Il problema dell'album di figurine

Un album contiene n figurine, si suppone che le figurine si comprino una alla volta e che siano tutte equiprobabili (cioè non ci siano figurine "rare"). Ci chiediamo quante figurine vanno acquistate in media per completare l'album.

Indichiamo con $T_1, T_2, \ldots$ i tempi di attesa, cioè T_k è il numero medio di figurine da acquistare per averne k distinte avendone già $k-1$ distinte. Si ha chiaramente $T_1 = 1$. La probabilità di trovare subito la seconda figurina è $(n-1)/n$ e dalla teoria dei processi di Bernoulli si trova $T_2 = n/(n-1)$. Analogamente si trova $T_k = n/(n-k+1)$. Dunque per avere l'album riempito con m figurine bisognerà in media acquistarne

$$N_m = \sum_{k=1}^{m} T_k = n \sum_{k=1}^{m} \frac{1}{n-k+1} = n \sum_{j=n-m+1}^{n} \frac{1}{j} = n\big(S(n) - S(n-m)\big)$$

dove abbiamo posto $j = n - k + 1$ e dove $S(h) = \sum_{j=1}^{h} 1/j$.

Usiamo lo sviluppo asintotico (per h grande)

$$S(h) = \log h + \gamma + o(1)$$

dove γ è la costante di Eulero-Mascheroni ($\gamma \simeq 0.577215665$). Si trova allora il numero medio di figurine per completare tutto l'album

$$N_n = n\big(\log n + \gamma\big) + o(n);$$

mentre per completarne solo metà, la media di figurine da acquistare sarà

$$N_{n/2} = n\big(\log n + \gamma - \log n/2 - \gamma + o(1)\big) = n \log 2 + o(n).$$

Se prendiamo ad esempio $n = 100$, per un album parzialmente completo ci vorrà in media un numero di acquisti circa pari a

$N_{25} \approx 29$	album completo per un quarto
$N_{50} \approx 69$	album completo per metà
$N_{75} \approx 139$	album completo per tre quarti
$N_{100} \approx 518$	album completo per intero

che corrisponde all'intuizione che una collezione di figurine si riempie molto facilmente all'inizio e molto più difficilmente man mano che le caselle vengono occupate.