n}\right)\left(\frac{s-1}{n-1}\right)$$

$$= \left(\frac{1}{n}\right)\left(\frac{s-1}{s-1} + \frac{s-1}{s} + \frac{s-1}{s+1} + \frac{s-1}{s+2} + \ldots + \frac{s-1}{n-1}\right)$$

$$= \left(\frac{1}{n}\right)\sum_{k=s}^{n}\frac{s-1}{k-1}$$

So here we have a generalized formula to determine the probability of winning given any number of desserts, and with any number of discards. For sanity's sake, we can run our example above ($n = 6$, $s = 3$) through this formula:

$$P(3,6) = \left(\frac{1}{6}\right)\left(\frac{3-1}{3-1}\right) + \left(\frac{1}{6}\right)\left(\frac{3-1}{3}\right) + \left(\frac{1}{6}\right)\left(\frac{3-1}{3+1}\right) + \left(\frac{1}{6}\right)\left(\frac{3-1}{3+2}\right)$$

$$= \left(\frac{1}{6}\right)\left(\frac{2}{2} + \frac{2}{3} + \frac{2}{4} + \frac{2}{5}\right)$$

$$= \left(\frac{1}{3}\right)\left(\frac{77}{60}\right)$$

$$= \frac{77}{180}$$

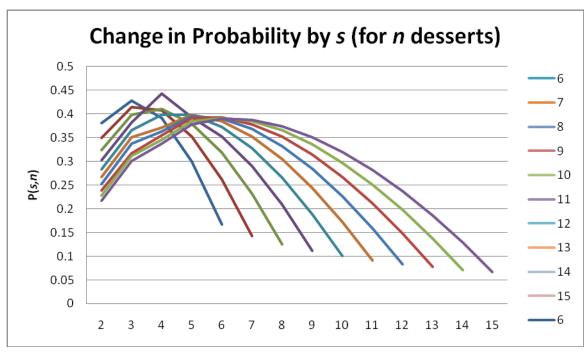$$\approx 0.427777778$$
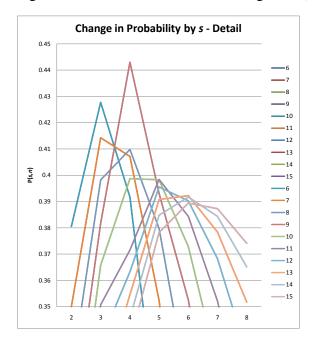
The same result as in our "counting" result above. Whew.

So we now have a clear – if, for large n computationally drawn-out – method for determining probability of looking for a candidate after a given number of discards. But the question remains – how many discards is ideal?

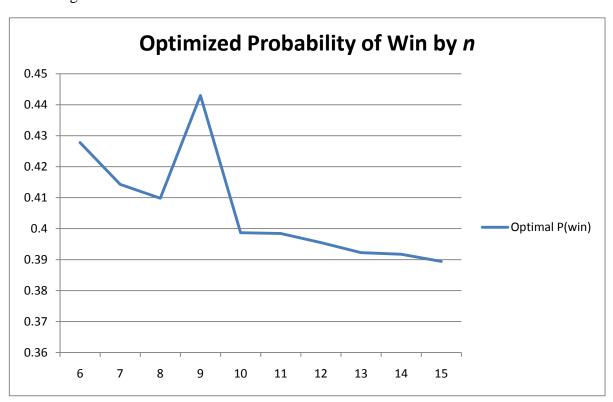An illustration of the P(s,n) function for various values of n shows that it's a moving target:

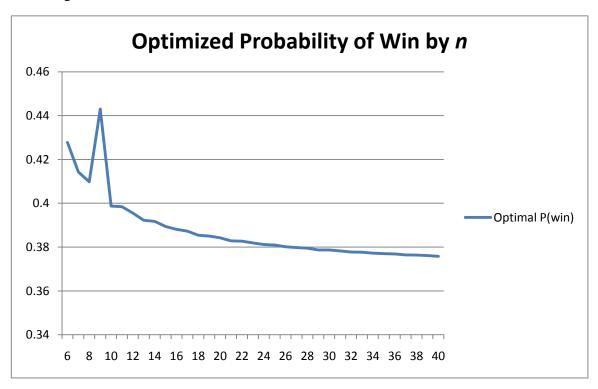In greater detail, zoomed in on the range of P(*s,n*) from 0.35 to 0.45:

The data behind the scenes and the optimal values for s highlighted:

s	6	7	8	9	10	11	12	13	14	15
					n					
2	0.380556	0.35	0.324107	0.301984	0.282897	0.26627	0.251656	0.238709	0.227152	0.216771
3	0.427778	0.414286	0.398214	0.381746	0.365794	0.350722	0.336646	0.317161	0.311448	0.300208
4	0.391667	0.407143	0.409821	0.442989	0.39869	0.371537	0.363303	0.354587	0.345743	0.336979
5	0.3	0.352381	0.379762	0.393122	0.398254	0.398413	0.395515	0.390731	0.3848	0.378194
6	0.166667	0.261905	0.318452	0.352513	0.372817	0.38438	0.390227	0.392261	0.391714	0.38941
7		0.142857	0.232143	0.289683	0.327381	0.352165	0.368272	0.378405	0.384343	0.387292
8			0.125	0.208333	0.265278	0.304798	0.332428	0.351729	0.365067	0.374062
9				0.111111	0.188889	0.244444	0.28468	0.314064	0.335587	0.351309
10					0.1	0.172727	0.226515	0.266783	0.297178	0.320223
11						0.090909	0.159091	0.210956	0.250833	0.281729
12							0.083333	0.147436	0.197344	0.236569
13								0.076923	0.137363	0.185348
14									0.071429	0.128571
15										0.066667

Visualizing these:

Optimized Probability of Win by *n*

Extending out further:

Optimized Probability of Win by *n*
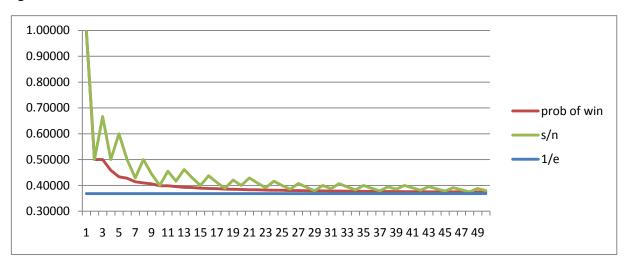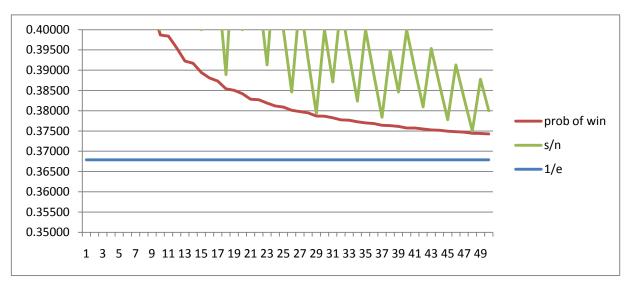
It appears clear that the probability of winning using this strategy trends toward a limit as *n* increases. In (Gilbert & Mosteller, 1966), this limit is shown to be 1/*e*.

Also of interest is an examination of the ratio *s*/*n*, where *s* is still the point beyond which we look for the next candidate. If we lay this ratio against the graph above, we notice something interesting:

Detail of the range [0.35,0.4]:

It can be shown that both the idealized probability of winning the Dessert Game and the ratio of ideal-*s* to *n* will approach 1/*e* as *n* approaches infinity.

3   The Full-Information Game

Another classic stopping puzzle pondered over the past century is similar to our renamed "Dessert Game" above, but with an interesting twist.

In the Dessert Game, we had absolutely no clue what might be under what serving dome – the next plate could hold a magically never-ending ice cream sundae or boiled shoe leather – in other words, we had no information about the values from which we were selecting. In the next game – the Full-Information Game – we assume some level of pre-awareness of what values are possible.

We play by the same basic rules as above – you draw from a known number *n* of values, you evaluate each value as you draw it, compare it to the other values you've seen before, and you make a choice: keep the value currently being considered (in your hand) or discard it forever and continue drawing values; if you stop with the highest value, you win; if you do not you lose. The same mechanics are in play, but the knowledge of what the values can be has a surprisingly positive effect on one's chances of winning the game.

3.1   *"A Known Distribution"*

A few words regarding this very important difference between the Dessert Game and the Full-Information Game, provided for the benefit of the mathematically rusty reader (or in this case, author).
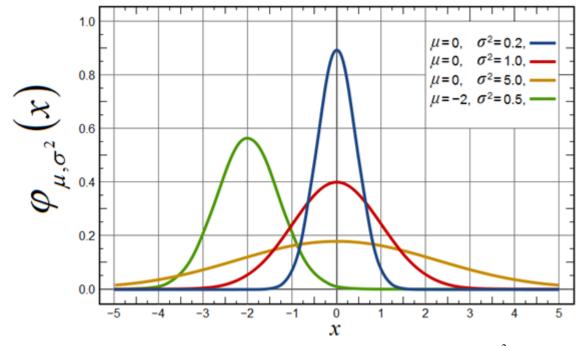
The assumption we make in the Full-Information game (Gilbert & Mosteller, 1966) is that we "know the distribution" of the values from which we are picking. In the context of mathematics and probability, a "distribution" is defined as follows:

> *Mathematical Distribution - A mathematical object that generalizes the idea of function. In other words, a mathematical distribution indicates the frequency of certain events occurring within a certain sample space.*

So in the context of our stopping games, this amounts to knowing:

- The range of the sample space, i.e. the boundaries of possible values
- How likely you are to pick any given value along the range

An example of a distribution is the "normal distribution", otherwise known as the "bell curve":

In the normal distribution, if you know the mean (μ) and the variance (σ^2) of the values being considered, you can reliably calculate the probability ($\varphi_{\mu,\sigma2}(x)$)of encountering any given value along the range.

This is an example of what is meant by choosing from "a known distribution" – in this game, we don't know what we're going to choose next, but we know a lot about what it may be (boundaries, probability of choosing one value over another, mean, variance, etc.).

Now, attempting to prove anything in this squishy, unlimited realm of continuous, known distributions at first would seem extremely problematic – too many options with too many variables. There exists, however, a little mathematical "trick" that actually makes it extremely simple – and it has to do with something called functional monotonic transformations.

Any distribution can, in essence, be "stretched" or "compressed" to a mapping on any other distribution. When this transformation occurs, if it preserves the order of the elements in the set (i.e., if $x < y$ in the first distribution, $x' < y'$ in the transformed mapping, it is considered "monotonic". Any truth we can prove on one distribution can be inferred – through the mapping – onto any other transformation of that distribution. So, we find the simplest distribution to work with, and can be confident that the findings can be extrapolated to any monotonic transformation of that distribution.

The simplest distribution to work with in this case is a "uniform distribution" – meaning a distribution in which every value in the range of the distribution is *equally likely* to occur. Taking it even farther, we choose to work with the "standard uniform distribution" – a uniform distribution defined on the range [0,1] (i.e. any value from 0 to 1 is equally likely to occur). This distribution – which again, through monotonic transformation may be used to represent any known distribution – has an added benefit: the probability of selecting a value greater than any given point on the range is simply one minus that value. For example, if the first value you select is 0.75, the probability of selecting a value greater than that is simply (1-0.75) = 0.25. This simplicity is key in unlocking the Full-Information Game; its simplicity belies the depth of application to the family of cases in which these findings can apply.

3.2   Simple Examples

Let's examine a few simple examples, using the rules as outlined above and focusing on the standard uniform distribution. First, say we are picking up to two values from the range 0 to 1, with any value's selection having an equal probability of being picked.

In this two-value game, there can be only two final outcomes: either the first value we select will be larger, or the second value we select will be larger. Regardless of the strategy that we employ to maximize our chances of winning, there are two possible outcomes, of which we will choose one. As such, whichever value we select, our probability of winning in this scenario is 0.5.

Let's now investigate ideal strategy for playing this sort of game. Since the chance of selecting any given value on this range from 0 to 1 is equal, we know that there's essentially a 50% chance of any given selection being greater than 0.5 and a 50% chance of it being less than 0.5. Once we examine the first value, this knowledge informs our next decision:

Say the first value we select is 0.3. Should we keep this value or discard it and select the next value? In the two-value game, the answer is clear: since there is a 70% chance that the next value will be greater than 0.3, we should discard and continue. Similarly:

If the first value is 0.8 – since there's only a 20% chance of doing better on the second value – we should keep it.

This "pivot" point of 0.5 – keep if the first value is greater, discard if less – becomes important, not just for the two-value game, but for determining optimal strategy for all larger games.

Next, let's say we are picking up to three values from the range 0 to 1, with any value's selection having an equal probability of being picked.

In this case, we see a break from the strategy of selecting randomly and using a strategy to increase our odds. Were we to pick randomly, we would have a 0.33 chance of choosing correctly. But in this case, we can achieve superior odds of winning based on the fact that we learn details regarding the game (and our odds) from seeing the first value:

Say the first value we discover is 0.7 – because we know that our defined range goes from 0 to 1, we know that this value is relatively high…but is it high enough to warrant ignoring both remaining values?

There are four possibilities:

- Both remaining values are larger than 0.7
- The second value is larger than 0.7 and the third is lower
- The second value is lower than 0.7 and the third is higher
- Both remaining values are lower than 0.7

If we were to keep the first value and ignore the next two, the probability of encountering each of these scenarios is simple to compute, since there is a 70% chance of getting lower than 0.7 and a 30% chance of getting higher:

Scenario	Probability of Occurring	Win if keep first value?
Both remaining are larger than 0.7	$(1-0.7)(1-0.7)$ $= 0.3^2$ $=0.09$	NO
Second higher than 0.7; third lower	$(1-0.7)(0.7)$ $= (0.3)(0.7)$ $= 0.21$	NO
Second lower than 0.7; third higher	$(0.7)(1-0.7)$ $= (0.7)(0.3)$ $= 0.21$	NO
Both remaining are lower than 0.7	$(0.7)(0.7)$ $= 0.7^2$ $= 0.49$	YES

Adding these cases up, we can see that there is a 51% chance of losing if we keep this first value, and a 49% chance of winning. As such, we should discard this value and try our luck with the remaining two values (in which, from above, we know we have a 50% chance of winning).

3.3   Generalizing the n=3 Game

In the previous example, we see we would benefit from discarding if the first value is 0.7 – but at what point does would it make sense to keep the first card? Where is the "pivot" point for the three-value game?

Mathematicians John P. Gilbert and Frederick Mosteller do all the heavy lifting on this class of solution, searching for a monotonic decreasing series of pivot points (or decision numbers) for any series of uniform distribution values (Gilbert & Mosteller, 1966). While their solutions are elegant and simply stated, their between-the-lines assumptions make for some dense reading for the struggling undergrad. The intent of this and the following sections is to unwind their dense arguments and examine their inner workings in expanded form.

Calculating this value is simple by generalizing the first value found and investigating relative probabilities of winning. Consider the case where there are three values to be discovered – A, B, and C – and the first value x is known:

A B C

Let's reconsider the three-value scenarios listed above, but with a generalized value x as the first discovered value. Here again we benefit from the selection of the standard uniform distribution, because we can state with certainty that the unknown values B and C each have a $(1 - x)$ chance of being greater than x and an x chance of being less than x.

At this point, there are four possible scenarios, each with a calculable probability of occurring:

Scenario	Probability of Occurring
B and C both less than x	$(x)(x)$ $= x^2$
B greater than x; C less than x	$(1 - x)(x)$ $= x - x^2$
B less than x; C greater than x	$(x)(1 - x)$ $= x - x^2$
B and C both greater than x	$(1 - x)(1 - x)$ $= (1 - x)^2$ $= 1 - 2x + x^2$

In case the probability of these scenarios is unclear or doubted, we can ease our minds by simply adding the four distinct probabilities together:

$$\left(x^2 \right) + \left(x - x^2 \right) + \left(x - x^2 \right) + \left(1 - 2x + x^2 \right)$$
$$= \left(1 \right) + \left(-2x + x + x \right) + \left(x^2 - x^2 - x^2 + x^2 \right)$$
$$= 1 + 0 + 0$$
$$= 1$$

The four probabilities sum to 1, confirming that all possible scenarios are accounted for.

Next, for each of these scenarios we ask: which are winners (if we discard x and try our luck at B and C)? To be clear, a scenario "wins" if we follow our game strategy and, in the end, wind up selecting the highest value of the three. Our game strategy is:

At this point, we assume that we have discarded x and are continuing on with B and C. Well – we've already played the game of choosing between two values, back at the start of §3.2, and we've determined a winning strategy: keep B if it is greater than 0.5; discard it and try your luck with C if B is less than 0.5. The one caveat in this case: what if x is greater than 0.5?

If that is the case, if say x was in fact 0.6, then there would be no point in keeping B if it were only 0.55; you would lose immediately. As such, the rules here apply only when $x > 0.5$; if the opposite were true, then your best strategy is to stick with the two-value game as examined previously.

So…on to our strategy and computation of a pivot point for the three-value game:

- Evaluate if the value is greater than its slot's pivot point
- Determine if the value is a candidate (i.e. is the greatest value uncovered so far)
  - If yes to both of the above, keep the value as the highest

The first scenario is the easiest: if x is in fact the greatest value of the three, then by discarding it we lose. Simple.

With a bit of thought, we can see that in the second and third scenarios, we will always win. Given our game strategy, we will pick the next candidate – if only one of the remaining two values is greater than x, then that is the value we'll pick, and we'll win. (To elaborate – if B is less than x we would discard it and move to C which is greater than x: WIN; if B is greater than x, then we would keep it, then see that C is less than x: WIN).

However the fourth case is slightly special. If both B and C are greater than x then the only way we can win is if B is also greater than C. At this point, knowing nothing about the values of B or C other than the fact that they are both greater than x, we can only say there is a 50% chance of the order being to our advantage (i.e. B > C). (Another way to state this, which will help when we evaluate the six-value and n-value cases below, is to say of the two possible states – B > C and B < C – we "choose" the one of the two that results in a win.)

Scenario	Probability of Occurring	Win if we discard x?
Both remaining are lower than x	$(x)(x)$ $= x^2$	NO
Second higher than x; third lower	$(1-x)(x)$ $= x - x^2$	YES
Second lower than x; third higher	$(x)(1-x)$ $= x - x^2$	YES
Both remaining are larger than x	$(1-x)(1-x)$ $= (1-x)^2$ $= 1 - 2x + x^2$	50% OF THE TIME

So, we add up our odds of winning by discarding the first and continuing:

- Scenario 1 has a chance of occurring of x^2, if it occurs we will win 0% of the time;

- Scenario 2 has a chance of occurring of $(x - x^2)$, if it occurs we will win 100% of the time;

- Scenario 3 has a chance of occurring of $(x - x^2)$, and if it occurs we will win 100% of the time;

- Scenario 4 has a chance of occurring of $(1 - x)^2$, and if it occurs we will win 50% of the time;

- Total odds:

$$(x - x^2) + (x - x^2) + \frac{1}{2}(1 - 2x + x^2)$$

$$= \frac{1}{2} + (x + x - x^2) + \left(-x^2 - x^2 + \frac{1}{2}x^2\right)$$

$$= \frac{1}{2} + x - \frac{3}{2}x^2$$

If this expression represents our chances of winning by passing on the first value and picking the next candidate, and if x^2 represents our chances of winning by keeping the first value, then our pivot point can be seen as the point where these two probabilities are equal to each other:

$$x^2 = \frac{1}{2} + x - \frac{3}{2}x^2$$

$$\frac{5}{2}x^2 - x - \frac{1}{2} = 0$$

$$5x^2 - 2x - 1 = 0$$

$$so$$

$$x = \frac{-(-2) \pm \sqrt{(-2)^2 - 4(5)(-1)}}{2(5)}$$

$$= \frac{2 \pm \sqrt{24}}{10}$$

$$= \frac{1 \pm \sqrt{6}}{5}$$
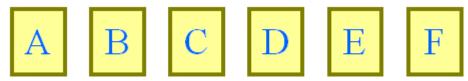
$$= (0.6899, -0.2899)$$

And so, finally: a pivot point decision number that takes into account our strategy of a) picking numbers that are greater than their decision numbers and b) are candidates.
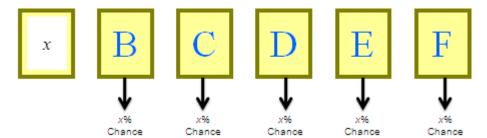
3.4   Further Generalization: n=6 and beyond

We will extend this concept of selecting pivot point decision numbers to a generalized number n, but first would be well-served to start with one final discrete example.

3.4.1   Example: $n = 6$

Before fully generalizing to determining the decision point for b_n, let's first consider the case where $n = 6$.
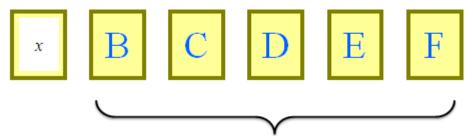
If we flip over card A and discover a value of x, there is only one way to win: all of the remaining 5 values must be less than x. Each individual value has a probability of x of being less than the value x:

By multiplying these individual probabilities together, we find the overall probability of winning by keeping the first value found: x^5.

However, if we discard this first value and select the next candidate, there are 5 distinct cases regarding the remaining 5 values that must be considered (we ignore the case where all remaining 5 values are less than x – that is the one case with no chance of winning):

1. All 5 remaining values are greater than x

2. 4 of the 5 are greater than x (and one is less than x)

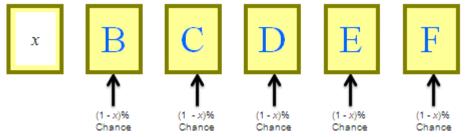3. 3 of the 5 are greater (and two are less)

4. 2 of the 5 are greater (and three are less)

5. 1 of the 5 are greater (and four are less)

In all cases, an individual value has a $(1 - x)$ probability of being greater than x and an x probability of being less than x. We examine each of these cases from the perspective of determining the chances of winning if we discard the first value and select the next candidate:

3.4.1.1 Case #1
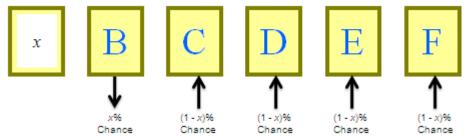

Case: All 5 remaining values are greater than x.

- This means that we'll be picking B (the next candidate we encounter)

- There is a 1 in 5 chance that B is the largest of the remaining values

- There are "5 choose 5" different places where the higher values can be, meaning 1 combination

- The overall chance of this case occurring is $(1 - x)^5$

$$\frac{1}{5}\binom{5}{5}(1-x)^5$$

3.4.1.2 Case #2
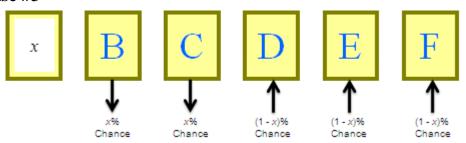
Case: 4 of the 5 are greater than x (and one is less than x).

- There's a 1 in 4 chance that the next candidate we choose is the largest of the 4

- There are "5 choose 4" different places where the higher values can be, meaning 5 combinations

- Each of these 5 sub-cases have odds of $(1 - x)^4(x)$ of occurring

$$\frac{1}{4}\binom{5}{4}(1 - x)^4(x)$$

3.4.1.3 Case #3

Case: 3 of the 5 are greater (and two are less).
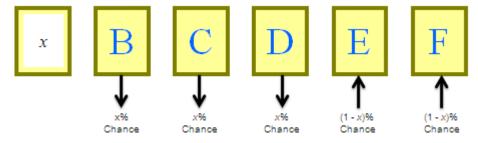
- There's a 1 in 3 chance that the next candidate we choose is the largest of the 3

- There are "5 choose 3" different places where the higher values can be, meaning 10 combinations

- Each of these 10 sub-cases have odds of $(1 - x)^3(x)^2$ of occurring

$$\frac{1}{3}\binom{5}{3}(1 - x)^3(x)^2$$

A simple illustration of the "5 choose 3" sub-cases of Case 3:

1 ↓ ↓ ↑ ↑ ↑

$x\%$ Chance $x\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance

↑ ↓ ↑ ↓ ↑ 6

$(1-x)\%$ Chance $x\%$ Chance $(1-x)\%$ Chance $x\%$ Chance $(1-x)\%$ Chance

2 ↓ ↑ ↓ ↑ ↑

$x\%$ Chance $(1-x)\%$ Chance $x\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance

↑ ↓ ↑ ↑ ↓ 7

$(1-x)\%$ Chance $x\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance $x\%$ Chance

3 ↓ ↑ ↑ ↓ ↑

$x\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance $x\%$ Chance $(1-x)\%$ Chance

↑ ↑ ↓ ↓ ↑ 8

$(1-x)\%$ Chance $(1-x)\%$ Chance $x\%$ Chance $x\%$ Chance $(1-x)\%$ Chance

4 ↓ ↑ ↑ ↑ ↓

$x\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance $x\%$ Chance

↑ ↑ ↓ ↑ ↓ 9

$(1-x)\%$ Chance $(1-x)\%$ Chance $x\%$ Chance $(1-x)\%$ Chance $x\%$ Chance

5 ↑ ↓ ↓ ↑ ↑

$(1-x)\%$ Chance $x\%$ Chance $x\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance

↑ ↑ ↑ ↓ ↓ 10

$(1-x)\%$ Chance $(1-x)\%$ Chance $(1-x)\%$ Chance $x\%$ Chance $x\%$ Chance

3.4.1.4 Case #4

| x | B | C | D | E | F |

B ↓ $x\%$ Chance   C ↓ $x\%$ Chance   D ↓ $x\%$ Chance   E ↑ $(1-x)\%$ Chance   F ↑ $(1-x)\%$ Chance

Case: 2 of the 5 are greater (and three are less).

- There's a 1 in 2 chance that the next candidate we choose is the largest of the 2

- There are "5 choose 2" different places where the higher values can be, meaning 10 combinations

- Each of these 10 sub-cases have odds of $(1 - x)^2(x)^3$ of occurring

$$\frac{1}{2}\binom{5}{2}\left(1-x\right)^2\left(x\right)^3$$

A simple illustration of the "5 choose 2" sub-cases of Case 4:

1 ↓ ↓ ↓ ↑ ↑ | ↓ ↑ ↑ ↓ ↓ 6
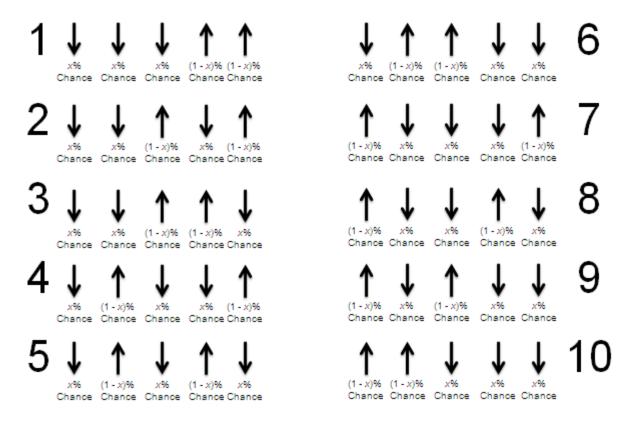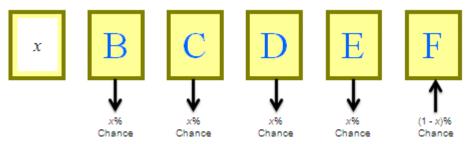x% x% x% (1 - x)% (1 - x)% | x% (1 - x)% (1 - x)% x% x%
Chance Chance Chance Chance Chance | Chance Chance Chance Chance Chance

2 ↓ ↓ ↑ ↓ ↑ | ↑ ↓ ↓ ↓ ↑ 7
x% x% (1 - x)% x% (1 - x)% | (1 - x)% x% x% x% (1 - x)%
Chance Chance Chance Chance Chance | Chance Chance Chance Chance Chance

3 ↓ ↓ ↑ ↑ ↓ | ↑ ↓ ↓ ↑ ↓ 8
x% x% (1 - x)% (1 - x)% x% | (1 - x)% x% x% (1 - x)% x%
Chance Chance Chance Chance Chance | Chance Chance Chance Chance Chance

4 ↓ ↑ ↓ ↓ ↑ | ↑ ↓ ↑ ↓ ↓ 9
x% (1 - x)% x% x% (1 - x)% | (1 - x)% x% (1 - x)% x% x%
Chance Chance Chance Chance Chance | Chance Chance Chance Chance Chance

5 ↓ ↑ ↓ ↑ ↓ | ↑ ↑ ↓ ↓ ↓ 10
x% (1 - x)% x% (1 - x)% x% | (1 - x)% (1 - x)% x% x% x%
Chance Chance Chance Chance Chance | Chance Chance Chance Chance Chance

3.4.1.5 Case #5

x	B	C	D	E	F
	↓	↓	↓	↓	↑
	x% Chance	x% Chance	x% Chance	x% Chance	(1 - x)% Chance

Case: 1 of the 5 are greater (and four are less).

- The next candidate we find will win (1 in 1 chance)

- There are "5 choose 1" different places where the higher values can be, meaning 5 combinations

- Each of these 5 sub-cases have odds of $(1 - x)(x)^4$ of occurring

$$\frac{1}{1}\binom{5}{1}\left(1 - x\right)\left(x\right)^4$$

3.4.1.6 Summing Odds for Cases 1 – 5

$$\underbrace{\frac{1}{1}\binom{5}{1}(1-x)(x)^4}_{\text{Case 5}} + \underbrace{\frac{1}{2}\binom{5}{2}(1-x)^2(x)^3}_{\text{Case 4}} + \underbrace{\frac{1}{3}\binom{5}{3}(1-x)^3(x)^2}_{\text{Case 3}} + \underbrace{\frac{1}{4}\binom{5}{4}(1-x)^4(x)}_{\text{Case 2}} + \underbrace{\frac{1}{5}\binom{5}{5}(1-x)^5}_{\text{Case 1}}$$

$$\sum_{j=1}^{5}\frac{1}{j}\binom{5}{j}(1-x)^j(x)^{5-j}$$

This final expression, then, represents the probability of winning if you were to discard the first value and select the next candidate.
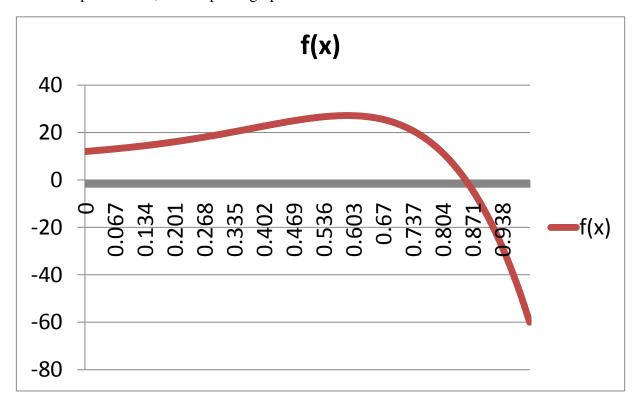
3.4.1.7 Finding the Pivot Point for *n* = 6

So the odds of winning by keeping the first selection is x^5 and the odds of winning by rejecting the first selection and selecting the next candidate are represented by the expression above. To find the pivot point decision number, we set these expressions equal to each other and calculate a solution:

$$x^5 = \sum_{j=1}^{5}\frac{1}{j}\binom{5}{j}(1-x)^j(x)^{5-j}$$

$$x^5 = \frac{1}{1}\binom{5}{1}(1-x)(x)^4 + \frac{1}{2}\binom{5}{2}(1-x)^2(x)^3 + \frac{1}{3}\binom{5}{3}(1-x)^3(x)^2 + \frac{1}{4}\binom{5}{4}(1-x)^4(x) + \frac{1}{5}\binom{5}{5}(1-x)^5$$

$$x^5 = \left(x^4-5x^5\right)+\left(x^3-10x^4+5x^5\right)+\left(\frac{10x^2-30x^3+30x^4-10x^5}{3}\right)+\left(\frac{5x-20x^2+30x^3-20x^4+5x^5}{4}\right)+\left(\frac{1-5x+10x^2-10x^3+5x^4-x^5}{5}\right)$$

$$60x^5 = \left(60x^4-300x^5\right)+\left(60x^2-600x^3+600x^4-200x^5\right)+\left(75x-300x^2+450x^3-300x^4+75x^5\right)+\left(12-60x+120x^2-120x^3+60x^4-12x^5\right)$$

$$60x^5 = 12+15x+20x^2+30x^3+60x^4-137x^5$$

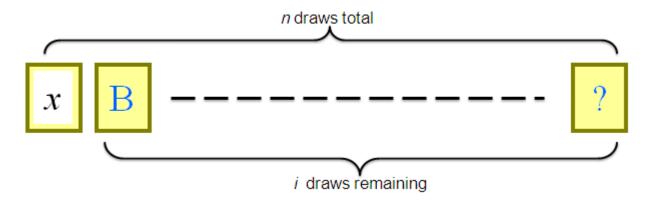$$0 = 12+15x+20x^2+30x^3+60x^4-197x^5$$

Not so simple to solve, but simple to graph for values between 0 and 1:

We find the solution for $n = 6$ to be approximately 0.8559492.

3.4.2 Generalizing to n

Now we have a clear path to generalize our solution to any number n (although the capability to compute values beyond modestly-sized n is best left to computer scripts).

$$\binom{i}{1}(1-x)^{i-1}(-x)+\frac{1}{2}\binom{i}{2}(1-x)^{i-2}(-x)^2+\frac{1}{3}\binom{i}{3}(1-x)^{i-3}(-x)^3+\cdots+\frac{1}{i}\binom{i}{i}(-x)^i$$

or

$$\sum_{j=1}^{i}\frac{1}{j}\binom{i}{j}x^{i-j}(1-x)^{j}$$

If we use this generalized formula for $n = 3$ from above, we get:

$$\sum_{j=1}^{2}\frac{1}{j}\binom{2}{j}x^{2-j}(1-x)^{j}$$

$$=\frac{1}{1}\binom{2}{1}x^{2-1}(1-x)^{1}+\frac{1}{2}\binom{2}{2}x^{2-2}(1-x)^{2}$$

$$=\left(2x-2x^2\right)+\left(\frac{1}{2}-x+\frac{1}{2}x^2\right)$$

$$=\frac{1}{2}+x-\frac{3}{2}x^2$$

Which, if we set equal to x^2 we get:

$$x^2=\frac{1}{2}+x-\frac{3}{2}x^2$$

$$0=\frac{1}{2}+x-\frac{5}{2}x^2$$

$$\frac{5}{2}x^2-x-\frac{1}{2}=0$$

$$5x^2-2x-1=0$$

This is the same expression we discovered above, using our raw strategy-counting strategy.

3.5 *Overall Odds of Winning*

So now we have a solid strategy and a way to calculate the crucial element of that strategy: our pivot point decision number. But if we were to employ this strategy over and over in multiple iterations of this game, what would our overall chance of success be?

3.5.1 Start small

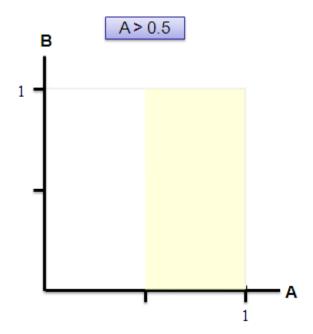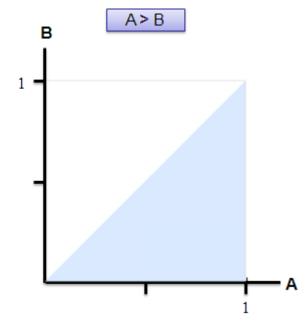Again, let's start with the simplest case of $n = 2$:

A B

The simplest way to prove the overall probability of winning is with a visual probability chart. In words, there are only two ways to win in the simple, $n = 2$ scenario:
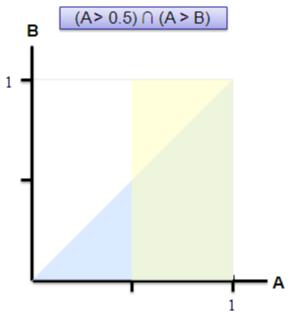
- If A > 0.5, it is greater than its pivot point and we would keep it; to win, A must then be greater than B

- If A < 0.5, it is less than its pivot point and we would discard it; to win, A must end up being less than B

If we consider a graph of A vs. B, where both A and B can range from 0 to 1, we may visualize the two separate concepts of "A > 0.5" and "A > B":
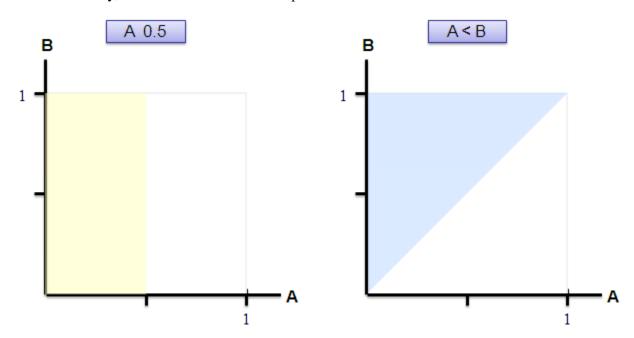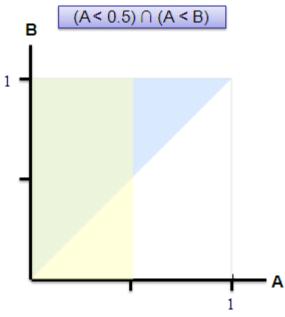
From here, it is easy to imagine the intersection of these two areas, here shown in green:

$(A> 0.5) \cap (A > B)$

Similarly, we can envision the concepts of "A < 0.5" and "A < B":

A 0.5

A < B

And again, the intersection of these two:
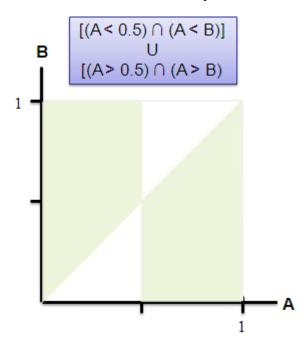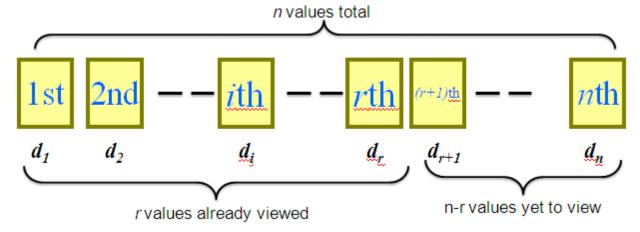

And so, if we consider the union of these two intersections, we have our final probability chart for a win in the $n = 2$ scenario, which we can easily see to be 0.75:

3.5.2 A Touch of Generalization

Before we expand into the *n*=3 realm, it's best to touch base again with Gilbert & Mosteller (Gilbert & Mosteller, 1966) regarding this topic.

In their paper, Gilbert & Mosteller present the following generalized case:

Gilbert & Mosteller present a theorem that allows determination of the probability of winning at any given $(r + 1)$th draw as the following:

$$P(1) = \frac{\left(1 - d_1^n\right)}{n}$$

$$P(r+1) = \frac{\sum_{i=1}^{r}\left(\frac{d_i^r}{r} - \frac{d_i^n}{n}\right)}{\left(1 - r\right)} - \frac{d_{r+1}^n}{n}$$
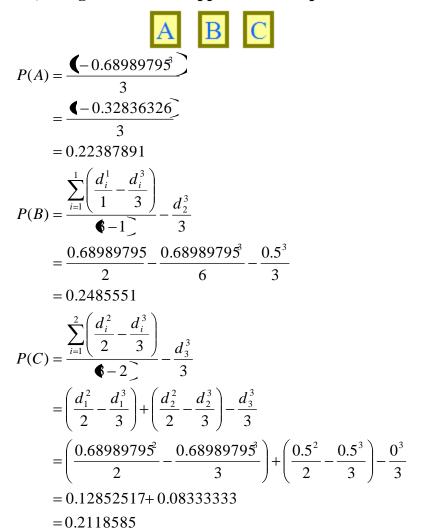
The logic provided is as follows:

- To determine the probability of winning with the first value, one takes the chance of the first value being the largest of all *n* values $(1/n)$ and subtracts the probability of the case where all *n* values are less than the first value's pivot point/decision number (the "no choice" scenario) – namely, the decision number itself raised to the number of values (*n*) divided by the number of values (*n*).

- To determine the probability of any winning with any given $(r + 1)$ value, each of the following cases must be considered:

    1. For each element *i* from the first value to the *r*th value, determine the difference between the chance that no value in the 1 to *r* range is greater than that value's decision number and the change of the same in the 1 to *n* range

    2. Sum these differences together

- 37 -

3.   Divide this sum by the number of values remaining from $(r + 1)$ to n – this accounts for the chance that the $(r + 1)$th value is the largest of all remaining values

4.   Subtract the chances of the case where the $(r + 1)$th value is largest, but not larger than its decision number

We shall follow this logic to compute the overall probability of winning the $n = 3$ scenario, knowing that $d_1 = 0.6898795$, $d_2 = 0.5$, and $d_3 = 0$.

3.5.3 Back to *n=3*, Using Generalized Approach and Optimum Decision Points

$$\boxed{A}\quad \boxed{B}\quad \boxed{C}$$

$$P(A) = \frac{\left(-0.68989795^3\right)}{3}$$

$$= \frac{\left(-0.32836326\right)}{3}$$

$$= 0.22387891$$

$$P(B) = \frac{\sum_{i=1}^{1}\left(\dfrac{d_i^1}{1} - \dfrac{d_i^3}{3}\right)}{\left(3-1\right)} - \frac{d_2^3}{3}$$

$$= \frac{0.68989795}{2} - \frac{0.68989795^3}{6} - \frac{0.5^3}{3}$$

$$= 0.2485551$$

$$P(C) = \frac{\sum_{i=1}^{2}\left(\dfrac{d_i^2}{2} - \dfrac{d_i^3}{3}\right)}{\left(3-2\right)} - \frac{d_3^3}{3}$$

$$= \left(\frac{d_1^2}{2} - \frac{d_1^3}{3}\right) + \left(\frac{d_2^2}{2} - \frac{d_2^3}{3}\right) - \frac{d_3^3}{3}$$

$$= \left(\frac{0.68989795^2}{2} - \frac{0.68989795^3}{3}\right) + \left(\frac{0.5^2}{2} - \frac{0.5^3}{3}\right) - \frac{0^3}{3}$$

$$= 0.12852517 + 0.08333333$$

$$= 0.2118585$$

P(total)=0.68429251

This result matches that of Gilbert & Mosteller.

4 Conclusion

A large part of the beauty involved with number theory – of which stopping problems are a part – has to do with taking an easily-communicated and easily-understood problem, that would seem at first (and second and third) glance to have an obvious solution, and then showing how an initially non-intuitive result is, in fact, incontrovertibly true. Better yet, it is often the case that the truth of the matter – while at first non-intuitive – becomes plainly obvious. Like the image of an Old Woman in an optical illusion, the answer is just sitting there, staring at us, waiting for our eyes to focus in on it. Once brought into focus, the correct answer is the only one you can see (and you feel a bit silly for ever having not seen it).

I'll be honest, however – there are also cases in this practice in which the seemingly simple truth, like a fish just about pulled into the boat, all of a sudden turns on you, launching you out of the boat, into the murky depths, and down the proverbial rabbit hole. There have been several times during this exercise that I have thought to have an answer firmly in-hand, only (upon further reflection) to have that answer disintegrate and slide away. And like a dream upon waking, the harder you focus on it, the faster it vanishes.

These stopping problems represent, in my view, examples of both of the cases above – and in doing so, handily represent the allure, satisfying rewards, and frustrating challenges of this genre. Stopping problems are comprehensible by the layman, provide sufficient fodder for the recreational mathematician, and yet still have the ability to vex those who have dedicated their professional lives to their study. They are at once both inherently practical and theoretically ephemeral; they represent the ideal aesthetic of mathematical elegance, and yet there doesn't seem a way to really solve anything without digging in and getting your hands dirty.

5 Bibliography

Ferguson, T. S. (1989). Who Solved the Secretary Problem? *Statistical Science , 4* (3), 282-296.

Gilbert, J. P., & Mosteller, F. (1966). Recognizing the Maximum of a Sequence. *Journal of the American Statistical Association , 61* (313), 35- 73.

Hill, T. P. (2009). Knowing When To Stop: How to gamble if you must - the mathematics of optimal stopping. *American Scientist , 97*, 126-133.