Paradoxes probabilistes :
Les lois de la Nature sont-elles compatibles avec la théorie des probabilités ?

par Bernard Beauzamy

Newsletter "les mathématiques du réel", article no 3, mars 2024

I. Introduction

La Théorie des Probabilités a environ 350 ans : on date les débuts formels de la correspondance entre Pierre de Fermat et Blaise Pascal en 1654. Malgré cette ancienneté, elle a du mal à pénétrer convenablement les communautés scientifiques, les entreprises, les journalistes et le grand public. Les idées fausses sont légion et elles ne tendent nullement à disparaître, bien au contraire. Les temps d'obscurantisme où nous vivons actuellement font que chaque discipline, chaque corporation, s'arcboute sur ses propres dogmes et refuse absolument toute remise en cause.

Parmi les idées fausses les plus souvent rencontrées, citons :

- Pour qu'une expérience ait une valeur scientifique, il faut qu'elle soit répétable. Ceci est absurde, comme nous le verrons sur des exemples simples.

- Si on prend un panel, mettons de 100 000 personnes, à égalité H/F, un sous-panel, mettons de 1 000 personnes, choisi au hasard, aura aussi une égalité H/F. Ceci est radicalement faux. La différence $|H-F|$, pour un sous-panel de taille $n,$ est de l'ordre de $\sqrt{n}.$

- Si je joue à Pile ou Face avec un adversaire, et que je reçois un € à chaque fois que je gagne, que je débourse 1 € à chaque fois que je perds, ma fortune tendra vers 0 à mesure que le nombre de parties augmente. Ceci est radicalement faux. Bien au contraire, il viendra nécessairement un moment où j'aurai gagné un milliard d'€ et un moment où j'aurai perdu cette même somme (peut-être avant, peut-être après). La courbe de sécurité est la "courbe de Khintchine" ; cela s'appelle "loi du logarithme itéré". Il est remarquable que, pour décrire un phénomène aussi simple que le jeu de pile ou face, on doive avoir recours à des mathématiques aussi difficiles et que le résultat soit aussi complexe.

- Si des animaux se partagent un territoire, ou des partis politiques une population, il s'établira progressivement un équilibre. Ceci est radicalement faux. Bien au contraire, la Nature ne recherche pas un équilibre, mais procède par larges oscillations, dont les amplitudes vont croissant. C'est dommage pour les économistes, dont les modèles d'équilibre sont entièrement dépourvus de valeur scientifique.

- Si on demande à un industriel de faire la preuve que, sur 100 appareils, au plus deux ont des problèmes, il a tout intérêt à attendre d'avoir fait tous les tests avant de rendre ses résultats aux Autorités. Le protocole "séquentiel" (on teste 100 appareils et on arrête dès que deux sont défectueux) et le protocole "global" (on va jusqu'au bout des 100 tests) conduisent à des résultats très différents.

Nous avons ici une nouvelle illustration de ce que nous disions dans les deux premiers numéros de notre "newsletter" : les mathématiques ont des difficultés à décrire les phénomènes naturels, même les plus simples.

Bien entendu, les énoncés mentionnés ci-dessus sont connus de très longue date. Prenons les choses dans l'ordre.

II. Répétabilité d'une expérience

Si je donne un grand coup de massue sur la tête de quelqu'un, il meurt ; expérience parfaitement répétable et qui donnera le même résultat à chaque fois, ce qui est très satisfaisant. Pourtant, considérons l'expérience suivante : je jette une pièce selon le jeu de pile ou face, trois fois de suite. L'expérience est parfaitement répétable et pourtant elle ne donnera pas le même résultat à chaque fois.

L'exigence de répétabilité, exigée par de nombreuses corporations (notamment les épidémiologistes et les spécialistes des vaccins) est totalement absurde dans son principe même. Chaque expérience comporte une part d'aléatoire, qui peut être due à l'imprécision des données d'entrée, ou bien à la nature même de l'expérience.

On peut tout au plus espérer un énoncé de nature probabiliste, du type : si les paramètres A, B, C sont dans telle plage, alors tel effet se manifestera dans 80% des cas. Vouloir un énoncé déterministe relève de l'arrogance et de l'ignorance des lois de la nature.

Les probabilités sont destinées à rendre compte de la variabilité des phénomènes ; à ce jour, c'est le meilleur moyen et en vérité le seul dont nous disposions ! Ces variabilités sont de plusieurs natures :

- Les conditions où se déroulent les process sont variables : c'est le cas de la température ambiante, de la pression atmosphérique, etc.

- Il arrive de temps en temps des phénomènes non souhaités, mais à prendre en compte, comme des séismes, des tornades, etc.

- Le réglage des process lui-même n'est jamais parfaitement constant : ce peut être la qualité des composants chimiques, les températures de fabrication, etc.

En résumé, toute activité humaine, tout process de fabrication industrielle, présentent une certaine variabilité, qu'on le veuille ou non.

L'attitude prise par la plupart des ingénieurs consiste à vouloir réduire cette variabilité autant que possible. On y parvient dans certains cas : on peut chercher à mieux contrôler un process de fabrication. Mais on ne parviendra jamais à l'éliminer entièrement.

Pour décrire un phénomène complexe, les ingénieurs aiment bien se doter de codes de calcul, aussi précis que possible. Et c'est là que la démarche devient fondamentalement "anti-probabiliste", dans son essence même. Car la prise en compte des incertitudes dans un code de calcul est impossible : il retourne un résultat précis à partir de valeurs d'entrée précises. La réponse habituellement apportée par les ingénieurs est "nous faisons des milliers de runs du code". Mais si le code comporte 40 paramètres, chacun avec 10 valeurs possibles, il faudra 10^{40} runs, ce qui est impossible en pratique, quelle que soit la vitesse d'exécution du code. L'idée selon laquelle on pourrait faire ces runs au hasard et les traiter selon la "méthode de Wilks" est fondamentalement erronée, comme nous l'avons expliqué dans l'article :
https://www.scmsa.eu/archives/BB_Wilks_2016_01_11.pdf

III. Déterminisme dans la Nature

Les lois de la physique sont purement axiomatiques : on nous dit : voici la forme de la loi de la gravitation universelle, voici la forme des lois de l'électromagnétisme, etc. Nous constatons cette forme mais nous ne savons pas l'expliquer ; elle ne résulte pas d'une démonstration.

Toutes les lois de la Nature que nous connaissons sont fondamentalement déterministes, sauf, à l'échelle macroscopique, celle qui régit le jeu de Pile ou Face (voir plus bas) et, à l'échelle microscopique, celles qui régissent la mécanique quantique.

La croyance au déterminisme dans la Nature remonte à Laplace, qui assurait que si tous les éléments relatifs à tous les atomes étaient connus à un instant, on pourrait en déduire toutes les positions ultérieures. Les conceptions modernes ont remis ceci en cause. Un exemple clair est le principe d'incertitude de Heisenberg, qui contredit complètement la conception de Laplace.

1. Mécanique quantique

Concernant la mécanique quantique, il est très difficile d'en parler, parce que l'on n'y comprend rien. Les points suivants sont à considérer :

- Le formalisme hilbertien, dû à Von Neumann, très pertinent pour des investigations préliminaires, n'est sans doute pas approprié. Il est naïf de croire que la Nature se laisse décrire, lorsqu'une observation est réalisée, par les valeurs propres d'un opérateur sur un espace de Hilbert.

- Il n'est pas du tout évident que la Théorie des Probabilités, telle que nous la connaissons, soit appropriée à l'échelle de la mécanique quantique.

Quoi qu'il en soit, on observe un grand nombre d'erreurs, chez les spécialistes de mécanique quantique, dues à leur ignorance des probabilités. Voir en particulier nos articles :

- Quantum electrodynamics: Some doubts about Feynman's probabilistic presentation
https://scmsa.eu/archives/BB_2015_Quantum_electrodynamics_doubts.pdf

- Analyse critique du livre "L'impensable hasard", par Nicolas Gisin
https://www.scmsa.eu/archives/BB_Correlations_Quantiques_2019_05.pdf

- Mécanique quantique, variables cachées et inégalités de Bell
https://www.scmsa.eu/archives/SCM_Variables_cachees_inegalites_Bell.pdf
dont nous extrayons le commentaire suivant :

Cet article se consacre à l'analyse de l'intrication ; il est divisé en deux parties. La première décrit une analogie physique simple : un canon à clous. Malheureusement, l'auteur prouve alors qu'il ignore les lois fondamentales des probabilités. Dans la seconde, l'auteur applique le raisonnement à l'intrication de deux photons et, à ce moment, il prouve qu'il ignore également les lois fondamentales de la mécanique quantique. Même en physique atomique, deux ignorances ne font pas une science.

Il y a, aux fondements même de la mécanique quantique, une confusion qu'il faut relever. On nous dit : la position de l'électron est régie par une densité de probabilité. Soit. On nous dit encore : pour que des franges d'interférence puissent être réalisées, pour un seul électron, il faut que l'électron passe dans deux fentes à la fois (dispositif des trous d'Young). Soit encore, mais les deux sont contradictoires.

On peut parler de la loi de probabilité régissant la distribution des tailles des habitants dans la région parisienne : tel pourcentage entre 1.70 m et 1.75 m. Une telle loi a un sens parfaitement clair et elle peut être construite numériquement, à partir d'observations. Mais il n'empêche que chacun de nous n'a qu'une seule taille !

Beaucoup d'articles relatifs à la mécanique quantique font une confusion entre "probabilité de présence en un point" et "présence simultanée en plusieurs endroits" ; ce n'est pas la même chose.

2. Composition d'un sous-panel

Supposons un panel constitué à égalité H/F ; on en extrait un sous-panel de taille n. Quelles peuvent être les compositions du sous-panel ? La réponse est de nature combinatoire et elle est donnée dans les manuels de niveau Terminale.

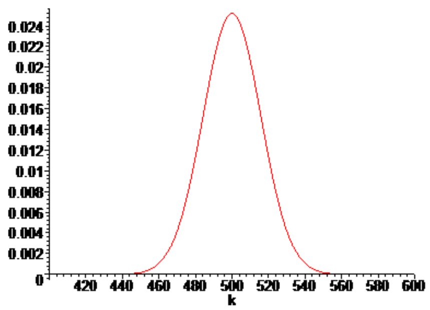
Représentons notre problème par la donnée de deux boîtes, l'une, celle de gauche, contient les H et l'autre, celle de droite, contient les F.

Pour extraire un sous-panel de taille n qui ne contient que des H, il faut se diriger n fois vers la gauche ; la probabilité est $\frac{1}{2^n}$.

Pour extraire un sous-panel de taille n qui contient $n-1$ H et 1 F il faut se diriger $n-1$ fois vers la gauche et 1 fois vers la droite ; la probabilité est $\binom{n}{1}\frac{1}{2^n}$.

Pour extraire un sous-panel de taille n qui contient $n-k$ H et k F il faut se diriger $n-k$ fois vers la gauche et k fois vers la droite ; la probabilité est $p_k = \binom{n}{k}\frac{1}{2^n}$.

On aura égalité si $k = \frac{n}{2}$; la probabilité est $\binom{n}{n/2}\frac{1}{2^n} \approx \sqrt{\frac{2}{\pi n}}$; la probabilité d'avoir égalité stricte est donc très faible. Voici un exemple de répartition des nombres p_k, pour $n = 1000$:



Ces résultats intéressent les industriels, mais dans le sens inverse : ils disposent d'un sous-panel, par exemple 100 radiateurs ont été testés lors des essais, et, parmi eux, 2 sont tombés en panne ; ils voudraient connaître le nombre total de radiateurs défectueux, attendu sur un total de production de 5000. Les formules explicites sont dues à Laplace ; on les trouvera dans le livre [NMP]. Nous y revenons plus bas.

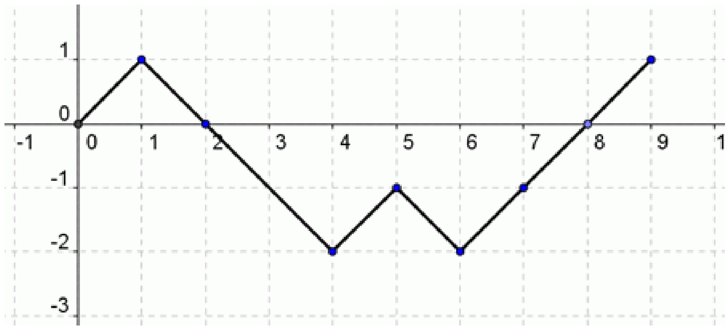
IV. Le jeu de Pile ou Face

C'est l'exemple le plus simple, le plus frappant, de loi de la Nature qui se laisse traiter complètement de manière axiomatique : on dispose d'une variable aléatoire X dont la loi est :

$$P\{X = 1\} = P\{X = -1\} = \frac{1}{2},$$

et dont les tirages successifs, notés X_k, sont indépendants.

Pratiquement, on peut considérer cela comme un jeu : le joueur qui perd donne un Euro à celui qui gagne ; la question est : comment évolue, au cours du temps, la fortune de chaque joueur ? Mathématiquement, on note $S_n = \sum_{k=1}^{n} X_k$: c'est le gain (algébrique) de chacun des joueurs : il peut être positif ou négatif. Le graphe de S_n (temps en abscisse, valeur du gain en ordonnée) représente une marche aléatoire dans le plan. Voici un exemple sur 9 coups :



La marche aléatoire part de l'origine. On s'intéresse à son comportement asymptotique lorsque le nombre de parties augmente. Les énoncés quantitatifs sont de deux types :

– Les sommes S_n repassent souvent par la valeur 0 : il y a souvent égalité entre les joueurs. C'est ce que nous appellerons le rapprochement ;

– Les oscillations des S_n sont de plus en plus accentuées ; contrairement à ce que l'on croit souvent, il est faux que $S_n \to 0$ lorsque $n \to +\infty$. C'est ce que nous appellerons l'éloignement.

3. Rapprochement des gains

L'énoncé est le suivant :

Théorème. - L*a probabilité que la première égalité se produise au-delà de* $2n$ *vaut :*

$$g_n = \frac{1}{4^n}\binom{2n}{n} \approx \frac{1}{\sqrt{\pi n}}$$

Elle décroît donc lentement à mesure que n augmente. Un exemple est donné dans le tableau suivant :

nb de coups	10	20	50	100	200
proba pas d'égalité	0,18	0,13	0,08	0,06	0,04

Ce tableau donne la probabilité que l'on n'ait pas d'égalité de gains pendant la partie, en fonction du nombre de coups joué. Par exemple, si l'on joue 200 fois, il n'y a qu'une probabilité de 0.04 que l'égalité ne se produise jamais pendant ces 200 coups. Autrement dit, avec proba 96%, l'égalité des gains se produit au moins une fois avant le 200$^{\text{ème}}$ coup.

Pour la démonstration, voir par exemple J.C. Kalbfleisch "Probability and Statistical Inference", volume 1.

4. Eloignement

Les résultats d'éloignement sont très peu intuitifs. Ils résultent de la loi du logarithme itéré, due au mathématicien russe Alexander Khintchine (1924).

Considérons la courbe $y=\sqrt{2xLog\left(Log\left(x\right)\right)}$, appelée "courbe de Khintchine". C'est une "courbe de sécurité" : presque sûrement, les gains ou les pertes de chaque joueur, à l'instant $n,$ ne dépasseront pas (en valeur absolue) la quantité $\sqrt{2nLog\left(Log\left(n\right)\right)}$. Nous allons donner un énoncé quantitatif précis.

Mais, chose plus remarquable, il est presque sûr que cette courbe sera atteinte. Si par exemple nous prenons $\varepsilon=0.01,$ il viendra presque sûrement un moment n où le gain du joueur 1 dépassera $0.99\sqrt{2nLog\left(Log\left(n\right)\right)}$, et de même pour ses pertes, bien sûr.

Théorème. - Loi du Logarithme Itéré, version quantitative, 1

Soit $\varepsilon>0$ *et* $m\geq 2$. *Posons :*

$$\eta\left(m,\varepsilon\right)=\frac{2}{\left(Log\left(1+\varepsilon\right)\right)^{1+\varepsilon}}\frac{1}{\varepsilon\left(m-1\right)^{\varepsilon}}$$

et :

$$B\left(m,\varepsilon\right)=\left\{\exists n\geq\left(1+\varepsilon\right)^{m},S_{n}>\left(1+\varepsilon\right)\varphi\left(n\right)\right\}.$$

Alors, pour tout m *:*

$$P\left(B\left(m,\varepsilon\right)\right)\leq\eta\left(m,\varepsilon\right).$$

L'ensemble $B\left(m,\varepsilon\right)$ est fait des chemins qui sont au dessus de la courbe $\left(1+\varepsilon\right)\varphi$ au moins une fois après le temps $\left(1+\varepsilon\right)^{m}$. Le théorème dit que la probabilité d'un tel événement tend vers 0 lorsque m augmente.

Prenons par exemple $\varepsilon=1$. Le théorème donne l'estimation :

$$P\left\{\exists n\geq 2^{m},S_{n}>2\varphi\left(n\right)\right\}\leq\frac{2}{\left(Log\left(2\right)\right)^{2}}\frac{1}{m-1}. \quad (1)$$

Si on veut que le membre de droite soit $\leq 0.05,$ il faut $m=85$. La probabilité d'avoir un $n\geq 2^{85}$ pour lequel $S_{n}>2\varphi\left(n\right)$ est < 0.05. Au-delà de $2^{85},$ 95% des chemins vérifient $S_{n}\leq 2\varphi\left(n\right)$.

A l'inverse, la plupart des chemins dépassent la courbe de sécurité. Plus précisément, soit $\alpha>0$ (petit). Nous allons montrer qu'il existe un $N_{0}(\varepsilon,\alpha)$ tel que si $N>N_{0}\left(\varepsilon,\alpha\right)$, alors :

$$P\left(\forall n=1,\ldots N\,;\,\frac{S_n}{\varphi(n)}<1-\varepsilon\right)<\alpha.$$

Précisément :

Théorème. - Loi du Logarithme Itéré, version quantitative, 2

Soit ε, $0<\varepsilon<1$, $0<\alpha<1$. *Posons* $\gamma=\frac{16}{\varepsilon^2}$, $\delta=\frac{2}{(Log(2))^2}\approx 4.2$, $k_0=1+\left[\left(2+\frac{\delta}{\alpha}\right)\frac{Log(2)}{Log(\gamma)}\right]$ *et*

$N_0=\left(k_0^{\frac{\varepsilon}{2}}+\frac{\varepsilon}{2}Log\left(\frac{2}{\alpha}\right)\right)^{\frac{2}{\varepsilon}}$. *On a :*

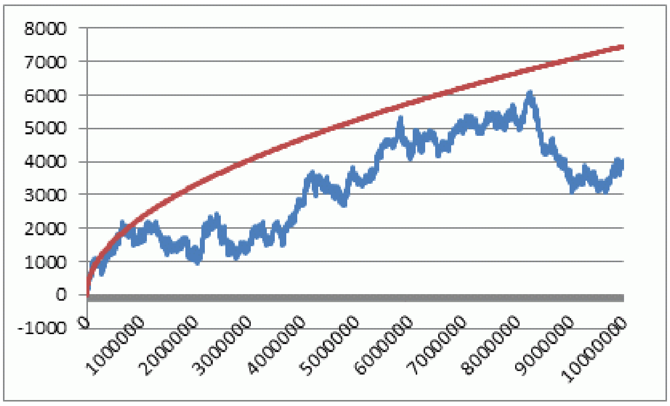
$$P\left(\exists k,\, k_0\le k\le N_0,\ tel\ que\ \frac{S_k}{\varphi(k)}>1-\varepsilon\right)\ge 1-\alpha.$$

Prenons par exemple $\varepsilon=\frac{1}{2}$, $\alpha=\frac{1}{2}$. On trouve $k_0=2$ et $N_0=7$ et donc :

$P\left(\exists k,\, 2\le k\le 7,\ tel\ que\ \frac{S_k}{\varphi(k)}>\frac{1}{2}\right)\ge\frac{1}{2}.$

Ce que dit ce théorème c'est qu'il est de moins en moins probable de rester constamment en dessous de la courbe $(1-\varepsilon)\varphi(x)$. Pour une largeur ε donnée, la probabilité que S_n entre dans la bande $[(1-\varepsilon)\varphi(x),\varphi(x)]$ tend vers 1 lorsque $n\to+\infty$.

Voici un exemple de marche aléatoire, avec la courbe de sécurité :



Il est très remarquable que, pour décrire un phénomène aussi simple que le jeu de pile ou face, on doive faire appel à des mathématiques aussi élaborées et que le résultat soit aussi complexe.

Les théorèmes probabilistes sont dus à Khinchine ; les énoncés quantitatifs se trouvent dans le livre [SRW].

5. Erreur commise dans les raisonnements

D'où vient l'erreur commise habituellement, qui fait penser que $S_n \to 0$?

On voit souvent écrire que $\frac{S_n}{n} = \frac{1}{n}\sum_1^n X_k$ suit une loi de Gauss (ce qui n'est pas vrai pour chaque n ; ce n'est vrai qu'asymptotiquement), et que par conséquent :

$$P\left\{-a \le \frac{S_n}{n} \le a\right\} = \int_{-a}^{a} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} \qquad (1)$$

ce qui est faux, ou encore :

$$P\left\{-an \le S_n \le an\right\} = \int_{-a}^{a} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}}$$

Posant $b = an$:

$$P\left\{-b \le S_n \le b\right\} = \int_{-\frac{b}{n}}^{\frac{b}{n}} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} \to 0 \text{ lorsque } n \to \infty$$

L'égalité (1) est fausse. Ce qui est correct est le fait que, pour tout a :

$$\left| P\left\{-a \le \frac{S_n}{n} \le a\right\} - \int_{-a}^{a} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} \right| \to 0,\ n \to \infty$$

mais cette convergence dépend de a.

6. Bien comprendre les énoncés

Les énoncés probabilistes sont toujours très difficiles à comprendre ; comme nous le disons dans le livre [MPPR], le vocabulaire généralement utilisé est peu approprié.

Les théorèmes ci-dessus ne sont pas des résultats qui guident un jeu de pile ou face en particulier. Si vous jouez avec votre voisin, il n'y a pas d'autre règle que ± 1 à chaque coup, avec probabilité 1/2. Sur une partie de N coups, tous les chemins ont probabilité $\frac{1}{2^N}$, y compris celui qui monte constamment et celui qui descend constamment, qui ne sont en rien moins probables que les autres. Il n'existe aucune loi occulte qui dit que, tôt ou tard, Dieu va guider votre gain vers la courbe de sécurité décrite plus haut.

Les théorèmes ci-dessus sont en réalité des théorèmes de dénombrement de chemins. Si vous avez joué pendant N coups, la plupart des 2^N chemins vont venir tangenter la courbe de sécurité au sens mentionné plus haut. Ou bien, en d'autres termes, si d'innombrables joueurs jouent en même temps à pile ou face, pour la plupart d'entre eux on observera le phénomène.

V. L'utilisation du hasard

Les Industriels pensent souvent qu'il est bon de restreindre le rôle du hasard. Cette idée n'est pas nécessairement correcte ; pis, elle est en général fausse.

Rappelons (voir le livre [MPPR]) que l'on désigne sous le nom générique de "hasard" toutes les causes, toutes les influences, que l'on ne peut pas connaître ou que l'on ne souhaite pas connaître.

Par exemple, un marchand de chaussures a besoin d'un stock ; pour prévoir le nombre de paires de telle pointure, il considérera que l'entrée d'un client et son choix résultent du hasard. Il n'est pas possible de faire un recensement des pointures aux alentours, sans oublier les touristes.

On peut considérer que le débit d'une rivière résulte du hasard et constituer des lois de probabilité de ce débit en se servant d'un historique. On peut aussi essayer de remonter aux causes, au moyen de modèles "pluies-débit", qui vont associer le débit de la rivière aux pluies pendant les jours précédents, mais de tels modèles sont très difficiles à construire (beaucoup d'autres paramètres interviennent, comme l'état des sols) et encore plus difficiles à valider. Dans le cas présent, l'approche "hasard" est beaucoup plus pertinente.

Un industriel qui fabrique des radiateurs considérera valablement que le dysfonctionnement d'un appareil résulte du hasard et il en recherchera la loi de probabilité : quel pourcentage est affecté ? Eventuellement, il peut vouloir des informations complémentaires : selon la saison, selon la région ; ce seront des lois de probabilité conditionnelle. Mais il s'en remet au hasard, en ce sens qu'il ne recherche pas l'explication physique du phénomène qui, dans certains cas, peut être très difficile à trouver.

S'en remettre au hasard signifie donc que l'on renonce à chercher les causes intimes d'un phénomène ; par contre, on va réclamer un historique, de manière à en apprécier les occurrences.

VI. Surveillance d'un échantillon

Un industriel veut mettre un produit nouveau sur le marché (par exemple des radiateurs électriques) ; il se donne un "taux de rejet" maximum acceptable, par exemple 2/100. Cela signifie, dans son idée, que s'il vend 10 000 radiateurs, il aura des retours (appels à la garantie) pour environ 200 d'entre eux. L'industriel veut tester un échantillon de taille 100 et, en fonction du nombre d'appareils défectueux dans l'échantillon, prendre une décision : mise sur le marché ou non. Si, sur 100 appareils testés, au plus deux sont défectueux, ok : mise sur le marché. Il estime que, pour chaque appareil, la probabilité d'être défectueux est $p = 1/100$. Il est confiant dans le résultat du test.

Deux approches sont possibles :

1. **Approche globale**

Il teste tout l'échantillon et calcule à la fin le quotient $\frac{D}{N}$, où D est le nombre d'appareils défectueux et N est la taille de l'échantillon. Si $\frac{D}{N} \leq \frac{2}{100}$, le produit est considéré comme satisfaisant, et la mise sur le marché est décidée.

2. **Approche séquentielle**

Les appareils sont rangés dans un ordre quelconque ; à chaque fois qu'un nouvel appareil est testé, on regarde le quotient $\frac{d}{n}$ (d : nombre d'appareils défectueux, n : nombre d'appareils testés) ; si à un moment quelconque ce quotient est $> \frac{2}{100}$ le produit est considéré comme non-satisfaisant. Les mêmes approches sont possibles pour le test d'un médicament, d'un vaccin, etc., sur une population.

La différence entre les deux approches est évidente. Supposons que les essais révèlent que le numéro 54 et le numéro 87 sont défectueux, et ceux-là seulement. Alors le test global est satisfaisant : le ratio $\frac{D}{N} = \frac{2}{100}$. Mais le test séquentiel s'arrête à l'instant 87, parce que $\frac{2}{87} > \frac{2}{100}$.

Lien avec la marche aléatoire dans le plan.

Notons X_k la variable aléatoire qui vaut 1 si le $k^{ème}$ appareil est défectueux (probabilité p) et 0 sinon, on a $S_n = \sum_{k=1}^{n} X_k$; la condition de succès du test séquentiel se traduit donc par le fait que la marche aléatoire S_n reste constamment au-dessous de la droite $y = \frac{2x}{100}$. La condition de succès du test global se traduit par le fait que, au terme des tests, la marche aléatoire se situe dans l'intervalle $[-2, 2]$.

Pour le test séquentiel, ce sont les premiers essais qui sont les plus dangereux. Un seul défaut parmi les 49 premiers conduit au rejet.

Références

J.C. Kalbfleisch "Probability and Statistical Inference", volume 1. Springer Texts in Statistics, 1993.

[NMP] Bernard Beauzamy : Nouvelles Méthodes Probabilistes pour l'évaluation des risques.

SCM SA, ISBN 978-2-9521458-4-8. ISSN 1767-1175, broché, 272 pages. Avril 2010.

[MPPR] Bernard Beauzamy : Méthodes Probabilistes pour l'étude des phénomènes réels. SCM SA, ISBN 2-9521458-0-6. ISSN 1767-1175, broché, 369 pages. Mars 2004. Seconde Edition, juin 2016.

[SRW] Bernard Beauzamy: Simple Random Walks in the Plane, in English, SCM SA. ISBN : 979-10-95773-01-6, ISSN: 1767-1175. Relié, 208 pages. Février 2020.