Le hasard, 0123456789 et π.

Claudine Schwartz

Quelles *chances* a-t-on de trouver la suite 0123456789 dans une suite de chiffres choisis au hasard ?

Est-ce $(1/10)^{10}$?
Est-ce 1 ?

On entend par *liste de chiffres au hasard* une liste de chiffres résultant de choix indépendants de chiffres où, pour chaque élément de la liste, chaque chiffre avait une probabilité 0,1 d'être choisi.

On peut dire que chacune de ces réponses est une réponse juste à une question plus clairement formulée que celle que nous avons posé :
►Le nombre $(1/10)^{10}$ est la probabilité de trouver la suite L=0123456789 entre le chiffre n et le chiffre n+9, n étant déterminé à l'avance.
►On a une probabilité 1, dans une suite infinie de nombres au hasard de trouver, quelque part la suite 0123456789.

L'apparition d'une série remarquable, définie avant d'avoir observé la liste de chiffres au hasard, pose en pratique question dans la situation où la suite de chiffres dont on pense qu'elle est au hasard est finie et où le rang d'apparition du premier terme de la liste L est « libre » (non choisi a priori) . Donc on va reformuler la question :

Quelle probabilité a-t-on de trouver la suite *L*=0123456789 quelque part dans une suite de *N* chiffres choisis au hasard ?

►Un raisonnement simple consiste à dire : la liste des N premiers termes de la suite peut être découpée en N/10 sous listes de taille 10. La probabilité qu'au moins une de ces sous-listes soit L est [1]:

$$r_N=1-(1-10^{-10})^{N/10} \approx 1-\exp(-N\times 10^{-11}).$$

La probabilité p_N de trouver la liste L dans les N premières décimales cherchées est plus grande, car L peut-être à cheval sur deux sous-listes :

$$r_N < p_N.$$

► On démontre par un calcul classique de probabilités [2] les résultats suivants, où $a_n = p_N$ et $N=10^n$

- $a_6 \approx 0.0001$
- $a_7 \approx 0.0010$
- $a_8 \approx 0.0099$
- $a_9 \approx 0.0951$

---

[1] Il suffit de considérer Ln(1- r_N) pour obtenir cette approximation.
[2] On peut utiliser une chaîne de Markov pour cela, et on se ramène pour obtenir les résultats ci-dessous au calcul numérique de la puissance N-ème d'une matrice carré d'ordre 11.

- $a_{10} \approx 0.6321$
- $a_{11} \approx 0.9999$

Ainsi,

-on a moins d'une chance sur 10 de trouver *L* sur une suite de moins de 1 milliard de termes,
-sur une liste de 10 milliards de chiffres au hasard, la probabilité de trouver au moins une fois *L* est 0,63.
-on est quasiment sûr d'observer *L* si la suite est de taille au moins 100 milliards.

Les décimales du nombre π se comportent comme si elles avaient été choisies au hasard : enfin, pour celles qu'on a calculées jusqu'à présent et pour les procédures de validation considérées. On connaît beaucoup de décimales de π, et on a déjà repéré 6 apparitions de *L* ; voici les rangs d'apparition de *L* :

17 387 594 880
26 852 899 245
30 243 957 439
34 549 153 953
41 952 536 161
43 289 964 000

(voir http://mathworld.wolfram.com/PiDigits.html).[3]

Par rapport au comportement au hasard des décimales de π, rien à signaler.

[3] voir http://mathworld.wolfram.com/PiDigits.html.