HOÀNG TỬ ẾCH VÀ XÁC SUẤT

LƯU MINH ĐỨC

Phỏng dịch theo bản gốc

Kissing frog: a mathematician's guide to mating

Chắc mọi người ai cũng đều biết câu chuyện Hoàng Tử Ếch. Đây là 1 phiên bản khác của câu chuyện ấy, hơi phức tạp hơn.

HOÀNG **TỬ** ẾCH

Một nàng công chúa đi lạc vào 1 khu rừng nọ. Đến 1 cái ao thì 1 mụ phù thủy hiện ra và nói "Đứng lại nào, con bé kia! Ta đã biến 1 hòang tử đẹp trai thành con ếch và giam cầm nó trong cái ao này cùng với 99 con ếch khác. Mỗi con ếch đều mang 1 số trên lưng và con số của hoàng tử là lớn nhất. Đấy là cách duy nhất để ngươi nhận ra nó từ trong lũ ếch. Nếu ngươi muốn rời khỏi khu rừng đã bị ếm bùa này, ngươi phải tìm ra hoàng tử và hôn nó. Mỗi con ếch sẽ lần lượt nhảy lên khỏi hồ. Khi mỗi conj ếch xuất hiện, ngươi phải quyết định hôn nó hay đá nó trở lại vào ao, và mỗi con ếch chỉ nhảy lên đúng 1 lần mà thôi. Nếu ngươi hôn phải ếch thật hoặc không chịu hôn con nào thì ngươi sẽ không thể rời khu rừng và hoàng tử thì vĩnh viễn ở lại trong hồ. Và với 1 tràng cười quỷ quyệt, mụ ta chìm trở lại xuống hồ. Rất may, công chúa của chúng ta rất giỏi toán và đã tìm được chiến lược tốt nhất để quyết định nên hôn chú ếch nào.

Chú ếch 1 nhảy ra với số 2 trên lưng và thế là bị đá trở lại vô hồ, chú ếch 2 mang số 12 (có khá hơn) nhưng cũng chung số phận với chú thứ nhất. Chú thứ 3 mang số -6 trên lưng (đâu có quy định các chú ếch chỉ mang toàn số dương!) và dĩ nhiên là vô hồ ngay. Công chúa lặp lại điều này với cả 37 chú ếch đầu tiên và nhận thấy rằng số lớn nhất đã xuất hiện là 23.2 (cũng không quy định là các chú ếch chỉ mang toàn số nguyên). Thế là cô đợi đến khi có chú ếch mang số lớn hơn 23.2 và hôn nó (dĩ nhiên chú ếch mang số 23.2 có thể chính là hoàng tử và cô có thể đã bỏ lỡ cơ hội duy nhất đó!). Chú ếch biến mất sau 1 tiếng nổ "bụp" và 1 hoàng tử đẹp trai xuất hiện. Thế là họ sống hạnh phúc mãi mãi (còn nếu bạn thích kết cục bất hạnh thì: mụ phù thủy hiện lên cười đắc thắng khi chú ếch vẫn y nguyên trước sự chú ý hết mực của công chúa. Ôi trời, vẫn còn chú ếch mang số lớn hơn chưa xuất hiện!).

QUY TẮC 37%

Chiến lược của công chúa là trước tiên thu thập 1 số thông tin về các chú ếch (cô ấy ghi nhớ số lớn nhất trong số 37 chú đầu tiên) và rồi lựa chọn dựa trên thông tin này (cô hôn chú ếch mang số cao hơn tiếp theo). Nếu số ếch là khác, cô ấy nên thu thập thông tin từ bao nhiêu con trước khi quyết định hôn? Câu trả lời là 37% của số ếch tổng cộng. Tại sao lại 37%? Nếu bạn biết các quy tắc đếm và biết rằng nguyên hàm của 1/x là $\ln x$, bạn có thể theo các suy luận logic dưới đây để hiểu được kết luận này.

CM quy tắc 37%

Ta hãy xét bài toán tổng quát

Có N chú ếch. Cần phải đá hết bao nhiêu chú xuống hồ để xác suất hôn trúng hoàng tử là lớn nhất?

Giả sử cần đá M chú ếch đầu tiên xuống hồ và rồi hôn chú ếch đầu tiên mà có số cao hơn M chú ban đầu này. Để xác định M ta cần tìm ra khả năng hôn trúng hoàng tử là bao nhiêu, tùy thuộc vào chuyện anh ta có xuất hiện trong dãy từ M+1 đến N hay không.

Nếu anh ta nằm trong M chú đầu tiên, thảm họa, công chúa đã bỏ qua anh ta.

Nếu anh ta là chú ếch thứ M+1, thành công, công chúa chắc chắn sẽ hôn anh ta.

Nếu anh ta là chú ếch thứ M+2, công chúa sẽ hôn anh ta trừ trường hợp xui xẻo là chú ếch thứ M+1 lại mang số cao hơn M chú đầu tiên. Tuy nhiên khả năng trường hợp xui xẻo này xảy ra chỉ là 1/(M+1) (vì số các hoán vị của M+1 chú ếch trong đó chú ở vị trí M+1 mang số cao nhất cũng bằng số các hoán vị của M+1chú mà chú mang số cao nhất ở vị trí bất kỳ. Nếu vẫn chưa rõ, hãy tính cụ thể xác suất này ra, bạn sẽ hiểu ngay). Do đó xác suất mà công chúa hôn trúng hoàng tử sẽ là

$$1-\frac{1}{M+1}=\frac{M}{M+1}$$

Tương tự nếu anh ta là chú ếch thứ M+3 thì công chúa sẽ hôn trúng anh ta trừ phi chú ếch thứ M+1 hoặc M+2 mang số cao nhất trong M+2 chú đầu tiên. Do đó xác suất công chúa hôn trúng hoàng tử sẽ là

$$1-\frac{2}{M+2}=\frac{M}{M+2}$$

Và cứ thế cho đến khi

……………………………………………………………………………………

Nếu anh ta là chú ếch thứ N thì công chúa sẽ hôn trúng anh ta trừ phi trong số các chú ếch từ M+1đến N-1có chú ếch mang số cao nhất trong N-1 chú ếch đầu tiên. Do đó xác suất hôn trúng hoàng tử trong trường hợp này sẽ là

$$1-\frac{N-M-1}{N-1}=\frac{M}{N-1}$$

Hơn nữa xác suất để hoàng tử ở 1 trong các vị trí từ M+1 đến N đều là 1/N(check this) . Do đó cuối cùng ta tìm được xác suất để hôn trúng hoàng tử sau khi đã đá xuống hồ M chú ếch đầu tiên là:

$$\begin{aligned}P(M,N)&=\frac{1}{N}\left(1+\frac{M}{M+1}+\frac{M}{M+2}+...+\frac{M}{N-1}\right)\\&=\frac{M}{N}\left(\frac{1}{M}+\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}\right)\end{aligned}$$

Ta muốn chọn M sao cho xác suất này là lớn nhất, tức là

$$P(M,N)\geq P(k,N),\forall 1\leq k\leq N$$

Trước hết để đơn giản ta chọn M sao cho

$$\begin{cases}P(M,N)>P(M-1,N)\\P(M,N)>P(M+1,N)\end{cases}$$

Xét bất phương trình đầu tiên

$$\frac{M}{N}\left(\frac{1}{M}+\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}\right)>\frac{M-1}{N}\left(\frac{1}{M-1}+\frac{1}{M}+\frac{1}{M+1}+...+\frac{1}{N-1}\right)$$

$$\Leftrightarrow M\left(\frac{1}{M}+\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}\right)>(M-1)\left(\frac{1}{M-1}+\frac{1}{M}+\frac{1}{M+1}+...+\frac{1}{N-1}\right)$$

$$\Leftrightarrow M\left(\frac{1}{M}+\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}\right)>1+(M-1)\left(\frac{1}{M}+\frac{1}{M+1}+...+\frac{1}{N-1}\right)$$

$$\Leftrightarrow\left(\frac{1}{M}+\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}\right)>1$$

Tương tự BPT thứ 2 cho ta

$$\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}<1$$

Vậy ta cần M thỏa:

$$\left(\frac{1}{M}+\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}\right)>1>\frac{1}{M+1}+\frac{1}{M+2}+...+\frac{1}{N-1}$$

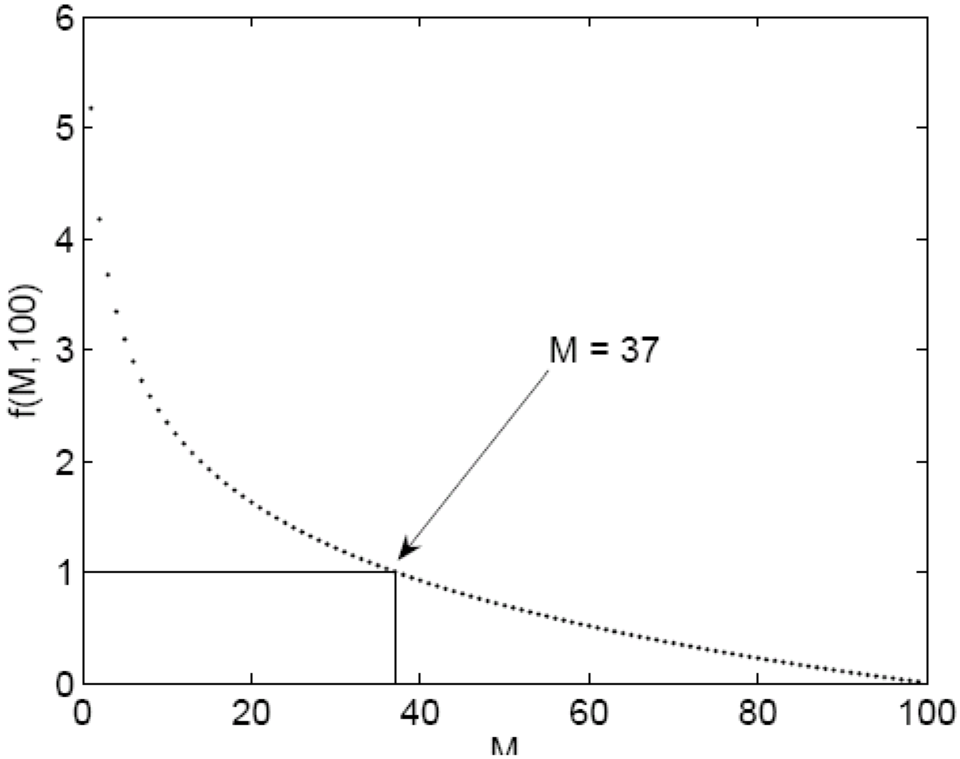
Ta định nghĩa hàm

$$f(M,N)=\frac{1}{M}+\frac{1}{M+1}+...+\frac{1}{N-1}$$

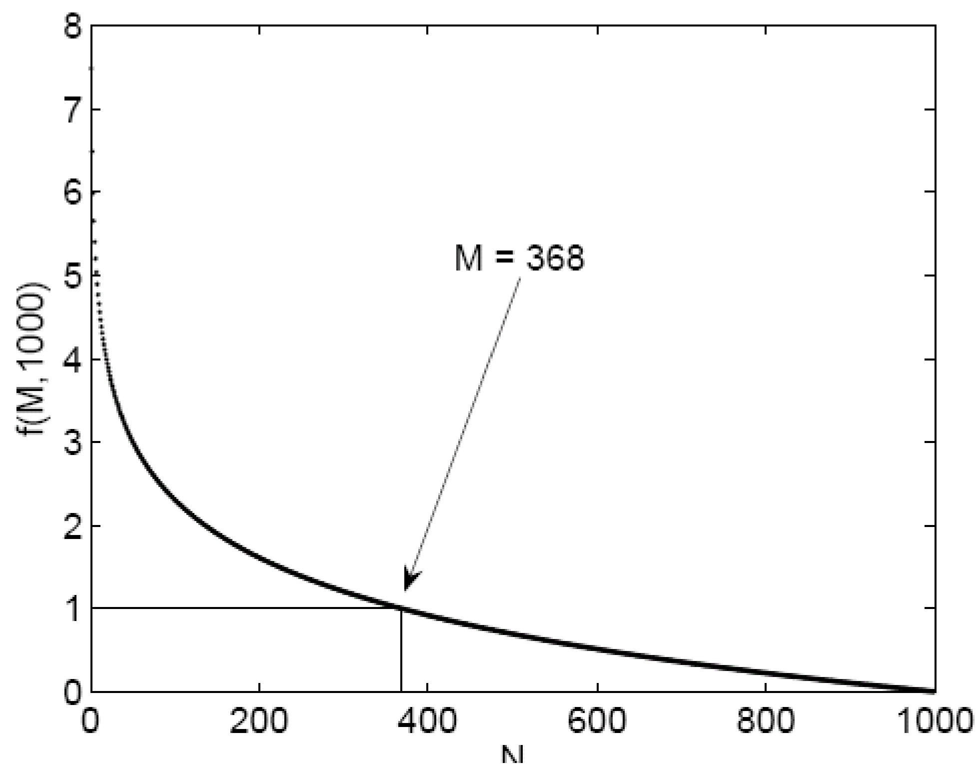
Khi đó M thỏa

$$f(M,N)>1>f(M+1,N)$$

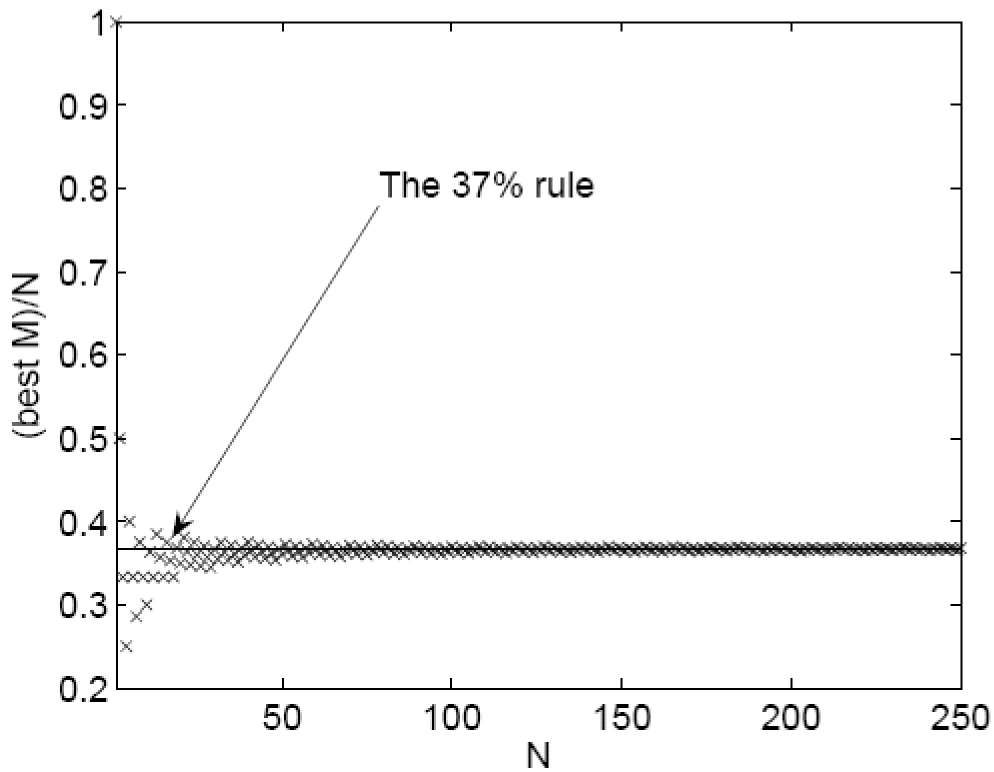
Với N=100 ta có thể nhờ máy tính vẽ đồ thị của hàm f(M,100) và tìm ra M=37.



Với N=1000, từ đồ thị ta tìm ra M=368

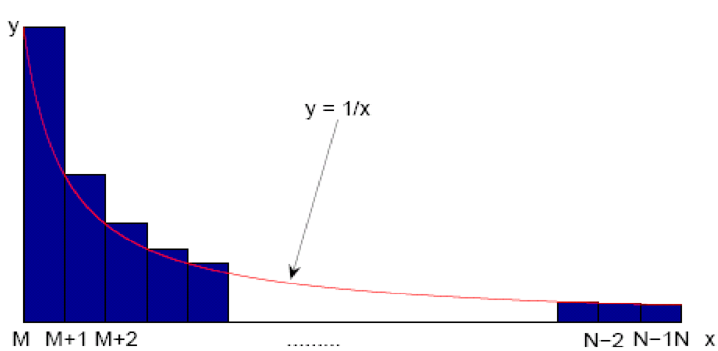


Bây giờ ta sẽ vẽ đồ thị của tỉ số M/N để các bạn có thể chọn M theo các giá trị N khác nhau



Khi N lớn dần ta có thể thấy rằng tỉ lệ ếch mà công chúa nên đá xuống hồ trước khi quyết định hôn là xấp xỉ 37%. Nhưng tại sao? Ta có thể chính xác hơn được không? Ta có thể ước lượng độ lớn của hàm f(M,N) khi N rất lớn hay không? Phần tiếp theo sẽ nêu rõ điều này.

Do f(M,N) là tổng các nghịch đảo của các số từ M đến N-1 nên ta ước lượng hàm f dựa theo hàm y=1/x. Hãy xem đồ thị dưới đây:



Tổng diện tích của các khối hình chữ nhật màu xanh chính là f(M,N) i.e

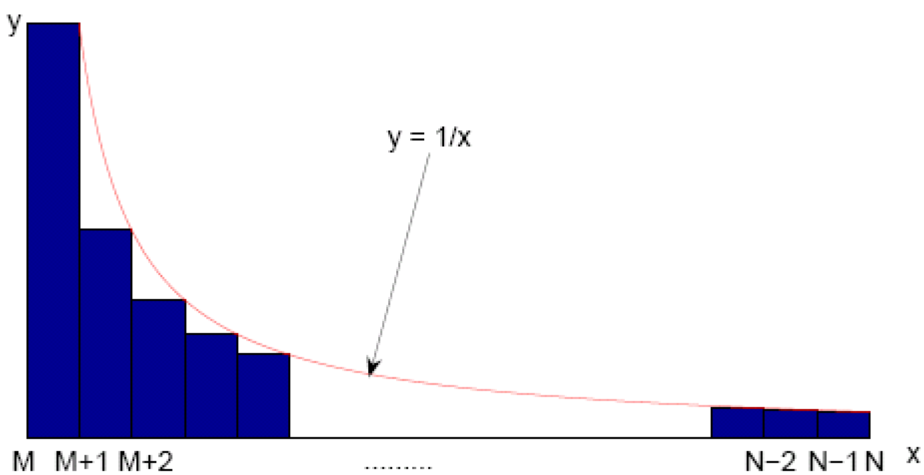
$\frac{1}{M}+\frac{1}{M+1}+\frac{1}{M+2}+\ldots+\frac{1}{N-1}$

Rõ ràng tổng diện tích này lớn hơn diện tích S nằm bên dưới đường cong y=1/x, mà S ta tính được bằng phép tính tích phân là:

$$S=\int_{M}^{N}(1/x)dx=\ln x\Big|_{M}^{N}=\ln N-\ln M=\ln\frac{N}{M}$$

Vậy ta có $f(M,N)>\ln\frac{N}{M}$

Bây giờ ta xét đồ thị kế tiếp



Lúc này rõ ràng phần diện tích bên dưới đường cong lớn hơn tổng diện tích các hình chữ nhật i.e:

$$\int_{M+1}^{N}\left(\frac{1}{x}\right)dx>\frac{1}{M+1}+\frac{1}{M+2}+\ldots+\frac{1}{N-1}$$

$$\Leftrightarrow \ln\frac{N}{M+1}>f(M+1,N)$$

Kết hợp cả 2 đánh giá trên ta được

$$f(M,N)>\ln\frac{N}{M}>\ln\frac{N}{M+1}>f(M+1,N)$$

Tuy nhiên ta đang muốn tìm M (theo N) sao cho

$$f(M,N) > 1 > f(M+1,N)$$

Khi N lớn thì $\ln\frac{N}{M} \approx \ln\frac{N}{M+1}$, do đó chỉ cần chọn M sao cho

$$\ln\frac{N}{M} \approx 1 \Leftrightarrow \frac{N}{M} \approx e \Leftrightarrow M \approx \frac{N}{e} \approx 0.3679N$$

Nhưng với M này thì xác suất hôn trúng hoàng tử là bao nhiêu. Nhớ lại công thức tính xác suất này ta được:

$$P(M,N) = \frac{M}{N}\left(\frac{1}{M} + \frac{1}{M+1} + \frac{1}{M+2} + \ldots + \frac{1}{N-1}\right) \approx \frac{1}{e}\ln\left(\frac{N}{M}\right) \approx \frac{1}{e} \approx 0.3679.$$

Vậy khi áp dụng quy tắc này thì xs hôn trúng hoàng tử là cỡ 37%, và đó là chiến lược tốt nhất mà công chúa có thể áp dụng.

Tuy nhiên đá hết 37% của 1 triệu chú ếch xuông hồ rồi mới quyết định hôn chú cao hơn tiếp theo thì cũng rất lâu. Vẫn còn chuyện phải suy nghĩ trong mô hình của chúng ta.

Hãy cưới đúng hòang tử của mình nhé

Bạn đã đoán ra ý của tôi chưa, tôi đang nói về một chiến lược để chọn bạn đời đấy. Bạn nên hẹn hò với bao nhiêu người (và đá trở lại xuống ao, ack!) trước khi bạn quyết định ổn định, lập gia đình và sinh con (i.e hôn chú ếch,☺), tùy thuộc vào bạn nghĩ người đó tốt đ/v bạn đến cỡ nào (con số trên lưng của họ đấy,☺). Mô hìhn của chúng ta như sau:

1. Từ 1 số lượng người có hạn, ta cần chọn ra người bạn đời cho mình.
2. Các "ứng viên" sẽ lần lượt xuất hiện và ta không biết khi nào hay liệu còn "ứng viên" nào tốt hơn sẽ xuất hiện hay không. Ta muốn chọn bạn đời tốt nhất có thể được.
3. Chiến lược sẽ là hẹn hò với 37% trong số ứng viên (rồi đá họ, ack!) và rồi chọn người tốt nhất xuất hiện kế tiếp. Xác suất thành công của chiến lược này cũng là 37%.

4. Đến đây xuất hiện điều thú vị, liệu chiến lược này có thật sự hợp lý không? Ta có bỏ sót điều gì so với quá trình lựa chọn bạn đời trong thực tế hay không?

(Tuy nhiên tôi không biết thực tế người ta có dùng đến hay có nên dùng quy tắc 37% này hay không ☺)

Chú ếch bạn chọn thật sự tốt đến cỡ nào?

Kissed too soon!!

Tôi có nghe nói rằng khi nam và nữ gặp nhau, 2 bên thường hay xếp hạng lẫn nhau theo 1 thang điểm từ 1 đến 10, các nhà toán học thì rất thông minh (☺) nên họ sẽ dùng thang từ 0 đến 1. Do đó ta thay đổi điều kiện của bài toán ban đầu 1 chút.

"Các chú ếch lúc này được đánh số ngẫu nhiên bằng 1 con số nằm trong khoảng từ 0 đến 1"

Như vậy lúc này công chúa có nhiều thông tin hơn. Lúc này cô biết có 1 số lớn nhất và 1 số nhỏ nhất (1 và 0) và các số của các chú ếch được phân bố đều (nhưng ngẫu nhiên) từ 0 đến 1. Nếu chú ếch đầu tiên mang số 0.99 thì chắc chắn là chú ếch xịn nhất và nên hôn chú ngay. Nhưng nếu chú đầu tiên mang số 0.8 thì sao? Có đủ tốt để được hôn chưa? Hóa ra chiến lược (mà bất kỳ cô nào trên 25 tuổi cũng biết ☺) là bắt đầu với các "tiêu chuẩn" cao và hạ thấp chúng dần dần. Ở đây chúng ta sẽ làm toán tiếp, do đó tiêu chuẩn ở đây nghĩa là "mỗi chú ếch có 1 **con số quyết định**, mà nếu số chú mang dưới con số này thì công chúa không nên hôn". Và sau đây là cách tính con số quyết định ấy.

Cách tính *số quyết định*

Ta xét bài toán tổng quát:

Bài toán: *Có N chú ếch, mỗi chú mang 1 con số ngẫu nhiên trong khoảng từ 0 đến 1. Với các điều kiện như từ đầu đến giờ, ta cần tìm con số quyết định ('tiêu chuẩn') tương ứng với mỗi chú ếch.*

Ký hiệu y_j là số trên lưng chú ếch thứ j,j=1,…,N. Và ta sẽ cố gắng tìm 1 dãy tăng các số $b_j, j = 0,1...,N-1$ (các số quyết định) sao cho nếu $y_j > b_{N-j}$ và y_j là số lớn nhất trong j số đầu tiên thì ta sẽ hôn chú ếch thứ j. Nói cách khác để quyết định hôn chú cuối cùng ta cần $y_n > b_0$, chú kế cuối ta cần $y_{N-1} > b_1$ và cứ thế. Lý do của việc đánh số ngược này sẽ được giải thích ngay.

Giả sử có N-k chú ếch đã nhảy lên và chú thứ N-k mang số lớn nhất y_{N-k}. Ta ký hiệu lại y_{N-k} là y cho gọn. Như vậy còn k chú ếch chưa nhảy lên. Do đó nếu công chúa hôn chú ếch thứ N-k thì xs để chú này là hoàng tử bằng với xs mà k chú còn lại mang số nhỏ hơn y, tức là y^k. Phần tiếp theo hơi khó hơn 1 chút.

Nếu công chúa không hôn chú thứ N-k mà quyết định tìm chú tiếp theo mang số lớn hơn y. Ta có thể tìm xs hôn trúng hoàng tử bằng lý luận như sau:

Giả sử còn lại j chú ếch mang số lớn hơn y. Khi đó j có thể nhận các giá trị từ 1,2,…,k. Ta tính xs thắng ứng với mỗi j rồi cộng lại.

Xs mà trong k chú còn lại có đúng j chú mang số lớn hơn y sẽ là $(1-y)^j y^{k-j}$ nhân với số cách chọn ra j chú từ k chú (tức là C_k^j hay $\binom{k}{j}$), như vậy là:

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

Khi đó xs mà công chúa hôn đúng chú ếch mang số lớn nhất trong số j chú này là 1/j.

Tóm lại khi có đúng j chú mang số lớn hơn y thì xs công chúa hôn đúng hoàng tử sẽ là:

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

Cộng hết tất cả các trường hợp này ta được xs thắng nếu công chúa không hôn chú ếch thứ N-k sẽ là:

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

Như vậy số quyết định cần tìm sẽ là giá trị của y mà tại đó xs thắng nếu công chúa hôn chú thứ N-k cân bằng với xs thắng nếu công chúa không hôn chú này, tức là giá trị y đó phải thỏa phương trình :

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

Chia 2 vế cho y^k ta được

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

Đặt $z = \frac{1-y}{y}$ ta được phương trình

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

Khi đã tìm ra nghiệm dương duy nhất $z = z_+$ của phương trình trên (bạn hãy cm phương trình trên có nghiệm dương duy nhất nhé☺) thì số quyết định thứ k sẽ là

$$b_k = \frac{1}{1+z_+}$$

Ta hãy thử ct vừa tìm được với các giá trị k khác nhau.

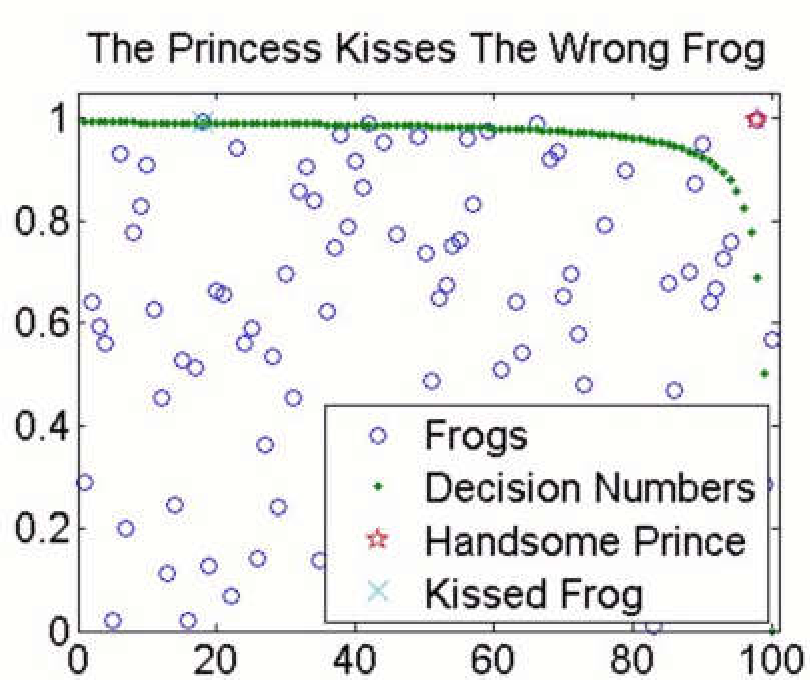
k=0, công chúa chắc chắn sẽ hôn chú cuối cùng nếu nó mang số cao nhất, do đó $b_0 = 0$

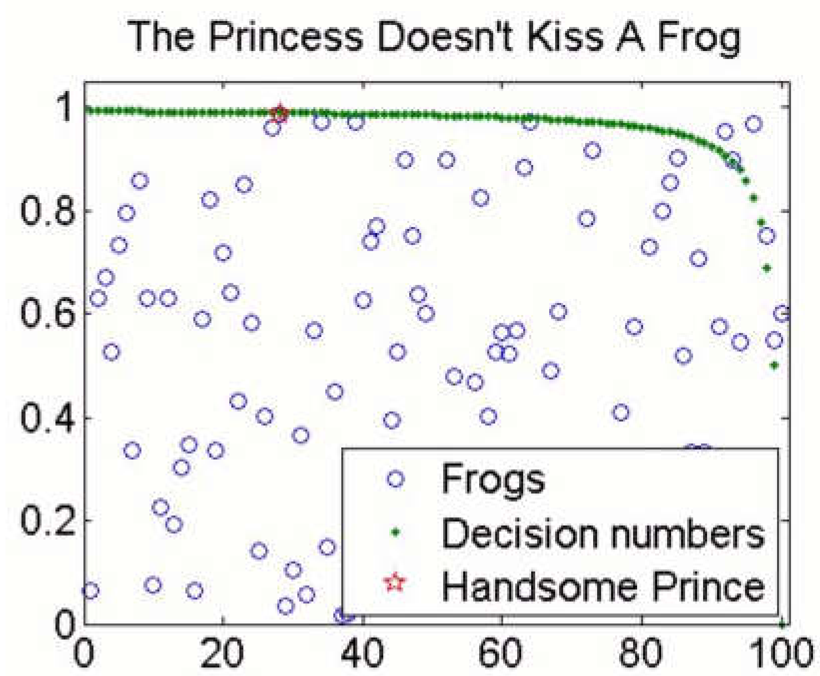
k=1, giá trị trung bình của chú ếch cuối cùng là 0.5, do đó công chúa nên hôn chú kế cuối nếu y_{N-1} là số lớn nhất cô từng nhìn thấy và $y_{N-1} > 0.5$, i.e $b_1 = 0.5$. Khi k=1, phương trình để tìm z_+ là: z-1=0 nên z_+=1 và $b_1 = 0.5$, lý luận và thực tế khớp nhau.

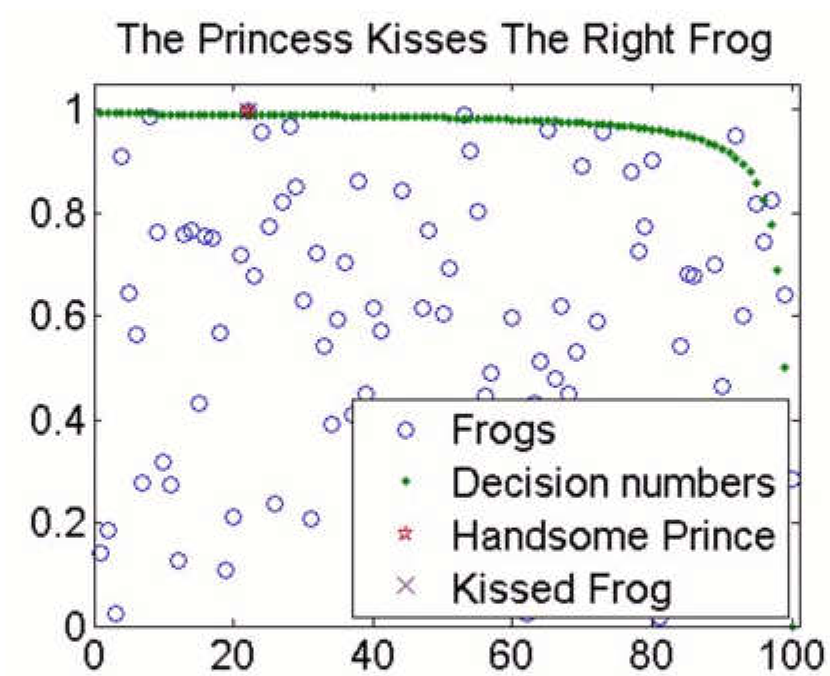
k=2, phương trình để tìm z_+ là: $-1 + 2z + \frac{1}{2}z^2 = 0$, phương trình này có nghiệm dương $z = z_+ = \sqrt{6} - 2$ nên $b_2 = (1 + \sqrt{6})/5 \approx 0.6899$.

Các bạn sẽ thấy các số b_0, b_1, b_2 này trên đồ thị.

Để tìm $b_k, k \geq 3$, ta phải giải các phương trình bậc cao hơn. Lúc này ta có thể dùng phương pháp lặp của Newton để tính xấp xỉ các nghiệm z_+ (pp Newton các em sẽ được học ở ĐH). Các em có thể nhìn thấy 100 số quyết định đầu tiên trên đồ thị.




The Princess Doesn't Kiss A Frog
Frogs
Decision numbers
Handsome Prince


The Princess Kisses The Right Frog
Frogs
Decision numbers
Handsome Prince
Kissed Frog

(còn tiếp)