

Schéma de Bernoulli et planchettes à clous

par P.L. HENNEQUIN

Le schéma de Bernoulli qui étudie la somme de n variables indépendantes de même distribution (1 avec probabilité p , 0 avec probabilité $q = 1-p$) est fondamental tant du point de vue de ses applications comme modèle de répétition des expériences, que du point de vue théorique car amenant aux deux théorèmes fondamentaux du calcul des probabilités : loi des grands nombres et théorème de la limite centrale.

Il est donc agréable de disposer d'un montage physique relevant de ce schéma. L'idée d'un tel dispositif remonte, semble-t-il, à F. Galton, connu par ailleurs pour avoir étudié et résolu le problème de l'extinction des noms de famille et l'évolution des processus multiplicatifs de population.

Le dispositif consiste à faire descendre des billes le long d'un plan, plus ou moins incliné, ces billes rencontrant des obstacles (clous) disposés en quinconce comme l'indique la figure 1.

Si on admet que chaque fois qu'une bille rencontre un obstacle elle le contourne par la droite ou par la gauche avec la même probabilité $1/2$ et ceci quel qu'ait été son trajet antérieur, alors les déviations successives de la bille constituent un schéma de Bernoulli avec $p=q=\frac{1}{2}$ et la bille

occupe la k ème case du n ème rang suivant une loi binomiale, c'est-à-dire avec une probabilité $\frac{1}{2^n} \binom{n}{k}$.

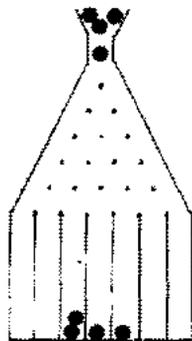


Figure 1

L'intérêt du dispositif consiste en sa rapidité : il est facile de faire descendre plusieurs centaines de billes en quelques minutes* et d'avoir ainsi une approche statistique de la distribution binomiale et, si n est grand, de la courbe en cloche.

* Il y a donc, en fait, deux schémas : la répétition des chocs pour une bille (schéma de Bernoulli) et la succession des billes (répétition du tirage d'une variable binomiale).

A. Rényi a remarqué, il y a déjà longtemps, que ce dispositif ne relevait du modèle de Bernoulli que si l'on prenait soin d'annuler la composante horizontale de la vitesse de la bille entre deux chocs. Il est intuitif en effet qu'une bille qui vient de droite a une probabilité plus forte de continuer son trajet vers la gauche que vers la droite. Nous réserverons donc le nom de "planchettes de Galton" aux dispositifs qui contraignent la bille à passer dans un couloir vertical avant chaque nouveau choc. Il est facile de réaliser un tel dispositif avec des carreaux hexagonaux comme le montre la figure 2.**

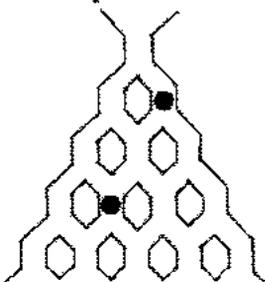


Figure 2

Par contre, la condition d'indépendance entre chocs successifs n'est pas réalisée dans la plupart des dispositifs qu'on rencontre aux Etats-Unis, sous le nom de "Probability Machine", par exemple dans les musées scientifiques, où les obstacles sont de petits cylindres en plexiglass, ou dans des dispositifs plus frustes qu'on peut fabriquer avec une planche et des clous. C'est cette dernière situation que nous nous proposons de modéliser pour la comparer à la planchette de Galton.

Nous allons faire l'hypothèse suivante :

— au premier obstacle, la bille part à droite ou à gauche avec la même probabilité $\frac{1}{2}$; ensuite, si elle vient de droite (et quelle qu'ait été sa trajectoire antérieure), elle part

$$\left\{ \begin{array}{l} \text{à gauche avec probabilité } p > \frac{1}{2} \\ \text{à droite avec probabilité } q = 1 - p \end{array} \right.$$

et si elle vient de gauche, de même, elle part à droite avec probabilité $p > \frac{1}{2}$ et à gauche avec probabilité q .

(Bien entendu, nous retrouvons la situation de Bernoulli dans le cas $p = \frac{1}{2}$).

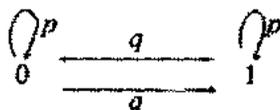
$$\text{Soit } X_j = \begin{cases} 1 & \text{si le } j^{\text{ème}} \text{ choc mène à droite} \\ 0 & \text{si le } j^{\text{ème}} \text{ choc mène à gauche} \end{cases}$$

Notre hypothèse se formule ainsi en termes de probabilités conditionnelles :

$$\begin{aligned} P(X_{j+1}=0 | X_j=0) &= P(X_{j+1}=1 | X_j=1) = p \\ \text{et } P(X_{j+1}=1 | X_j=0) &= P(X_{j+1}=0 | X_j=1) = q \end{aligned}$$

** Dans son livre "Initiation aux Statistiques et aux probabilités au moyen d'expériences simples", publié chez DUNOD en 1969, E.C. BERKELEY signale qu'il a fait fabriquer un tel dispositif aux U.S.A. et qu'il l'a baptisé "HEXSTAT".

Nous pouvons donc représenter les transitions de X_i à X_{i+1} par le diagramme suivant :



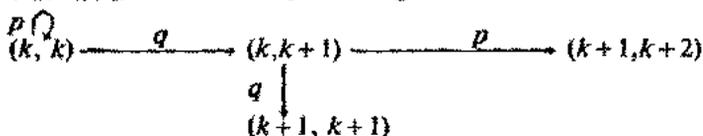
Nous supposons qu'au premier choc il y a équiprobabilité, donc

$$P(X_1=0) = P(X_1=1) = \frac{1}{2}$$

L'abscisse de la bille après le $n^{\text{ième}}$ choc est égale à

$$S_n = \sum_{i=1}^n X_i$$

S_{n+1} dépend à la fois de S_n et de S_{n-1} ; les transitions pour le couple (S_n, S_{n+1}) sont décrites par le diagramme :



qui se déduit du diagramme précédent.

$$\text{Si on pose pour } n > 1 \quad \begin{cases} P_k^n = P(S_n = k, S_{n-1} = k) \\ Q_k^n = P(S_n = k, S_{n-1} = k-1) \end{cases}$$

et $k < n$

on a, avec la convention $S_0 = 0$,

$$\begin{cases} P_0^1 = \frac{1}{2} & , & P_k^1 = 0 & \text{pour } k \neq 0 \\ Q_1^1 = \frac{1}{2} & , & Q_k^1 = 0 & \text{pour } k \neq 1 \end{cases} \quad (1)$$

et, en utilisant notre diagramme et la règle des probabilités composées,

$$\text{pour } n > 1 \quad \begin{cases} P_k^{n+1} = p P_k^n + q Q_k^n \\ Q_k^{n+1} = q P_{k-1}^n + p Q_{k-1}^n \end{cases}$$

On en déduit donc les P_k^n et Q_k^n par récurrence sur n puis

$$P(S_n = k) = P_k^n + Q_k^n$$

(en effet, si $S_n = k$, S_{n-1} ne peut être qu'égal à $k-1$ ou à k).

On en déduit par exemple

$$P_0^2 = \frac{p}{2} \quad , \quad P_1^2 = \frac{q}{2} \quad , \quad Q_1^2 = \frac{q}{2} \quad , \quad Q_2^2 = \frac{p}{2}$$

$$\text{d'où } P(S_2=0) = \frac{p}{2} = P(S_2=2)$$

$$P(S_2=1) = q$$

$$P_0^3 = Q_3^3 = \frac{p^2}{2}, \quad P_1^3 = Q_2^3 = \frac{q}{2}, \quad P_2^3 = Q_1^3 = \frac{qp}{2};$$

d'où la distribution de S_3 :

| k | 0 | 1 | 2 | 3 |
|------------|-----------------|--------------------|--------------------|-----------------|
| $P(S_3=k)$ | $\frac{p^2}{2}$ | $\frac{q(1+p)}{2}$ | $\frac{q(1+p)}{2}$ | $\frac{p^2}{2}$ |

$$P_0^4 = Q_4^4 = \frac{p^3}{2}, \quad P_1^4 = Q_3^4 = \frac{pq}{2}(1+q),$$

$$P_2^4 = Q_2^4 = \frac{q(p^2+q)}{2},$$

$$P_3^4 = Q_1^4 = \frac{qp^2}{2}, \text{ et la distribution de } S_4 :$$

| k | 0 | 1 | 2 | 3 | 4 |
|------------|-----------------|------|-----------|------|-----------------|
| $P(S_4=k)$ | $\frac{p^3}{2}$ | qp | $q(1-pq)$ | pq | $\frac{p^3}{2}$ |

On vérifie que, pour $p=q=\frac{1}{2}$, S_1, S_2, S_3 ont des distributions binomiales ; par contre, pour $p > \frac{1}{2}$, non seulement ces distributions ne sont pas binomiales mais on peut même, en choisissant p suffisamment voisin de 1, leur donner une forme "en U" et non en cloche (bien entendu, pour $p=1$, $P(S_n=0) = P(S_n=n) = \frac{1}{2}$).

Il est facile de voir que pour tout n

$$P(S_n=0) = P(S_n=n) = \frac{p^{n-1}}{2}$$

On peut donc utiliser le nombre de billes tombées dans les cases extrêmes pour estimer le paramètre p (on remarque que $\frac{p^n}{2} > \frac{1}{2^{n+1}}$: il y a plus de billes dans les cases extrêmes pour une planchette à clous que pour une planchette de Galton).

Les calculs ci-dessus sont à la portée d'une classe terminale. Décrivons rapidement ceux qu'on peut poursuivre avec des élèves de D.E.U.G. ou des candidats au C.A.P.E.S.

Pour déterminer la fonction caractéristique de S_n , définie pour t réel par

$$\chi_n(t) = E(e^{itS_n}) = \sum_{k=0}^n P(S_n=k) e^{itk}$$

posons $\sum_k P_k^n e^{itk} = \varphi^n(t)$

et $\sum_k Q_k^n e^{itk} = \psi^n(t)$

On déduit de (1) et (2)

$$(1') \quad \varphi^1 = \frac{1}{2}, \quad \psi^1 = \frac{e^{it}}{2}$$

$$\text{et } (2') \quad \varphi^{n+1} = p \varphi^n + q \psi^n \\ \psi^{n+1} = e^{it}(q \varphi^n + p \psi^n)$$

Posons $A = \begin{pmatrix} p & q \\ e^{it} q & e^{it} p \end{pmatrix}$

On a donc $\begin{pmatrix} \varphi^{n+1} \\ \psi^{n+1} \end{pmatrix} = A \begin{pmatrix} \varphi^n \\ \psi^n \end{pmatrix}$

$$\text{et } \begin{pmatrix} \varphi^1 \\ \psi^1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ e^{it} \end{pmatrix}$$

$$\text{d'où } \begin{pmatrix} \varphi^n \\ \psi^n \end{pmatrix} = \frac{1}{2} A^{n-1} \begin{pmatrix} 1 \\ e^{it} \end{pmatrix}$$

$$\text{et } \chi_n = \varphi^n + \psi^n = (1,1) \begin{pmatrix} \varphi^n \\ \psi^n \end{pmatrix} \\ = \frac{1}{2} (1,1) A^{n-1} \begin{pmatrix} 1 \\ e^{it} \end{pmatrix}$$

Nous avons déjà remarqué, au moins sur les premières valeurs de n , que S_n a une distribution symétrique, c'est-à-dire que

$$P(S_n - \frac{n}{2} = k) = P(S_n - \frac{n}{2} = -k)$$

Ceci nous amène à introduire la fonction caractéristique de la variable centrée $S_n - \frac{n}{2}$:

$$\begin{aligned}\xi_n(t) &= E(e^{it(S_n - \frac{n}{2})}) = E(e^{itS_n} e^{-\frac{itn}{2}}) = \chi_n(t) e^{-\frac{itn}{2}} \\ &= \frac{1}{2} (1, 1) \left(e^{-\frac{it}{2}} A \right)^{n-1} \begin{pmatrix} e^{-\frac{it}{2}} \\ e^{\frac{it}{2}} \end{pmatrix}\end{aligned}$$

$$\text{Or } e^{-\frac{it}{2}} A = \begin{pmatrix} p e^{-\frac{it}{2}} & q e^{-\frac{it}{2}} \\ q e^{\frac{it}{2}} & p e^{\frac{it}{2}} \end{pmatrix}$$

a pour valeurs propres les deux racines de l'équation

$$(\lambda - p e^{-\frac{it}{2}})(\lambda - p e^{\frac{it}{2}}) - q^2 = 0$$

ou
$$\lambda^2 - 2p\lambda \cos \frac{t}{2} + p^2 - q^2 = 0.$$

Comme

$$p^2 \cos^2 \frac{t}{2} + q^2 - p^2 = q^2 - p^2 \sin^2 \frac{t}{2} > 0 \text{ pour } |t| \text{ assez petit}$$

(si on exclut le cas $p=1, q=0$ déjà traité), $e^{-\frac{it}{2}} \psi A$ a deux valeurs propres réelles distinctes, $\lambda_1(t), \lambda_2(t)$ pour $|t|$ assez petit, données par

$$\lambda_1(t) = p \cos \frac{t}{2} + \sqrt{q^2 - p^2} \sin^2 \frac{t}{2}$$

$$\lambda_2(t) = p \cos \frac{t}{2} - \sqrt{q^2 - p^2} \sin^2 \frac{t}{2}$$

et l'on peut écrire

$$\xi_n(t) = \lambda_1^n(t) a(t) + \lambda_2^n(t) b(t)$$

où a et b sont deux fonctions qui ne dépendent pas de n ; en particulier

pour $t=0$, $\lambda_1(0) = p + q = 1$ et $\lambda_2(0) = p - q$

d'où $1 = \xi_n(0) = a(0) + (p-q)^n b(0)$, pour tout n , d'où $a(0) = 1$ et $b(0) = 0$.

Intéressons-nous maintenant, comme il est usuel dans l'établissement du théorème de la limite centrale, à la variable

$$\frac{1}{\sqrt{n}} \left(S_n - \frac{n}{2} \right)$$

Elle a pour fonction caractéristique

$$\xi_n\left(\frac{t}{\sqrt{n}}\right) = \lambda_1^n\left(\frac{t}{\sqrt{n}}\right) a\left(\frac{t}{\sqrt{n}}\right) + \lambda_2^n\left(\frac{t}{\sqrt{n}}\right) b\left(\frac{t}{\sqrt{n}}\right)$$

Quand $t \rightarrow 0$, $\lambda_1(t) \rightarrow p+q=1$ et $\lambda_2(t) \rightarrow p-q < 1$ on en déduit que $\lambda_2^n\left(\frac{t}{\sqrt{n}}\right) \rightarrow 0$ et, comme b est continue, que

$$\lambda_2^n\left(\frac{t}{\sqrt{n}}\right) b\left(\frac{t}{\sqrt{n}}\right) \rightarrow 0.$$

Pour étudier la limite de $\lambda_1^n\left(\frac{t}{\sqrt{n}}\right)$, développons $\lambda_1(t)$ à l'ordre deux :

$$\begin{aligned} \lambda_1(t) &= p\left(1 - \frac{t^2}{8}\right) + q\left(1 - \frac{p^2}{q^2} - \frac{t^2}{8}\right) + o(t^2) & t \rightarrow 0 \\ &= 1 - \frac{t^2}{8} \frac{p}{q} + o(t^2) \end{aligned}$$

on en déduit que $\lambda_1^n\left(\frac{t}{\sqrt{n}}\right) \rightarrow e^{-\frac{t^2 p}{8q}}$

et, comme $\xi_n\left(\frac{t}{\sqrt{n}}\right)$ est une fonction caractéristique, et que a est continue, avec $a(0) = 1$,

$$\lim_{n \rightarrow \infty} \xi_n\left(\frac{t}{\sqrt{n}}\right) = e^{-\frac{t^2 p}{8q}},$$

fonction caractéristique d'une variable gaussienne centrée, de variance $p/4q$.

La répartition de $\frac{1}{\sqrt{n}}(S_n - \frac{n}{2})$ converge donc vers une répartition gaussienne, centrée, de variance $\frac{p}{4q}$ (dans le cas $p=q=\frac{1}{2}$, on retrouve le théorème de la limite centrale puisqu'alors $\frac{1}{\sqrt{n}}(S_n - \frac{n}{2})$ a pour variance $\frac{1}{4}$).

On voit que, par un choix convenable de p voisin de 1, $\frac{p}{q} = \frac{p}{1-p}$ peut être rendu arbitrairement grand.

Ainsi, même si pour n assez grand notre planchette à clous redonne une courbe en cloche, celle-ci est, pour p assez grand, beaucoup plus aplatie que la courbe en cloche de la planchette de Galton.