Loi de Benford

On ouvre un journal et on relève le *premier chiffre significatif* de tous les nombres rencontrés dans ce journal (pris en valeurs absolues) ; le *premier chiffre significatif* est, par exemple, 2 pour le nombre 2543,34 et 6 pour le nombre 0,00678.

Ce premier chiffre significatif est un entier compris entre 1 et 9 et la distribution observée est proche de la loi de Benford, définie par :

$$\forall k \in [\![1, 9]\!],\ P(\{k\}) = \mathrm{Log}\left(1+\frac{1}{k}\right) \qquad \text{(Log désignant le logarithme décimal)}$$

k	1	2	3	4	5	6	7	8	9
Log (1+1/k)	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

On vérifie : $\forall Y \in \mathbb{R}_+^*,\ \exists!\ Z \in [1, 10[,\ \exists!\ n \in \mathbb{Z},\ Y = Z \times 10^n$.

On a en effet : $Y = 10^{\mathrm{Log}Y} = 10^{\mathrm{Log}Y - [\mathrm{Log}Y]}10^{[\mathrm{Log}Y]}$; on pose $n = [\mathrm{Log}Y]$ (il s'agit bien d'un entier relatif) et $Z = 10^U$ avec $U = \mathrm{Log}Y - [\mathrm{Log}Y]$. Comme U appartient à [0 , 1[, le nombre réel Z appartient à l'intervalle [1 , 10[de $\mathbb{R}$. Le premier chiffre significatif de Y est alors la partie entière de Z. Avant les calculatrices, du temps où l'on calculait les logarithmes à partir de tables, la partie entière (resp. décimale) de Log Y était appelée *caractéristique* (resp. *mantisse*) de Log Y.

Proposition :

Si U suit une loi uniforme sur [0, 1[, alors la partie entière de 10^U suit une loi de Benford.

Preuve

Soit U une v.a.r. de loi uniforme sur [0, 1[, on pose : $Z = 10^U$ et $T = [Z]$ (partie entière de Z). On a :

$$\forall z \in [1, 10[,\ F_Z(z) = P(Z \le z) = P(U \le \mathrm{Log}(z)) = \mathrm{Log}(z) \text{ car } \mathrm{Log}(z) \in [0, 1[\text{ et } U \sim U_{[0,1[}$$

$$\forall k \in [\![1, 9]\!],\ P(T = k) = F_Z(k+1) - F_Z(k) = \mathrm{Log}(k+1) - \mathrm{Log}(k) = \mathrm{Log}\left(1+\frac{1}{k}\right)$$

donc T suit la loi de Benford.

Représentation graphique de la f.d.r. de $Z = 10^U$ (continue)
et de la f.d.r. de $T = [Z]$ (en escalier) où U est une v.a.r. de loi uniforme sur [0, 1[).