Paradoxe des anniversaires

Dans un groupe de 23 personnes, il y a une chance sur deux que deux personnes aient le même jour d'anniversaire, et dans un groupe de seulement 80 personnes, il y a de très fortes chances que deux personnes aient le même jour d'anniversaire (1 chance sur 10000 que ça ne se produise pas). Calculons pourquoi il en est ainsi. Soit

$$\bar{p}_E(n) = \mathbb{P}\big[\text{éléments de } (\boldsymbol{e_i})_{i=1}^n \in E^n \text{ tous différents}\big]$$

Pour calculer cette probabilité, on peut dénombrer les n-uplets (listes ordonnées) *sans répétitions*, c'est-à-dire les choix ordonnés de n objets dans une collection de $N = |E|$ objets, c'est-à-dire les *arrangements* :

$$\mathrm{A}_N^n = \frac{N!}{(N-n)!}$$

que l'on rapporte à tous les n-uplets possibles, au nombre de $|E^n| = N^n$:

$$\bar{p}_E(n) = \begin{cases} 0 \text{ si } n > |E| & \text{(lemme des tiroirs)} \\ \dfrac{\mathrm{A}_N^n}{|E^n|} = \dfrac{N!}{(N-n)!\,N^n} = \displaystyle\prod_{k=0}^{n-1} \frac{N-k}{N} & \text{si } n \leqslant |E| \end{cases}$$

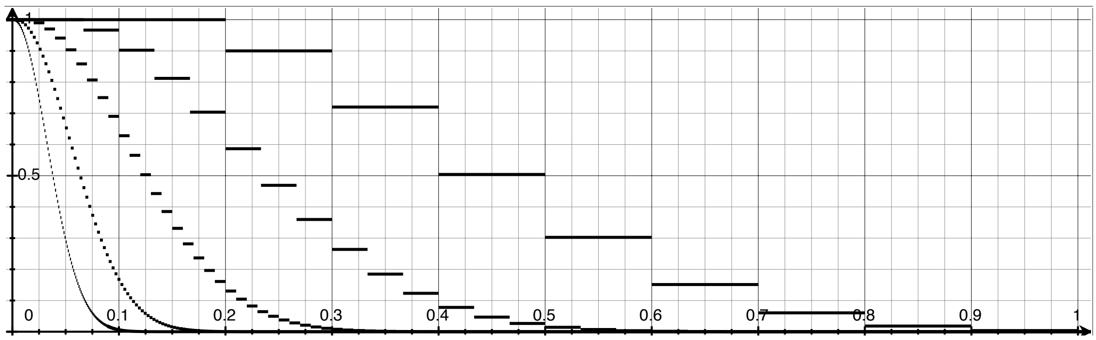

Figure 1. $\bar{p}_E(n)$ pour $N = 10, 30, 100, 365, 1000$, en fonction de $x = n/N$.

Si l'on veut une approximation pour $n \ll N$ (en pratique, bon pour n jusqu'à quelques $\frac{N}{\ln^2 N}$), on écrit

$$\bar{p}_N(n) = \prod_{k=0}^{n-1} \left(1 - \frac{k}{N}\right) \simeq \prod_{k=0}^{n-1} \mathrm{e}^{-\frac{k}{N}} = \mathrm{e}^{-\frac{n(n-1)}{2N}}$$

Et alors,

$$\boxed{p_N(n) = \mathbb{P}\big[\text{au moins deux éléments identiques dans } (\boldsymbol{e_i})_{i=1}^n \in E^n\big] = 1 - \bar{p}_E(n) \approx 1 - \mathrm{e}^{-\frac{n(n-1)}{2N}}}$$

la dernière approximation se trouvant être très bonne $\forall n$ dès $N = 30$. On a

$$\boxed{p_{365}(23) = 51\% \quad \text{et} \quad p_{365}(80) = 99.99\%}$$