



© 1997–2009, Millennium Mathematics Project, University of Cambridge.

Permission is granted to print and copy this page on paper for non-commercial use. For other uses, including electronic redistribution, please contact us.

September 2008

Features



Kissing the frog: A mathematician's guide to mating

by John Billingham



Finding the Decision Numbers

As discussed in the main article, if the frogs are randomly numbered between 0 and 1, with the prince having the highest number, there is a sequence of decision numbers associated with the best strategy for choosing the top frog. Since we're meant to be doing some maths, we need some notation. We'll consider the general case with N frogs, and let the number on the j th frog be y_j , $j = 1, 2, \dots, N$. In addition, we'll try to find an increasing sequence of decision numbers, b_j , $j = 0, 1, 2, \dots, N - 1$, such that we kiss the j th frog if it has the largest number seen so far and $y_j > b_{N-j}$. In other words, to think about kissing the last frog we need $y_N > b_0$, the next to last frog, $y_{N-1} > b_1$, and so on. The reasons for this back to front numbering of the decision numbers should become clear below.

Let's assume that the $(N - k)$ th frog has the largest number so far on its back, y_{N-k} , which we'll write as just y to make things look nicer on the page. There are therefore k frogs still to come out of the pond. If we kiss the frog, the probability that he's the handsome prince is equal to the probability that the remaining k frogs have numbers less than y , which is y^k . That's the easy part of the calculation. The next bit is rather more tricky, but basically just involves counting (or combinatorics as it's rather grandly called by mathematicians).

If we don't kiss the frog, but choose the next one with number larger than y , we can work out the probability that we kiss the prince by considering the situation when there are exactly j frogs remaining that have numbers higher than the current frog's, y , and adding up our answers for all possible values of $j = 1, 2, \dots, k$. The probability that there are exactly j frogs remaining that are numbered higher than y is

Kissing the frog: A mathematician's guide to mating

equal to $y^j(1-y)^{k-j}$ times the number of different ways that we can choose j frogs from the final k , which is

$$\binom{k}{j} = \frac{k!}{j!(k-j)!}.$$

The probability that the prince comes *first* in the sequence of j higher-numbered frogs, which is the only position in which he can get kissed, is therefore

$$\frac{1}{j} \binom{k}{j} y^{k-j} (1-y)^j.$$

The appropriate decision number for the current frog is the value of y at which the probability of winning by kissing and the probability of winning by not kissing are equal. This gives us, remembering that we must sum over all possible values of j ,

$$y^k = \sum_{j=1}^k \frac{1}{j} \binom{k}{j} y^{k-j} (1-y)^j,$$

or, by defining $z = (1-y)/y$,

$$y^k = \sum_{j=1}^k \frac{1}{j} \binom{k}{j} y^k z^j.$$

Bringing everything over to the right hand side and factoring out the y^k gives

$$0 = y^k \left(-1 + \sum_{j=1}^k \frac{1}{j} \binom{k}{j} z^j \right).$$

Since y^k cannot possibly be zero, this leads to the equation

$$-1 + \sum_{j=1}^k \frac{1}{j} \binom{k}{j} z^j = 0. \tag{1}$$

Once we've found the unique positive solution of this, $z = z_+$ (can you prove there's one and only one positive solution?), the k th decision number is given by

$$b_k = \frac{1}{1 + z_+}. \tag{2}$$

Let's try this for various values of k . Note that if you're up to the last frog, you'd obviously kiss it if it's the highest so far, so $b_0 = 0$. If $k = 1$ there's just one frog remaining after the current frog. The average value of this final frog is 0.5, so we should kiss the last but one frog if it's the highest we've seen so far and its number satisfies $y_{N-1} > 0.5$. We therefore expect that the corresponding decision number is $b_1 = 0.5$. Now, with $k = 1$, equation (1) becomes $-1 + z = 0$, with the pretty obvious positive solution

Kissing the frog: A mathematician's guide to mating

$z = z_+ = 1$, and hence from equation (2), $b_1 = 0.5$, as expected.

When $k = 2$, there are two frogs remaining, and the value of the decision number is less obvious. Equation (1) becomes, after rearranging, the quadratic equation $-1 + 2z + \frac{1}{2}z^2 = 0$, which has a positive solution $z = z_+ = \sqrt{6} - 2$, and hence $b_2 = (1 + \sqrt{6})/5 \approx 0.6899$. You should be able to see these last three decision numbers, 0, 0.5, 0.6899, on the graphs plotted in the main article. We need to solve a cubic equation to find b_3 , a quartic equation to find b_4 , and so on. Pencil and paper isn't much help here but, because equation (1) has a unique positive solution $z = z_+$, it's straightforward to find z_+ iteratively using Newton's method. You can see the first 100 decision numbers in the graphs in the main article.

One other thing we'd like to know is how the decision numbers behave as k gets large.

The maths gets a little harder here, so if you've had enough, you can [return to the main article](#).

What happens as k gets large?

We can see from the graphs in the main article that $b_k \rightarrow 1$ as $k \rightarrow \infty$, but how fast does this happen? In order to do this, we need to find the positive solution of equation (1) when k is large. You could have a think about how you might do this before reading on, but I warn you that it's not easy.

The summed part of equation (1) looks like a binomial series, except for the factor of $1/j$ multiplying the terms. However, if we differentiate the series, this goes away. This suggests that we can write the equation as an integral of a binomial series, and a little experimentation shows that equation (1) is equivalent to

$$1 = \int_0^z \frac{(1+s)^k - 1}{s} ds. \quad (3)$$

If you don't believe me, use the binomial theorem on $(1+s)^k$ and then integrate the series you get term by term.

How does equation (3) help us? Well, integrals are always easier to estimate than sums, and we need to estimate the integral in equation (3) when k is large. In general, $(1+s)^k$ will be large when k is large, but the left hand side of the equation is just 1. Let's look for a solution with z small, and therefore s small. If we define new variables $s = \bar{s}/k$ and $z = \bar{z}/k$, equation (3) becomes

$$1 = \int_0^{\bar{z}} \frac{(1 + \frac{\bar{s}}{k})^k - 1}{\bar{s}} d\bar{s}. \quad (4)$$

Now

$$\left(1 + \frac{\bar{s}}{k}\right)^k = \exp\left\{k \ln\left(1 + \frac{\bar{s}}{k}\right)\right\} \approx \exp(\bar{s}),$$

when k is large. We have used the first term in the Taylor series expansion of $\ln\left(1 + \frac{\bar{s}}{k}\right)$ to approximate it here. We can therefore write (4) as

$$1 \approx \int_0^{\bar{z}} \frac{e^{\bar{s}} - 1}{\bar{s}} d\bar{s}. \quad (5)$$

What happens as k gets large?

Kissing the frog: A mathematician's guide to mating

This is now just one equation that we must solve, since we've managed to remove k . We can solve this on a computer, approximating the integral using, for example, the trapezium rule, and find the solution $\bar{z} = \bar{z}_+ \approx 0.80435$. We now just have to work our way back through our changes of variable; first, $z_+ \approx \bar{z}_+/k$, and then $b_k \approx 1/(1 + \bar{z}_+/k) = (1 + \bar{z}_+/k)^{-1}$, and hence, applying the binomial theorem to this final result,

$$b_k \approx 1 - \frac{\bar{z}_+}{k} \approx 1 - \frac{0.80435}{k} \quad \text{for } k \text{ large.}$$

We have therefore shown that the decision numbers do, as expected, approach 1 from below as k gets large, decaying algebraically, like $1/k$ as $k \rightarrow \infty$.

[Return to main article](#)



Plus is part of the family of activities in the Millennium Mathematics Project, which also includes the [NRICH](#) and [MOTIVATE](#) sites.