Giacomo Salvati Pavia, 15-12-2017

L'indovinello delle Uova Fabergé

"La città ha appena aperto un museo di Uova Fabergé che ospita 100 uova, una per ogni piano del palazzo del museo, e il più grande ladro di gioielli al mondo sta già preparando il suo piano per rubarli. La sicurezza nel palazzo è però ai massimi livelli e perciò il ladro può rubare solo un uovo, gettandolo fuori dalla finestra giù nel furgone del complice, che aspetta in strada, pronto a fuggire.
Le uova hanno tutte stessa robustezza e consistenza, ma il loro prezzo varia. Al primo piano si trova l'uovo che vale di meno e, salendo, il valore dei gioielli aumenta, fino ad arrivare all'ultimo piano, che ospita il pezzo più prezioso. Il ladro punta subito gli occhi su quest'ultimo, ma teme che le uova non possano resistere ad una caduta di 100 piani.
Cerca così un modo per trovare quale sia l'uovo più prezioso che riuscirebbe a rubare.
Per sua fortuna al negozio di souvenir del museo sono rimaste due copie delle uova del famoso artista, che non hanno alcun valore ma sono identiche alle originali in robustezza e consistenza; così decide di comprarle e usarle per testare quale urto riescano a sopportare le uova, lanciandole giù dalle finestre del palazzo.
Il ladro può gettare fuori dalla finestra quante volte vuole un uovo, ma quando questo non reggerà l'urto e si romperà non sarà più utilizzabile. Inoltre le guardie potrebbero insospettirsi, quindi il ladro deve cercare di arrivare alla soluzione in meno tempo possibile.
Quanti lanci deve fare al minimo il ladro per essere sicuro di trovare l'uovo più prezioso che regga l'urto, in qualsiasi piano esso si trovi?"

Prendendo spunto da questo indovinello, ho pensato di presentarne la soluzione e generalizzare il problema per n uova e k piani del palazzo.

Chiamiamo:

- x il piano dell'uovo più prezioso che il ladro può rubare;
- T il numero di test richiesti per trovare il piano x nel peggiore dei casi.

(per "peggiore dei casi" intendo il caso in cui x coincida con il piano che richieda più test per essere verificato)

Quindi T dipenderà dalla strategia con cui si pianificano i test, e l'obiettivo è trovare quella per cui T è minimo.

Supponiamo di avere $n = 1$.
Con un solo uovo siamo costretti a testare il lancio da ogni piano, uno ad uno, partendo dal primo (il più basso).
Quindi nel peggiore dei casi $x = k$, quindi $T = k$.

Ora poniamo $n = 2$.
Sappiamo che una volta rotto il primo uovo, il secondo dovrà essere testato di piano in piano, a partire dall'ultimo in cui il primo uovo non si è rotto. Possiamo così pensare di testare il primo ad intervalli regolari di piani, per esempio all'm-esimo, al 2m-esimo, al 3m-esimo, e così via… ($m \in \mathbb{N}$)
Il peggiore dei casi sarà quando il piano x sia appena prima di un piano testato con il primo uovo, che al massimo può essere il k-esimo.
Quindi T si ottiene per $x = k - 1$.

Consideriamo ora T_i il numero di test necessari con il i-esimo uovo per individuare il piano x, con $i \in [1, n]$

Calcoliamo:
- $T_1 = \lceil k/m \rceil$
- $T_2 = m - 1$

Per definizione sappiamo che $T = \sum_{i=1}^{n} T_i$ perciò $T = \lceil k/m \rceil + m - 1$

Il nostro obiettivo è trovare la strategia per cui T sia minimo, perciò in questo caso si ottiene per m tale che $km^{-1} + m - 1$ sia minimo.
Per m molto piccolo T_1 sarà molto grande, viceversa per un m grande aumenterà molto T_2. È facilmente intuibile perciò come si comporta la funzione $T(m)$ al variare di m: tracciandone il grafico si può osservare una curva convessa, perciò nel punto di minimo assoluto la retta tangente alla funziona avrà pendenza uguale a 0.
Quindi il minimo della funzione si ottiene per:
$\frac{d}{dm}(km^{-1} + m - 1) = -km^{-2} + 1 = 0$ quindi per $m = \lceil \sqrt{k} \rceil$

Prendiamo in considerazione il caso del problema posto all'inizio, ovvero in cui $k = 100$.
Abbiamo dimostrato che, se testiamo il primo uovo a intervalli regolari, questi intervalli devono essere composti da $m = \lceil \sqrt{100} \rceil = 10$ piani.
Ora introduciamo una nuova notazione: $t(j)$ è il numero di test necessari per verificare se il j-esimo piano coincida con il piano x o meno. Perciò $T = max(t(j))\ per\ j \in [1, k]$.

Quindi per $x = 99$ si ha $t(99) = T = 19$, poiché bisogna testare il primo uovo 10 volte (ai piani 10, 20, 30, …, 100) e il secondo altre 9 (ai piani 91, 92, 93, …, 99).
Adottando questa strategia, però, bisogna fare in media più test nel caso in cui x è vicino a 100 rispetto a quando x è piccolo. Infatti in media $\overline{t_2(x)} = (1 + 2 + 3 + \cdots + 9)/9 = 5$, indipendentemente dalla grandezza di x, mentre $t_1(x)$ è linearmente proporzionale a x, poiché per un generico piano x fissato $t_1(x) = \lceil x/m \rceil$.

Possiamo dunque pensare che questa non sia la strategia ideale, poiché per x piccoli si ha t(x) piccoli e, viceversa, per x grandi t(x) grandi. Infatti è richiesta una strategia in cui sia minimo $T = max(t(j))$. È più conveniente una strategia che "bilanci" questa situazione, infatti non è necessario che i piani dai quali testiamo il primo uovo debbano essere separati da intervalli uguali.
Bisogna perciò trovare una strategia per cui $t(j)$ è costante, così se da una parte $min(t(j))$ crescerà, $T = max(t(j))$ diminuirà.

Immaginiamo di testare per la prima volta il primo uovo dal piano p e la seconda volta dal piano p+q. Ora tra i piani compresi fra il primo e il p-esimo, quello che richiederà più test per dimostrare che coincida con il piano x è il (p-1)-esimo. Perciò per $j \in [1, p]$

$$max(t(j)) = t(p-1) = t_1(p-1) + t_2(p-1) = 1 + (p-1) = p$$

Se invece prendiamo in considerazione il secondo intervallo, il piano che richiederà più test sarà quello appena prima del piano p+q. Perciò per $j \in [p+1, p+q]$

$$max(t(j)) = t(p+q-1) = t_1(p+q-1) + t_2(p+q-1) = 2 + (q-1) = q + 1$$

Per ottenere $max(t(j))$ costante all'interno dei due intervalli dobbiamo considerare $p = q + 1$, cioè $q = p - 1$. Quindi il secondo intervallo avrà ampiezza uguale al primo meno un piano.

Da notare il fatto che, preso un qualsiasi intervallo di m piani, compreso fra due piani dai quali viene testato il primo uovo (z e z+m), per $j \in [z+1, z+m]$ $t_1(j)$ è costante, mentre $max(t_2(j)) = t_2(z+m-1) = m-1$ cioè è uguale all'ampiezza dell'intervallo meno 1.

Allo stesso modo possiamo considerare il terzo intervallo, così come i successivi, ed eguagliando il $max(t(j))$ di ognuno di questi otterremo che ogni intervallo contiene un piano in meno dell'intervallo precedente.
Quindi se il primo uovo viene lanciato prima da un piano h, poi sarà testato dal piano $h + (h - 1)$, poi dal piano $h + (h-1) + (h-2)$, … fino ad arrivare a un intervallo di un solo piano.
Ora è necessario trovare il valore di h. Sappiamo che il palazzo ha k piani in totale, perciò $h + (h-1) + (h-2) + \cdots + 1 = h(h+1)/2 \geq k$ perciò dobbiamo trovare un numero triangolare maggiore di k.
Inoltre $T = max(t(j)) = t(h-1) = t(2h-2) = \cdots = t(ah - a(a-1)/2 - 1) = h$ perciò vogliamo prendere h minore possibile.
Quindi per $h(h+1) \geq 2k$ ovvero $h^2 - h - 2k \geq 0$, l'h minimo che soddisfa l'equazione sarà

$$h = \left\lceil \left(-(-1) + \sqrt{(-1)^2 - 4 * 1 * (-2k)}\right)/2 \right\rceil = \left\lceil (1 + \sqrt{1+8k})/2 \right\rceil$$

Perciò con due uova (per $n = 2$), con la migliore strategia, $T = h = \lceil (1 + \sqrt{1+8k})/2 \rceil$.
In particolare la soluzione dell'indovinello proposto in prima pagina è $T = \lceil (1 + \sqrt{1 + 8 * 100})/2 \rceil = \lceil (1 + \sqrt{801})/2 \rceil = 19$

Poniamo ora $n = 3$.
Una volta rotto il primo uovo conosciamo quale strategia attuare per trovare il piano x nell'intervallo di uova "papabili" rimaste, ovvero tra l'ultimo piano dal quale il primo uovo non si è rotto e il piano dal quale questo si è rotto.
Ma da quali piani bisogna testare il primo uovo?

Supponiamo di lanciare per la prima volta il primo uovo dal piano p_1, allora:

$$\exists!\ h \ tale\ che\ (h-1)h/2 < p_1 \leq h(h+1)/2$$

Ora se il primo test avvenisse in un qualsiasi piano p compreso fra i piani $(h-1)h/2$ e $h(h+1)/2$

$$per\ j \in [1, p]\ max(t(j)) = t_1(j) + max(t_2(j) + t_3(j)) = 1 + h$$

Perciò, per massimizzare l'ampiezza del primo intervallo, è più conveniente testare il primo uovo da un piano p_1 dove $p_1 = h(h+1)/2\ con\ h \in \mathbb{N}$.

In altre parole:
il numero massimo di test da effettuare per verificare che un piano p del primo intervallo, cioè compreso fra il piano 1 e il piano da cui viene lanciato per la prima volta il primo uovo, è costante per valori di p compresi fra gli stessi due numeri triangolari consecutivi. Perciò per massimizzare l'ampiezza dell'intervallo, senza variare $max(t(j))\ per\ j \in [1, p]$, bisogna prendere p numero triangolare, cioè scrivibile nella forma $h(h+1)/2\ con\ h \in \mathbb{N}$.

Ora possiamo ripetere lo stesso ragionamento per gli altri intervalli, che quindi avranno tutti un'ampiezza $A = w(w+1)/2\ con\ w \in \mathbb{N}$ che varia a seconda dell'intervallo.

Prendiamo in considerazione i primi due intervalli. Il primo (compreso fra il piano 1 e il piano A_1) avrà ampiezza $A_1 = h_1(h_1+1)/2$ mentre il secondo (dal piano A_1+1 al piano A_1+A_2) avrà ampiezza $A_2 = h_2(h_2+1)/2$.
Ora $per\ j \in [1, A_1]\ max(t(j)) = t_1(j) + max(t_2(j) + t_3(j)) = 1 + h_1$

Infatti con il primo uovo dobbiamo fare solo un test, poiché stiamo considerando il primo intervallo, mentre con gli altri due, al massimo, ne dobbiamo fare h_1, come dimostrato nel punto precedente, dove $n = 2$.
Mentre $per\ j \in [A_1 + 1, A_1 + A_2\ ]\ max(t(j)) = t_1(j) + max(t_2(j) + t_3(j)) = 2 + h_2$
Vogliamo che $max(t(j))$ resti costante di intervallo in intervallo, perciò $1 + h_1 = 2 + h_2$, ovvero $h_2 = h_1 - 1$.
Il ragionamento può essere ripetuto per tutti gli altri intervalli e si giunge alla regola più generale: $h_{n+1} = h_n - 1$, ovvero $h_n = h_1 - (n - 1)$.

Perciò $A_1 = h(h + 1)/2, A_2 = (h - 1)h/2, A_3 = (h - 2)(h - 1)/2$, ... fino ad arrivare all'intervallo che abbia ampiezza $A_h = 1 * 2/2 = 1$

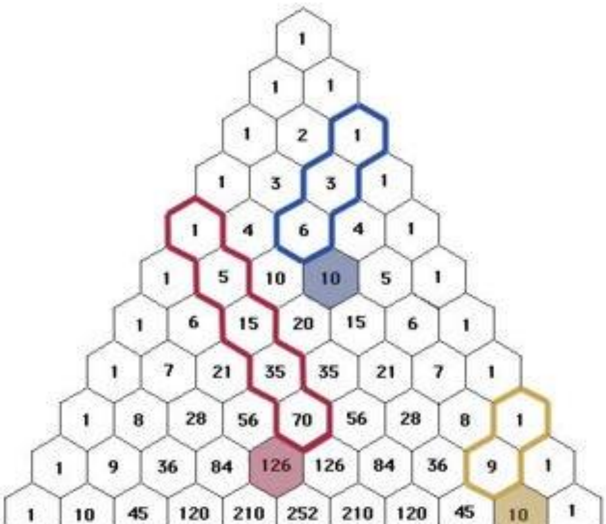
Quindi l'unica variabile che entra in gioco è h. Quale valore dobbiamo assegnarle?
Sommando i vari intervalli dobbiamo arrivare ad un numero di piani maggiore di quelli del palazzo. Perciò $(1 * 2)/2 + (2 * 3)/2 + (3 * 4)/2 + \cdots + h(h + 1)/2 \geq k$

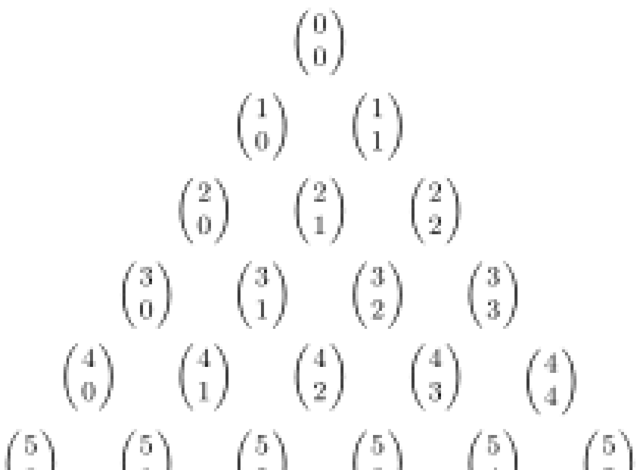
Ma la somma dei primi h numeri triangolari è uguale all'h-esimo numero tetraedrico, proprietà che si può ricavare dal triangolo di Tartaglia.

Cioè $\sum_{i=1}^{h} A_i = \sum_{i=1}^{h} i(i + 1)/2 = h(h + 1)(h + 2)/(2 * 3) = \binom{h+2}{3} \geq k$

(Per numero tetraedrico, o numero piramidale triangolare, si intende un numero figurato che rappresenta un tetraedro: si immagini di avere n sfere; n è un numero tetraedrico se le n sfere possono essere disposte nello spazio in modo da formare un tetraedro.)

Perciò h è il minore valore che soddisfi l'equazione di cui sopra e $per\ j \in [1, k]$
$$T = max(t(j)) = max(t_1(j) + (t_2(j) + t_3(j))) = 1 + h = 2 + (h - 1) = \cdots = h + 1$$





Nel Triangolo di Tartaglia ogni numero è la somma dei due soprastanti e, di conseguenza, sommando i numeri in direzione dei lati obliqui (come mostrato nella figura a sinistra) si ottiene il numero sottostante non allineato con i precedenti. Inoltre il Triangolo di Tartaglia possiede alcune proprietà importanti riguardo i coefficienti binomiali (come mostrato in figura 2).

Generalizziamo per un qualsiasi n.
Il ragionamento condotto nel punto precedente è utilizzabile anche per $n = 4$.
Gli intervalli tra i piani dai quali sarà testato il primo uovo avranno ampiezza
$A = w(w+1)(w+2)/6 \; con \; w \in \mathbb{N}$ che diminuisce di 1 di intervallo in intervallo.
Perciò la disequazione da risolvere è $\sum_{i=1}^{h} \binom{i+2}{3} = \binom{h+3}{4} \geq k$.
Quindi:

$per \; j \in [1, k] \; T = max(t(j)) = max(t_1(j) + (t_2(j) + t_3(j) + t_4(j))) = 1 + (1 + h) = h + 2$
(essendo $max(t(j)) \; per \; j \in [1, k]$ costante di intervallo in intervallo, posso semplicemente calcolare $max(t(j)) \; per \; j \in al \; primo \; intervallo$).

Allo stesso modo si procede per $n = 5$, dove otterremo $T = h + 3$ con h minimo valore che soddisfa la disequazione $\sum_{i=1}^{h} \binom{i+3}{4} = \binom{h+4}{5} \geq k$

In generale, per un qualsiasi $n \geq 2$, $T = h + n - 2$ con h minimo valore che soddisfa l'equazione $\binom{h+n-1}{n} \geq k$. $(*)$

Non è semplice trovare h in funzione di n e k, perciò mi sono aiutato a trovare la soluzione con un grafico Geogebra.
Dopo aver costruito gli slider per le variabili n e k, si traccia il grafico della funzione $f(x) = \binom{x+n-1}{n}$ e la retta $y = k$ e si può trovare il punto P di intersezione con uno strumento dell'applicazione (x sostuisce la h della disequazione $(*)$). Quindi si può trovare il valore di x come valore dell'ascissa del punto P arrotondato per eccesso.
Ora $T = x(P) + n - 2$ e questo valore lo possiamo calcolare nei fogli di calcolo.

Rimane un problema.
Tuttavia, variando il numero di uova e di piani del palazzo sul file Geogebra, si può notare una "falla nel sistema", un errore concettuale nel procedimento condotto finora.
Abbiamo sempre considerato di usare tutte le uova a disposizione, supposizione logica ed efficace, ma solo per dei valori di n e k "ragionevoli". Infatti se si prende un numero di piani non abbastanza grande rispetto al numero di uova, può essere che alcune di queste non siano necessarie e, anzi, non siano proprio da utilizzare. Prendiamo ad esempio il caso in cui $k = 20$ e $n = 10$. Per testare 20 piani non sono assolutamente necessarie tutte le uova. Usandone solo 3, infatti, sono necessari $T = 5$ test, mentre usandole tutte e 10 ne servono addirittura $T = 11$.

Cosa sta succedendo sul piano matematico?

Finora abbiamo trovato $T = h + n - 2$ con h minimo che risolve $\binom{h+n-1}{n} \geq k$.
Con l'aumentare di n, diminuisce sicuramente h, ma il numero di test necessari dipende anche dal numero di uova usate.
Consideriamo il Triangolo di Tartaglia: stiamo cercando il minimo valore di y tale che $\binom{y}{n} \geq k$, perciò stiamo cercando il primo valore maggiore di k che sta sulla n-esima riga obliqua (partendo da sinistra, con la zeresima). Così otterremo $T = y - 1$.
Tuttavia possono esserci numeri maggiori di k anche sulle righe precedenti alla n-esima. In particolare consideriamo il valore $\binom{w}{v}$: è il massimo numero di piani controllabili con v uova e $T = w - 1$ test a disposizione. Se $v < n \; e \; w < y$, significa che possiamo verificare ogni piano del palazzo senza usare tutte le n uova e, inoltre, facendo meno test.

Quando può accadere ciò?

Sappiamo che T dipende solo da w, perciò dalla riga orizzontale (partendo dall'alto, con la zeresima) alla quale il numero $\binom{w}{v}$ appartiene. Il massimo valore che può assumere un numero della w-esima riga è quello centrale, ovvero $\binom{w}{\lfloor w/2 \rfloor}$.
Ora, una volta forniti n e k, cerchiamo il primo numero $\binom{w}{v}$ maggiore di k che compare nel Triangolo di Tartaglia, partendo dalla zeresima riga orizzontale (cioè con w minimo). Questo numero sarà sicuramente minore o uguale al valore centrale della stessa riga, ovvero $\binom{w}{\lfloor w/2 \rfloor}$.
Ora se $\lfloor w/2 \rfloor < n$, allora $y \geq w$, poiché non esiste valore maggiore di $\binom{w}{\lfloor w/2 \rfloor}$ sulla w-esima riga orizzontale, quindi nemmeno in quelle precedenti.
Se $\lfloor w/2 \rfloor < n$, allora, non è necessario né efficiente usare tutte le uova a disposizione per avere la certezza di trovare l'uovo Fabergé desiderato. Infatti facendo solo $T = w - 1$ test, anziché $y - 1$, si può trovare il piano richiesto.
In conclusione:
Troviamo i valori interi minimi di w e y tali che $\binom{w}{\lfloor w/2 \rfloor} \geq k$ e $\binom{y}{n} \geq k$.
Se $\lfloor w/2 \rfloor < n \Rightarrow\ T = w - 1$, altrimenti $T = y - 1$.