V. RÉUNIONS D'ÉVÉNEMENTS.

V. 1. Qu'est-ce qu'un événement ?

Les problèmes scolaires de probabilités sont le plus souvent exprimés en langage courant. Par exemple : *on jette trois dés ; calculer la probabilité d'avoir trois six* ; ou bien : *on jette trois dés ; calculer la probabilité d'avoir un double (c'est-à-dire que deux des trois dés montrent la même face)* ; ou encore *on jette trois dés ; calculer la probabilité pour que chacun des trois dés marque un nombre impair* ; etc. Dans ces problèmes l'énoncé détermine toujours un événement dont il s'agira de calculer la probabilité. Ainsi, dans le premier des trois exemples précédents, l'événement est "les trois dés marquent six". Pour résoudre le problème, il faut traduire ce langage imagé sous forme mathématique, c'est-à-dire trouver d'abord l'espace Ω qui convient le mieux, puis déterminer le sous-ensemble A de Ω qui représente l'événement. Ainsi, la première proposition de l'énoncé *on jette trois dés* se traduit mathématiquement en prenant pour espace Ω l'ensemble des groupes de trois chiffres de 1 à 6, autrement dit l'ensemble des mots de trois lettres qu'on peut écrire avec l'alphabet $\{1, 2, 3, 4, 5, 6\}$, soit un ensemble de cardinal 216 (Ω est toujours l'ensemble de tous les résultats équiprobables possibles). La seconde proposition de l'énoncé détermine l'événement : chaque résultat *possible* est la famille des trois chiffres donnés par chaque dé, mais l'événement A cherché est l'ensemble des familles distinguées par l'énoncé : ici celles dont les trois chiffres sont six ; il n'y en a qu'une sur les 216, donc A est un sous-ensemble à un seul élément, de cardinal 1, donc de probabilité 1/216. Dans le deuxième problème, l'événement décrit par l'énoncé est "avoir un double". Le sous-ensemble B de Ω correspondant à cela est l'ensemble des familles ayant deux chiffres égaux, ou l'ensemble des mots ayant une lettre répétée ; le meilleur argument pour le dénombrer consiste à prendre son complémentaire B', qui est l'ensemble des mots dont les trois lettres sont toutes distinctes : c'est donc l'ensemble des mots ayant toutes leurs lettres distinctes (second cas de dénombrement du chapitre **II**), dont le cardinal est $6 \cdot 5 \cdot 4 = 120$. Donc on conclut que $\#B = 216 - 120 = 96$. Dans le troisième problème, l'événement suggéré par l'énoncé est "chacun des trois dés donne un chiffre impair". Le sous-ensemble C de Ω correspondant est l'ensemble des familles dont les trois chiffres sont tous impairs. On peut remarquer que $C = C_1 \cap C_2 \cap C_3$, où C_1 est l'ensemble des familles dont le *premier* chiffre est impair (et les autres indifférents), C_2

l'ensemble des familles dont le *second* chiffre est impair, et C_3 l'ensemble des familles dont le *troisième* chiffre est impair. On peut dire que C_1 est la traduction mathématique de "le premier dé donne un chiffre impair", et de même pour C_2 et C_3. L'intersection des trois ensembles correspond à la phrase "le premier dé donne un chiffre impair *et* le deuxième aussi *et* le troisième aussi". Il ne reste plus qu'à remarquer que les trois événements C_1, C_2, C_3 sont stochastiquement indépendants (reflet de l'indépendance causale des trois dés), de sorte que $\mathcal{P}(C) = \mathcal{P}(C_1) \times \mathcal{P}(C_2) \times \mathcal{P}(C_3)$. Les probabilités de C_1, C_2, C_3 sont faciles à calculer (elles valent $\frac{1}{2}$ car chacun des trois dés séparément a une chance sur deux de donner un chiffre impair), donc $\mathcal{P}(C) = \frac{1}{8}$.

On constate dans chacun de ces trois problèmes, que ce qui permet de *calculer* est la traduction mathématique préalable; c'est elle qui permet le dénombrement systématique. Mais elle a aussi permis le recours à des opérations ensemblistes (complémentaire pour le second problème et intersection pour le troisième) qui, grâce a leurs propriétés algébriques, ont facilité le calcul.

Considérons encore le quatrième problème que voici : *on jette trois dés; calculer la probabilité pour que l'un au moins des trois dés donne un chiffre impair*. Cette fois l'événement D décrit par l'énoncé est "le premier dé donne un chiffre impair *ou* le second dé donne un chiffre impair *ou* le troisième dé donne un chiffre impair" (*ou* non exclusif). Donc $D = C_1 \cup C_2 \cup C_3$. Pour pouvoir calculer comme dans le troisième problème, il faudrait disposer d'une formule donnant la probabilité d'une réunion d'événements [(1)]. Or une telle formule existe, c'est la formule de Poincaré.

V.2. Formule de Poincaré.

Commençons par le cas extrêmement simple de la réunion de deux événements. Il est clair que si deux ensembles A_1 et A_2 sont disjoints, le cardinal de $A_1 \cup A_2$ est la somme des cardinaux : $\#(A_1 \cup A_2) = \#A_1 + \#A_2$. Mais si les deux ensembles ne sont pas disjoints et qu'on écrit la liste numérotée des éléments de A_1, suivie immédiatement par la liste numérotée des éléments de A_2 (cette opération est appelée *concaténation*), on obtient une liste contenant $\#A_1 + \#A_2$ éléments, mais dans laquelle tous ceux qui sont à la fois dans A_1 et dans A_2, c'est-à-dire ceux qui sont dans $A_1 \cap A_2$, figurent deux fois. Pour compenser cette répétition, il faut donc retrancher le cardinal de $A_1 \cap A_2$; ainsi l'égalité suivante est correcte :

$$\#(A_1 \cup A_2) = \#A_1 + \#A_2 - \#(A_1 \cap A_2) \qquad (V.1.)$$

(1) On peut aussi considérer le complémentaire de D, qui est l'intersection des complémentaires des C_j, ce qui nous ramène au problème précédent. Mais nous y reviendrons.

Comment traiter le cas d'une réunion de trois ensembles $A_1 \cup A_2 \cup A_3$? On pourrait procéder de même, concaténer les trois listes d'éléments, puis regarder combien sont comptés plusieurs fois. On comprend aisément que ceux qui sont dans $A_1 \cap A_2 \cap A_3$ sont comptés trois fois, ceux qui sont dans $A_1 \cap A_2$, $A_2 \cap A_3$, ou $A_1 \cap A_3$ sans être dans $A_1 \cap A_2 \cap A_3$ sont comptés deux fois. Il faut donc retrancher $\#(A_1 \cap A_2)$, $\#(A_2 \cap A_3)$, et $\#(A_1 \cap A_3)$ à $\#A_1 + \#A_2 + \#A_3$, mais du coup on aura retranché trois fois ceux qui sont dans $A_1 \cap A_2 \cap A_3$ alors qu'il n'aurait fallu les retrancher que deux fois ; on compensera donc encore en les *rajoutant* une fois. Ainsi l'égalité suivante sera correcte :

$$\begin{aligned} \#(A_1 \cup A_2 \cup A_3) \quad &= \quad \#A_1 + \#A_2 + \#A_3 \\ &- \#(A_1 \cap A_2) - \#(A_2 \cap A_3) - \#(A_1 \cap A_3) \\ &+ \#(A_1 \cap A_2 \cap A_3) \end{aligned} \tag{V.2.}$$

En principe on pourrait traiter ainsi, par concaténations compensées, les réunions d'un nombre arbitraire d'événements ; mais l'examen de tous les cas possibles d'intersections partielles est un exercice un peu abstrait. Il vaut mieux se contenter d'abord de *deviner* une formule générale, qu'on démontrera ensuite rigoureusement par récurrence. La formule qu'on devine par les concaténations compensées est la suivante

$$\begin{aligned} \#(A_1 \cup A_2 \cdots \cup A_n) \quad &= \quad \sum_{j=1}^{j=n} \#A_j \\ &- \sum_{1 \le j_1 < j_2 \le n} \sum \#(A_{j_1} \cap A_{j_2}) \\ &+ \sum_{1 \le j_1 < j_2 < j_3 \le n} \sum \sum \#(A_{j_1} \cap A_{j_2} \cap A_{j_3}) \\ &- \sum_{1 \le j_1 < j_2 < j_3 < j_4 \le n} \sum \sum \sum \#(A_{j_1} \cap A_{j_2} \cap A_{j_3} \cap A_{j_4}) \\ &\cdots \\ &+ (-1)^{n-1} \#(A_1 \cap A_2 \cap A_3 \cdots \cap A_n) \end{aligned} \tag{V.3.}$$

Une manière plus condensée d'exprimer cette relation est la suivante :

$$\#(A_1 \cup A_2 \cdots \cup A_n) = \mathrm{S}_1 - \mathrm{S}_2 + \mathrm{S}_3 - \cdots \pm \mathrm{S}_n \tag{V.4.}$$

où

— S_1 est la somme des cardinaux de tous les A_j [cette somme contient n termes],

— S_2 est la somme des cardinaux de toutes les intersections de deux des A_j, telles que $A_1 \cap A_2$ ou $A_3 \cap A_7$ [contient $\binom{n}{2} = n(n-1)/2$ termes],

— S_3 est la somme des cardinaux de toutes les intersections de trois des A_j, telles que $A_1 \cap A_5 \cap A_6$ ou $A_3 \cap A_7 \cap A_{11}$ [contient $\binom{n}{3} = n(n-1)(n-2)/6$ termes],

— S_4 est la somme des cardinaux de toutes les intersections de quatre des A_j [contient $\binom{n}{4}$ termes],

...

— S_{n-1} est la somme des cardinaux de toutes les intersections de $n-1$ des A_j, [contient $\binom{n}{n-1} = n$ termes],

— S_n ne contient qu'un seul terme, $\#(A_1 \cap A_2 \cap A_3 \cap A_4 \cdots \cap A_n)$, car il n'y a qu'une seule intersection de n d'entre les A_j $[\binom{n}{n} = 1]$.

On peut maintenant se proposer de démontrer cette formule par récurrence: pour amorcer la récurrence, nous disposons déjà du cas $n = 2$ (de toute façon la formule n'a vraiment de sens que pour $n \geq 2$). Supposons qu'elle soit vraie pour $n-1$ événements et montrons qu'elle doit alors automatiquement être vraie aussi pour n événements $A_1, A_2, A_3, \ldots A_n$.

Or posons $B = A_2 \cup A_3 \cup A_4 \cdots \cup A_n$. On voit que la réunion des A_j de $j = 1$ à $j = n$, est identique à $A_1 \cup B$. Or il est déjà établi que $\#(A_1 \cup B) = \#A_1 + \#B - \#(A_1 \cap B)$. Mais d'autre part B est la réunion de $n-1$ événements, on peut donc utiliser l'hypothèse de récurrence pour décomposer $\#B$; de même $A_1 \cap B$ est égal à $(A_1 \cap A_2) \cup (A_1 \cap A_3) \cup (A_1 \cap A_4) \cdots \cup (A_1 \cap A_n)$, c'est-à-dire à une réunion de $n-1$ événements, auxquels aussi on peut donc appliquer l'hypothèse de récurrence. Ce qui donne

$$\#B = \mathrm{Q}_1 - \mathrm{Q}_2 + \mathrm{Q}_3 \cdots \pm \mathrm{Q}_{n-1}$$

où Q_j est la somme des cardinaux des intersections de j parmi les $n-1$ événements A_2, A_3, $A_4, \ldots A_n$. Il s'agit donc des intersections de j des n événements A_1, A_2, A_3, $A_4, \ldots A_n$, mais *excluant* l'événement A_1. De même (on reprend à partir d'ici la notation multiplicative pour l'intersection):

$$\#(A_1 B) = \mathrm{R}_1 - \mathrm{R}_2 + \mathrm{R}_3 \cdots \pm \mathrm{R}_{n-1}$$

où R_j est la somme des cardinaux des intersections de j parmi les $n-1$ événements A_1A_2, A_1A_3, A_1A_4, $\ldots A_1A_n$: ce sont donc les intersections de $j+1$ parmi les A_1, A_2, A_3, $A_4, \ldots A_n$, *dont l'un est obligatoirement* A_1. On peut donc dire que $\mathrm{R}_{j-1} + \mathrm{Q}_j$ est la somme des cardinaux de *toutes* les intersections de j parmi les n événements A_1, A_2, A_3, $A_4, \ldots A_n$: c'est la somme pour celles qui contiennent le facteur A_1 (R_{j-1}) plus la somme pour celles qui ne contiennent pas le facteur A_1 (Q_j). On a donc pour les expressions S définies plus haut

$$\mathrm{S}_j = \mathrm{R}_{j-1} + \mathrm{Q}_j$$

En appliquant alors l'hypothèse de récurrence comme indiqué, on obtient

$$\begin{aligned}
\#(A_1 \cup A_2 \cup A_3 \cdots \cup A_n) &= \#(A_1 \cup B) \\
&= \#A_1 + \#B - \#(A_1 B) \\
&= \#A_1 + \mathrm{Q}_1 - \mathrm{Q}_2 + \mathrm{Q}_3 \cdots \pm \mathrm{Q}_{n-1} \\
&\quad - \mathrm{R}_1 + \mathrm{R}_2 - \mathrm{R}_3 + \cdots \pm \mathrm{R}_{n-1} \\
&= \#A_1 + \#A_2 + \#A_3 + \cdots + \#A_n \\
&\quad - (\mathrm{Q}_2 + \mathrm{R}_1) + (\mathrm{Q}_3 + \mathrm{R}_2) \\
&\quad - (\mathrm{Q}_4 + \mathrm{R}_3) \cdots \pm (\mathrm{Q}_n + \mathrm{R}_{n-1}) \\
&= \mathrm{S}_1 - \mathrm{S}_2 + \mathrm{S}_3 - \mathrm{S}_4 \cdots \pm \mathrm{S}_n
\end{aligned}$$

Ceci est bien ce que nous voulions obtenir ; nous avons donc ainsi rigoureusement démontré la formule de Poincaré par récurrence.

On peut revenir au problème : *on jette trois dés ; quelle est la probabilité pour que l'un au moins donne un résultat impair ?* L'événement C_1 est formé des familles de trois chiffres dont le premier est impair ; il y a donc trois possibilités pour le premier (1, 3, et 5), et six possibilités pour chacun des deux autres, de sorte que $\#C_1 = 3 \times 6 \times 6 = 108$. De même $\#C_2 = 6 \times 3 \times 6 = 108$ et $\#C_3 = 6 \times 6 \times 3 = 108$. Pour appliquer la formule de Poincaré (ici pour trois événements, soit *V.2*) il faut connaître également les cardinaux des intersections par deux et par trois ; mais cela est aisé : $\#(C_1C_2) = 3 \times 3 \times 6 = 54$, $\#(C_2C_3) = 6 \times 3 \times 3 = 54$, $\#(C_1C_3) = 3 \times 6 \times 3 = 54$; enfin $\#(C_1C_2C_3) = 3 \times 3 \times 3 = 27$. Ainsi

$$\#(C_1 \cup C_2 \cup C_3) = 3 \times 108 - 3 \times 54 + 27 = 189$$

et la probabilité pour que *l'un au moins* des trois dés marque un chiffre impair est $189/216 = 7/8$.

V. 3. Le problème des coïncidences fortuites.

Le problème typique que l'on résoud par la formule de Poincaré est celui des *coïncidences fortuites* ou problème de Montmort (mathématicien français du $XVIII^e$ siècle) ; dans la littérature anglo-saxonne *random matches problem*. Ce problème s'énonce ainsi :

> *n lettres sont adressées chacune à un destinataire déterminé ; mais on les distribue au hasard entre leurs n destinataires. Quelle est la probabilité pour qu'aucun des n destinataires ne reçoive la lettre qui lui était destinée ?*

Ce problème est l'un des plus anciens du Calcul des probabilités. Il est évoqué pour la première fois dans le livre de Montmort à propos du *jeu des Treize*. Dans ce jeu, on mélangeait treize cartes de valeurs 1 à 13 (les trois dernières étant valet, dame, roi, considérés comme numéros 11, 12, 13). Puis il fallait retirer les cartes une à une et voir si la k^{e} tirée avait aussi la valeur k. Montmort a posé le problème de trouver la probabilité pour que cela se produise au moins une fois.

L'historien Todhunter rapporte que c'est probablement Nicolas Bernoulli et non Montmort qui a trouvé la réponse ; en effet, la première édition du livre de Montmort rapporte le résultat mais ne comporte aucune démonstration ; puis dans la seconde édition Montmort donne deux démonstrations en disant qu'il les tient de Nicolas Bernoulli (Montmort et N. Bernoulli entretenaient une correspondance régulière).

La célébrité de ce problème provient de ce que la limite de cette probabilité lorsque n tend vers l'infini, est égale à $1/e$ (nous ferons ce calcul ci-dessous). Qu'on puisse rencontrer des nombres tels que e ou π dans un problème concret de probabilités, était perçu comme une découverte fantastique, et s'apparentait au miracle de la géométrie qui avait déjà fasciné les Anciens[(2)].

L'explication de ce paradoxe se trouve dans le principe d'invariance qui conduit à postuler l'équiprobabilité des destinataires. L'énoncé du problème dit : *on distribue les lettres au hasard entre les destinataires*, ce que tout le monde traduit par "tous les destinataires sont équiprobables". Comme cela a déjà été discuté dans ce livre (section **I. 3** *La signification de l'équiprobabilité*), l'équiprobabilité est postulée a priori, à partir d'une invariance que le sens commun, d'origine empirique, fait percevoir comme évidente : chaque lettre est supposée avoir autant de chances d'arriver chez n'importe lequel des n destinataires. D'un point de vue empirique, cela signifie que si on refait un grand nombre de fois l'expérience consistant à distribuer les n lettres, on constatera que statistiquement elles se répartissent à peu près uniformément ; nous verrons au chapitre **XI** que les variations correspondant à cet "à peu près" doivent être de l'ordre de $1/\sqrt{N}$, N étant le nombre de fois qu'on répète l'expérience ; l'équivalence des destinataires sera donc vraie à $1/\sqrt{N}$ près : pour vérifier expérimentalement l'équiprobabilité au millionième près, il faudrait refaire l'expérience 10^{12} fois ! Et selon toute vraisemblance, on constaterait alors que l'équiprobabilité n'est pas parfaite : pour la distribution au hasard les lettres devraient être à chaque fois remélangées et tirées au sort dans un panier, mais l'une pourrait

(2) Platon, *La République*, Livre VII.

être légèrement plus lourde et se retrouver plus souvent au fond du panier, ou une autre légèrement écornée pourrait s'accrocher aux autres et avoir plus de chances de remonter sur le dessus du panier. En un mot, l'équiprobabilité empirique est par nature approximative. Cependant, comme l'équiprobabilité est un principe qualitatif et non quantitatif, il ne se trouve pas diminué par le caractère approximatif de sa vérification empirique. Il en va de même dans d'autres domaines ; si par exemple on énonce "la constante c de l'électromagnétisme (vitesse de la lumière dans le vide) vaut $2.997\,925 \cdot 10^8\ m/s$", une mesure plus précise augmentera l'information ainsi exprimée, car elle est uniquement quantitative et ne contient rien d'autre que la connaissance de décimales. Mais si on énonce "la constante c de l'électromagnétisme est la même dans tous les repères galiléens", une vérification plus précise de ce principe ne l'améliore plus (tout au plus elle peut conduire à l'infirmer).

Les philosophes de la Grèce antique ont été fascinés devant la possibilité de connaître le nombre π par les pures mathématiques, avec une précision infinie, alors que la mesure physique d'une circonférence ne pouvait jamais donner plus de trois décimales. Platon en déduisait dans le texte célèbre de *La République* (op. cit.) que l'on peut donc atteindre par la pensée une vérité située au-delà de tout ce que l'expérience sensible permet de connaître. La raison de ce paradoxe (ou de cette illusion) est que le calcul de π par la géométrie résulte du principe *qualitatif* d'une invariance de l'espace par rotation. Ce principe qualitatif est d'origine expérimentale, mais le fait qu'il soit qualitatif permet le glissement conceptuel d'une validité approchée à une validité parfaite. Si on a mesuré π expérimentalement avec six décimales, on ne peut pas deviner les décimales suivantes ; mais si on a établi à six décimales près l'invariance de l'espace par rotation, on peut deviner ou croire que cette invariance se maintiendra pour des mesures plus précises (du moins jusqu'au jour où ces mesures seront devenues précises au point de montrer les limites de l'invariance) et en déduire mathématiquement une valeur de π "exacte", avec une infinité de décimales.

La fascination des contemporains de Montmort devant la probabilité égale à $1/e$ est la même que celle de Platon devant la valeur mathématiquement exacte de π. L'équiprobabilité rigoureuse n'existe jamais en pratique ; mais si on la traduit sous forme de vérité mathématique (on appelle cela aujourd'hui "construire un modèle mathématique"), elle conduit à des nombres qui peuvent être calculés avec une précision infinie. Il est cependant bien évident que si on distribue *réellement*, *matériellement*, les lettres par tirage au sort dans un panier, l'équiprobabilité ne sera satisfaite qu'approximativement (tout comme pour les dates de naissances

du problème de la section **II.2.**), de telle sorte que les premières décimales du modèle mathématique seront correctes, mais les suivantes fantaisistes. Le calcul par les pures mathématiques donnera $1/e = 0.367\,879\ldots$, mais la probabilité pour qu'aucune lettre n'arrive à son destinataire sera de 37%.

Voyons maintenant comment on résoud le problème de Montmort. L'espace Ω est l'ensemble de toutes les distributions possibles. Chaque distribution correspond à une permutation des lettres par rapport à la distribution correcte ; comme il y a n lettres, il y a $n!$ permutations, soit $\#\Omega = n!$ L'événement défini par l'énoncé est $\mathcal{A}$: "aucun destinataire ne reçoit la lettre qui lui est destinée". Son complémentaire est $\mathcal{B}$: "*au moins un* destinataire reçoit la lettre qui lui est destinée". L'expression *au moins un* correspond à une réunion, celle des événements :

$\mathcal{B}_1$: "le destinataire $N°1$ reçoit la lettre qui lui est destinée"
$\mathcal{B}_2$: "le destinataire $N°2$ reçoit la lettre qui lui est destinée"

... ...

$\mathcal{B}_n$: "le destinataire $N°n$ reçoit la lettre qui lui est destinée"

Ainsi $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3 \cdots \cup \mathcal{B}_n$, et on peut obtenir la probabilité de $\mathcal{B}$ par la formule de Poincaré.

Ici, il convient peut-être de faire deux remarques :

1. Le lecteur peut trouver bizarre que le problème de Montmort, résolu au début du $XVIII^e$ siècle, fasse appel à la formule de Poincaré (mathématicien français 1854 – 1912). Il est clair qu'on savait calculer la probabilité d'une réunion bien avant ! Les démonstrations de Nicolas Bernoulli, mentionnées plus haut, ne faisaient évidemment pas appel à la "formule de Poincaré", mais à des procédés spécifiques.

L'appellation "formule de Poincaré", ou "théorème de Poincaré", ou encore "identité de Poincaré" pour *V.*3 ou *V.*4 provient du fait qu'on la trouvait sous cette forme générale dans le *Calcul des probabilités* (op. cit. chap **I**, note 6 et bibliographie) de Henri Poincaré. Mais les cas particuliers *V.*1 et *V.*2 étaient connus dès la fin du $XVII^e$ siècle (Jacques Bernoulli), et leur généralisation par récurrence était évidemment dès cette époque à la portée de n'importe quel mathématicien. La formule n'est donc pas une découverte de Poincaré : il l'a simplement rendue populaire.

2. Si $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3 \cdots \cup \mathcal{B}_n$, alors $\mathcal{A} = \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3 \cdots \cap \mathcal{A}_n$, où $\mathcal{A}_j$ est le complémentaire de $\mathcal{B}_j$ et $\mathcal{A}$ celui de $\mathcal{B}$. Pourquoi ne pas utiliser la formule beaucoup plus simple $\mathcal{P}(\mathcal{A}) = \mathcal{P}(\mathcal{A}_1) \times \mathcal{P}(\mathcal{A}_2) \times \mathcal{P}(\mathcal{A}_3) \cdots \times \mathcal{P}(\mathcal{A}_n)$? **Réponse :** ce serait faux, car les événements $\mathcal{A}_j$ ne sont pas stochastiquement indépendants. $\mathcal{A}_j$ est l'événement "le destinataire $N°j$ ne reçoit pas la lettre qui lui était destinée" : donc un autre la reçoit, à qui elle n'était pas destinée non plus, ce qui implique que le fait d'appartenir à l'undes $\mathcal{A}_j$ augmente la probabilité d'appartenir aussi à l'undes autres. La propriété du produit est simple et pratique, mais ne marche malheureusement que pour des événements indépendants, tandis que la formule de Poincaré est valable pour n'importe quelle sorte d'événements.

Pour calculer $\#\mathcal{B}$ par la formule de Poincaré, il faut d'abord connaître les cardinaux des $\mathcal{B}_j$, puis des $\mathcal{B}_{j_1} \cap \mathcal{B}_{j_2}$, puis des $\mathcal{B}_{j_1} \cap \mathcal{B}_{j_2} \cap \mathcal{B}_{j_3}$, et ainsi de suite. De l'équivalence des différents destinataires, on peut déjà déduire que les $\mathcal{B}_j$ ont tous le même cardinal (sinon, cela indiquerait que certains destinataires seraient plus égaux que d'autres). De même, le cardinal de $\mathcal{B}_{j_1} \cap \mathcal{B}_{j_2}$ ne peut pas dépendre de j_1 ou de j_2, ni celui de $\mathcal{B}_{j_1} \cap \mathcal{B}_{j_2} \cap \mathcal{B}_{j_3}$ de j_1, j_2 ou j_3.

Sachant que les intersections par deux sont au nombre de $\binom{n}{2}$, que les intersections par trois sont au nombre de $\binom{n}{3}$, etc. il suffit donc de calculer les cardinaux de $\mathcal{B}_1$, de $\mathcal{B}_1\mathcal{B}_2$, de $\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3$, ... $\mathcal{B}_1\mathcal{B}_2 \cdots \mathcal{B}_n$, et on aura

$$\#\mathcal{B} = n\,\#\mathcal{B}_1 - \binom{n}{2}\#(\mathcal{B}_1\mathcal{B}_2) + \binom{n}{3}\#(\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3) \cdots \pm \binom{n}{n}\#(\mathcal{B}_1\mathcal{B}_2 \cdots \mathcal{B}_n) \tag{$V.5.$}$$

Or $\mathcal{B}_1$ étant l'événement "le destinataire $N°1$ reçoit la lettre qui lui est destinée", son cardinal est évidemment le nombre de permutations des $n-1$ lettres restantes, soit $(n-1)!$ L'événement $\mathcal{B}_1\mathcal{B}_2$ est "les destinataires $N°1$ et $N°2$ ont reçu les lettres qui leur sont destinées" donc son cardinal est le nombre de permutations des $n-2$ lettres restantes, soit $(n-2)!$ En général, $\#(\mathcal{B}_1\mathcal{B}_2 \cdots \mathcal{B}_k) = (n-k)!$ Substituant dans $(V.5)$ on obtient

$$\begin{aligned}\#\mathcal{B} &= n\,(n-1)! - \binom{n}{2}(n-2)! + \cdots - (-1)^k \binom{n}{k}(n-k) \cdots - (-1)^n \\ &= -\,n!\,\sum_{k=1}^{k=n}(-1)^k\,\frac{1}{k!}\end{aligned}$$

On obtient les probabilités en divisant par $\#\Omega = n!$ d'où

$$\mathcal{P}\,(\mathcal{B}) = -\sum_{k=1}^{k=n}(-1)^k\,\frac{1}{k!}$$

Passant au complémentaire, cela donne

$$\begin{aligned}\mathcal{P}\,(\mathcal{A}) = 1 - \mathcal{P}\,(\mathcal{B}) &= 1 + \sum_{k=1}^{k=n}(-1)^k\,\frac{1}{k!} \\ &= 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \cdots + (-1)^n\,\frac{1}{n!}\end{aligned} \tag{$V.6.$}$$

En faisant tendre n vers l'infini, on reconnaît le développement en série de MacLaurin de e^{-1}, ce qui montre que pour n grand, la probabilité $\mathcal{P}\,(\mathcal{A})$ est proche de $1/e$. Une autre observation est aussi que cette probabilité est la même quel que soit le nombre de destinataires, pourvu qu'il soit grand.

L'expression $V.6$, en tant qu'expression algébrique, garde la trace de la formule de Poincaré dont elle est issue ; elle ne se simplifie pas davantage, montrant par là que le problème de Montmort relève intrinsèquement d'une réunion d'événements non disjoints et qu'il ne peut pas faire l'objet d'une réduction à un problème plus simple. Ce type de constat permet souvent en mathématique de s'assurer qu'il n'existe pas de réduction sous-jacente (ou au contraire de deviner qu'il en existe certainement une). Remarquons en passant que c'est sur ce type de raisonnement algébrique que reposent les démonstrations de l'impossibilité de la quadrature du cercle ou de l'impossibilité de résoudre algébriquement les équations de degré supérieur ou égal à cinq. En Calcul des probabilités, on peut également tirer des renseignements de l'étude de la fraction, dans la mesure où celle-ci reflète une structure algébrique : tant qu'il s'agit de probabilités a priori et exactes, elles s'expriment sous forme de fraction, dont le dénominateur est en principe le cardinal de l'espace des épreuves. Mais il peut arriver (et il arrive souvent) que la fraction se simplifie et que le dénominateur devienne alors plus petit que le cardinal de l'espace sur lequel on a travaillé. Ce phénomène signifie alors que la probabilité obtenue aurait pu être calculée sur un espace des épreuves réduit. Nous avons en effet observé à propos des probabilités conditionnelles (voir remarques à la fin de **IV.3**), que les problèmes de probabilité peuvent parfois être réduits à des espaces d'épreuves plus petits, par la considération de probabilités conditionnelles.

Un exemple de ce phénomène est fourni par le quatrième problème de la section **V.1.** : *on jette trois dés; calculer la probabilité pour que l'un au moins des trois dés donne un chiffre impair*. Nous avons introduit les événements C_j : "le dé $N°j$ donne un chiffre impair", et l'événement du problème "au moins des trois dés donne un chiffre impair" était la réunion $D = C_1 \cup C_2 \cup C_3$. En calculant avec la formule de Poincaré nous avons trouvé que $\#D = 189$, d'où $\mathcal{P}(D) = \frac{189}{216} = \frac{7}{8}$. On voit que cette fraction, contrairement à $V.6$, ne garde aucune trace de la formule de Poincaré dont elle est issue ; en outre, le dénominateur est 8, ce qui laisse soupçonner qu'on devait pouvoir traiter le problème sur un espace des épreuves de cardinal 8 seulement. La probabilité du complémentaire $\overline{D}$ de D est $\frac{1}{8}$, montrant par là qu'il ne contient qu'une seule de ces épreuves réduites ; cela devrait nous mettre sur la voie. Et en effet, $\overline{D}$ est l'intersection des complémentaires des C_j : $\overline{D} = \overline{C_1} \cap \overline{C_2} \cap \overline{C_3}$. Par ailleurs, $\overline{C_1}$, $\overline{C_2}$, et $\overline{C_3}$ sont relatifs à trois dés indépendants et sont donc des événements stochastiquement indépendants, de sorte que $\mathcal{P}(\overline{D}) = \mathcal{P}(\overline{C_1}) \times \mathcal{P}(\overline{C_2}) \times \mathcal{P}(\overline{C_3})$. Or $\overline{C_j}$ n'est autre que "le dé $N°j$ donne un chiffre pair", événement dont la probabilité est *évidemment* $\frac{1}{2}$. Par ailleurs, le dénominateur 8 signifie tout simplement qu'on pouvait se contenter de prendre pour épreuves les huit triplets de *pair* et *impair*

possibles, au lieu des 216 triplets de chiffres.

V. 4. Les textes aléatoires.

Un autre problème classique où on est amené à considérer des réunions d'événements est celui des suites aléatoires. Nous avons déjà rencontré les suites aléatoires au chapitre **I** à propos du chaos déterministe (voir **I.4**). Un problème célèbre discuté par Émile Borel et popularisé par *La bibliothèque de Babel* de J. L. Borges[(3)] est celui de l'apparition aléatoire d'un texte sensé dans une suite de lettres écrites au hasard.

Ce problème se présente ainsi : une machine écrit au hasard à la suite les uns des autres des caractères de base de la typographie : 26 lettres de l'alphabet majuscules et minuscules, signes de ponctuation, parenthèses, lettres accentuées telles que é, ê, è, ë, î, ï, chiffres de 0 à 9, blanc séparant les mots, retour-chariot avec ou sans alinea ; disons cent caractères en tout. Chaque élément successif de la suite est choisi au hasard parmi les cent caractères de base. Une suite de n caractères est donc un "mot" de n lettres écrit avec l'alphabet des cent caractères de base. Le problème relève du cas de dénombrement étudié en **II.1** : on peut écrire 100^n suites différentes. Ces suites étant supposées équiprobables, la probabilité d'en obtenir une particulière parmi toutes celles possibles est alors 100^{-n}.

Le problème étudié par Émile Borel est celui de l'apparition *partielle* d'un texte sensé particulier : quelle est la probabilité pour que dans une suite de longueur n, apparaisse au moins une fois quelque part un texte donné de longueur k ? Si la machine a écrit une suite de $10^{1\,000\,000}$ de caractères, quelle est la probabilité d'y trouver quelque part, noyé dans un océan de gallimatias insensé, le texte exact des Voyages de Gulliver ?

Ce problème revient à calculer la probabilité d'une *réunion* d'événements. En effet, appelons A_j l'événement "entre le rang $N^o\, j$ inclu et le rang $N^o\, j+k$ exclu de la suite se trouve exactement le texte cherché". Il est clair que j ne peut pasêtre supérieur à $n-k$, puisqu'à partir du rang$N^o\, j$ il doit rester assez de place pour placer le texte de k caractères. Donc $1 \leq j \leq n-k$. L'événement "le texte prédéfini se rencontre au moins une fois dans la suite" est alors laréunion $E = A_1 \cup A_2 \cup \cdots \cup A_{n-k}$.

D'après la formule de Poincaré, pour avoir la probabilité de la réunion E, nous devons déterminer non seulement les probabilités de chacun des A_j, mais aussi celles des intersections $A_{j_1}A_{j_2}$, puis $A_{j_1}A_{j_2}A_{j_3}$, etc.

---

(3) Jorge Luis Borges, *Fictions*; traduction française chez Gallimard, collection folio.

Dans le problème discuté avant (celui des lettres et des destinataires), les intersections étaient toutes équivalentes car les destinataires étaient interchangeables. Ici, la situation est un peu plus complexe car la probabilité d'une intersection ne sera pas la même selon que les textes peuvent ou non se superposer : si le texte comporte à son début une partie de longueur ℓ qu'on retrouve à la fin, il y a une probabilité non nulle pour une intersection de la forme $A_j A_{j+k-\ell}$ impliquant la superposition de ces parties. Mais pour un texte qui n'est pas autosuperposable, les intersections $A_{j_1} A_{j_2}$ pour lesquelles $j_2 - j_1 < k$ auront une probabilité nulle. En revanche les autres intersections (pour lesquelles $j_2 - j_1 \geq k$) auront toutes la même probabilité. Il en va de même pour les intersections de trois, quatre, ..., des A_j.

Ainsi la somme des cardinaux des intersections de r parmi les A_j, que dans la formule de Poincaré *V.*4 nous avons désignée par S_r, sera égale au nombre d'intersections possibles de r événements A_j sans recouvrement de texte, multiplié par le cardinal commun de ces intersections. Appelons provisoirement $G^r_{n,k}$ ce nombre d'intersections possibles. Le calcul de ces nombres $G^r_{n,k}$ n'est pas immédiat et nous y procéderons d'ici peu. Par contre le cardinal de l'une quelconque de ces intersections résulte directement de la formule de dénombrement *II.*1 : en effet, soient A_{j_1}, A_{j_2}, ..., A_{j_r} tels que les textes correspondants ne se recouvrent pas. Cela équivaut à ce que les différences entre deux quelconques des indices j_ℓ soient toutes supérieures à k. Un élément de cette intersection (une épreuve) est une suite dont les caractères situés dans les intervalles disjoints $\{j_1 \dots j_1 + k - 1\}$, $\{j_2 \dots j_2 + k - 1\}$, ..., $\{j_r \dots j_r + k - 1\}$ constituent les r occurrences du texteprédéfini et sont donc déterminés. Le nombre de ces caractères déterminés est par conséquent $r \times k$. La suite comporte n caractères en tout, de sorte qu'il reste $n - rk$ caractères à choisir de toutes les façons possibles, ce qui d'après *II.*1 fait 100^{n-rk} possibilités. Ainsi $\#\left(A_{j_1} \cup A_{j_2} \cdots \cup A_{j_r}\right) = 100^{n-rk}$. L'espace Ω de toutes les suites possibles ayant le cardinal 100^n, cela donne la probabilité 100^{-rk} pour l'intersection.

En conclusion, nous aurons d'après la formule de Poincaré :

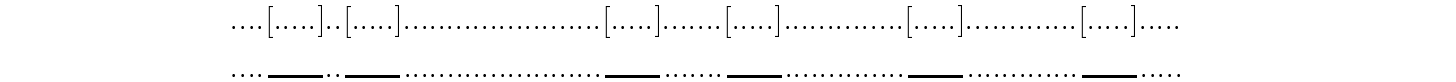
$$\mathcal{P}\left(E\right) = \frac{G^1_{n,k}}{100^k} - \frac{G^2_{n,k}}{100^{2k}} + \frac{G^3_{n,k}}{100^{3k}} - \frac{G^4_{n,k}}{100^{4k}} + \cdots$$

Bien entendu, si la machine utilise non pas 100 caractères mais un nombre quelconque m, la probabilité cherchée sera

$$\mathcal{P}\left(E\right) \quad = \sum_{1 \leq r \leq \frac{n}{k}} (-1)^{r-1} \, \frac{G^r_{n,k}}{m^{rk}} \qquad (V.7.)$$

Cette expression de $\mathcal{P}(E)$ n'est toutefois absolument exacte que si les recouvrements de textes sont impossibles ; si les textes peuvent se recouvrir en partie, il faut tenir compte d'une probabilité non nulle pour des intersections d'événements A_j dont les indices diffèrent de moins de k. Le texte des voyages de Gulliver (comme la quasi totalité des textes de la littérature) ne permet aucune superposition partielle ; pour avoir une telle superposition il faut fabriquer des textes ad hoc. Par ailleurs, même pour de tels textes spéciaux, le changement quantitatif de la probabilité est infime et ne change rien tant qu'on n'envisage que des valeurs approchées.

Il reste à calculer les nombres $G^r_{n,k}$. Ce problème ne relève pas directement des cas de dénombrement passés en revue au chapitre **II**, mais s'y ramène par une simple réduction. Représentons la suite des caractères écrits par la machine par des points sur une droite. On alignera ainsi en tout n points. Puis remplaçons chaque apparition du texte prédéfini (occurrence) par un seul segment regroupant les k points correspondants, comme le montre la figure ci-dessous où les apparitions du texte sont marquées par des crochets ($n = 100$, $k = 5$, $r = 6$) :



Si on a r occurrences du texte (ne se recouvrant pas), la droite portera $n - rk$ points restés isolés et r segments. À chaque distribution particulière de r occurrences correspond ainsi univoquement une succession déterminée de $n - rk$ points isolés et r segments. Le dénombrement de toutes les intersections possibles des événements A_j équivaut donc au dénombrement des configurations de $n - rk$ points et r segments, soit $n - r[k-1]$ éléments en tout, ce qui relève du cas **II.3**. Par conséquent

$$G^r_{n,k} = \binom{n - r[k-1]}{r} \qquad (V.8.)$$

Le problème est surtout intéressant pour de grandes valeurs de n. Dans ce cas on peut avoir des valeurs approchées pour $G^r_{n,k}$. Le coefficient binômial qui apparaît dans $V.8$ peut s'écrire (ici k remplace $k-1$) :

$$\begin{aligned}\binom{n - rk}{r} &= \frac{(n-rk)(n-rk-1)(n-rk-2)\cdots(n-rk-r+1)}{r!} \\ &= \frac{(n-rk)^r}{r!}\Big(1 - \frac{1}{n-rk}\Big)\Big(1 - \frac{2}{n-rk}\Big)\cdots\Big(1 - \frac{r-1}{n-rk}\Big) \qquad (V.9) \\ &= \frac{n^r}{r!}\Big[1 - \frac{rk}{n}\Big]^r \cdot \Big(1 - \frac{1}{n-rk}\Big)\Big(1 - \frac{2}{n-rk}\Big)\cdots\Big(1 - \frac{r-1}{n-rk}\Big)\end{aligned}$$

Reportons le tout dans la formule de Poincaré *V*.7. On obtient alors

$$\mathcal{P}(E) \quad = \sum_{1\leq r\leq \frac{n}{k}} (-1)^{r-1}\, \frac{x^r}{r!}\, a_r \qquad (V.10.)$$

où l'on a posé

$$x = \frac{n}{m^k}$$

$$a_r = \Big[1 - \frac{r[k-1]}{n}\Big]^r \cdot \Big(1 - \frac{1}{n - r[k-1]}\Big) \cdots \Big(1 - \frac{r-1}{n - r[k-1]}\Big)$$

On peut remarquer que pour les petites valeurs de r, a_r est pratiquement égal à 1 ; par contre pour les grandes valeurs de r, c'est le terme $\frac{x^r}{r!}$ qui devient très petit (en toute rigueur, la condition pour que ces deux cas ne puissent pas se présenter en même temps est que n, quoique très grand, reste petit devant m^{2k}/k). Sous cette condition on peut dire que la somme *V*.10 est pratiquement égale à la série $\sum(-1)^{r-1}\frac{x^r}{r!} = 1 - e^{-x}$. Si n est trop grand (du même ordre que m^{2k}/k, ou plus grand encore), le terme $\frac{x^r}{r!}$ sera encore grand lorsque les a_r deviendront plus petits que 1 et on ne pourra pas assimiler simplement la somme *V*.10 à la série exponentielle, mais il n'y en aura même pas besoin, car dans ce cas il est clair que la probabilité $\mathcal{P}(E)$ sera pratiquement égale à 1.

On peut donc conclure que si une machine écrit au hasard n caractères (choisis dans un alphabet de m caractères) à la suite les uns des autres, la probabilité pour qu'il apparaisse au moins une fois un texte donné de k caractères est égale à $1 - e^{-x}$ avec $x = n/m^k$. Pour $m = 100$ par exemple, la probabilité restera négligeable tant que x sera petit, c'est-à-dire tant que n sera petit devant 100^k. Les exégètes de la bibliothèque de Babel avaient découvert un volume qui contenait — comme unique passage sensé — la phrase "O temps tes pyramides" ; cette phrase comporte (blancs compris) 21 caractères. Tant que n reste petit devant $100^{21} = 10^{42}$ la probabilité de trouver la phrase "O temps tes pyramides" est nulle. À l'inverse, lorsque n deviendra nettement plus grand que 10^{42}, la probabilité de voir apparaître "O temps tes pyramides" deviendra égale à 1. Ce sera seulement lorsque n sera de l'ordre de 10^{42} que l'on verra le passage *progressif* de la probabilité 0 à la probabilité 1 ; ce passage se produira selon la loi $1 - e^{-x}$.

Une approche heuristique du problème, ne faisant pas appel à la formule de Poincaré, est envisageable. On pourrait par exemple se dire que l'apparition d'un texte sensé dans un texte écrit au hasard étant forcément un événement rare, on peut négliger les cas où un tel texte apparaîtrait

deux, trois, quatre fois, et approcher la probabilité $\mathcal{P}(E)$ par la somme $\sum_j \mathcal{P}(A_j)$. Autrement dit ne retenir que le début de la formule de Poincaré. Le résultat qu'on obtiendrait alors serait $(n-k)/m^k \simeq x$. Ceci constitue bien une approximation de $1-e^{-x}$ lorsque x est petit et en ce sens le raisonnement heuristique est correct. Mais lorsque x n'est pas petit, il devient grossièrement faux. L'explication en est simple : négliger l'apparition de plusieurs occurrences est effectivement légitime pour des événements rares ; mais nous avons vu que lorsque n devient grand par rapport à m^k, non seulement l'apparition d'une occurrence cesse d'être rare, mais devient même quasi-certaine. C'est pourquoi on ne peut pas se passer de la formule de Poincaré complète.

Jusqu'ici nous avons considéré le cas asymptotique où l'entier n est grand. Dans ce cas la somme *V.*7 comporte un très grand nombre de termes et devient pratiquement une série, comme nous avons vu.

Bien sûr *V.*7 est une expression "exacte", en ce sens qu'elle se déduit mathématiquement sans approximation de l'hypothèse d'équiprobabilité a priori des épreuves ; elle est donc tout aussi exacte pour des valeurs modestes de n. Dans la pratique des probabilités ne peuvent pas être exactes ; cela n'a aucun sens réel. L'expression *V.*7 (de même que tout calcul de probabilité a priori) n'est exacte que dans la mesure où l'hypothèse d'équiprobabilité des épreuves est elle-même exacte. Comme nous l'avons déjà discuté au début de la section **V. 3**, ce type d'exactitude est comparable à celui de la géométrie : on peut montrer que le rapport de la circonférence au diamètre est "exactement" π ou que le rapport de la diagonale au côté du carré est "exactement" $\sqrt{2}$ — ces nombres ayant une infinité de décimales déterminées — uniquement parce qu'on admet que les invariances de l'espace (par rotations et translations) sont elles-mêmes absolument exactes.

On peut dire que s'il existe une légère inexactitude dans l'équiprobabilité des caractères alignés par la machine, celle-ci se répercutera sur le résultat du calcul exact ; mais comme l'hypothèse de l'équiprobablilité des épreuves est indépendante de la valeur des paramètres tels que n, k, m, etc., il n'y a aucune raison pour qu'une inexactitude à ce niveau se manifeste plutôt pour les petites valeurs de n que pour les grandes. Par contre, l'approximation exponentielle reposait explicitement sur l'hypothèse que n est grand. Donc la formule *V.*7 s'applique en principe pour n'importe quelle valeur de n, tandis que l'approximation exponentielle ne s'applique en principe que pour n grand. On peut se rendre compte de son domaine de validité en comparant un calcul effectué directement à partir de *V.*7 au résultat donné par l'approximation exponentielle.

Cherchons par exemple les probabilités pour que, dans une suite de vingt,

cent, trois cents, mille, trois-mille, et dix-mille chiffres décimaux écrits au hasard, la suite 314 apparaisse au moins une fois. Désignons respectivement par $\mathcal{P}(20)$, $\mathcal{P}(100)$, $\mathcal{P}(300)$, $\mathcal{P}(1\,000)$, $\mathcal{P}(3\,000)$, $\mathcal{P}(10\,000)$ ces probabilités. Le tableau suivant donne les valeurs obtenues d'après *V.*7, suivies entre parenthèses par le calcul selon l'approximation exponentielle :

$$\mathcal{P}(20) = \frac{18}{10^3} - \frac{16\cdot 15}{2\cdot 10^6} + \frac{14\cdot 13\cdot 12}{6\cdot 10^9} - \cdots \simeq 0.017\,88 \quad (0.019\,80)$$

$$\mathcal{P}(100) = \frac{98}{10^3} - \frac{96\cdot 95}{2\cdot 10^6} + \frac{94\cdot 93\cdot 92}{6\cdot 10^9} - \cdots \simeq 0.093\,57 \quad (0.095\,16)$$

$$\mathcal{P}(300) = \frac{298}{10^3} - \frac{296\cdot 295}{2\cdot 10^6} + \frac{294\cdot 293\cdot 292}{6\cdot 10^9} - \cdots \simeq 0.258\,25 \quad (0.259\,18)$$

$$\mathcal{P}(1\,000) = \frac{998}{10^3} - \frac{996\cdot 995}{2\cdot 10^6} + \frac{994\cdot 993\cdot 992}{6\cdot 10^9} - \cdots \simeq 0.632\,31 \quad (0.632\,12)$$

$$\mathcal{P}(3\,000) = \frac{2998}{10^3} - \frac{2996\cdot 2995}{2\cdot 10^6} + \cdots \simeq 0.950\,49 \quad (0.950\,21)$$

$$\mathcal{P}(10\,000) = \frac{9998}{10^3} - \frac{9996\cdot 9995}{2\cdot 10^6} + \cdots \simeq 0.999\,84 \quad (0.999\,95)$$

On constate aisément que l'approximation exponentielle est déjà très bonne pour $n = 100$; sur le tableau, l'approximation ne diffère notablement du calcul "exact" que pour $n = 20$ (erreur relative de 11%).

Le problème de l'apparition d'un passage sensé dans un texte aléatoire a été utilisé par Émile Borel[(4)] pour discuter la nature du hasard. Borel considérait surtout des suites infinies (il était mathématicien). Il expliquait que si une suite de caractères infinie est *vraiment*écrite au hasard, elle doit comporter obligatoirement n'importe quel texte donné ; donc tous les textes qui ont été écrits et qui seront écrits un jour y figurent, et y figurent même une infinité de fois, car dès que l'un apparaît, la suite se poursuit comme si elle repartait de zéro. Une suite infinie dans laquelle ne figurerait nulle part un certain passage ne peut pas avoir été écrite au hasard, car il a fallu prescrire à la machine qui l'a engendrée d'éviter le passage manquant. Si toutes les suites de n caractères d'un alphabet qui en comporte m sont équiprobables, alors la probabilité de voir apparaître un certain passage de longueur k au bout de $n = x\,m^k$ caractères est $1-e^{-x}$, qui tend vers 1 quand x (et donc n) tend vers l'infini. La contraposée de cette assertion vraie est donc : si un passage donné n'apparaît jamais, alors les suites ne sont pas toutes équiprobables ; en particulier celle qui contient le passage est moins probable que les autres.

(4) Émile Borel *Le hasard* éd. Alcan, Paris, 1914, page 162.

Ces observations sur le hasard faites au début du siècle par Émile Borel ont une postérité considérable, qui est la théorie des suites aléatoires déjà évoquée (section **I.4**). Il s'agit de savoir quels sont les critères qui garantissent qu'une suite est "vraiment" aléatoire. Pour qu'il en soit ainsi, il faut donc, non seulement que chaque lettre y revienne aussi souvent que n'importe quelle autre, mais il faut aussi que n'importe quelle chaîne de caractères fixée à l'avance y revienne aussi souvent que n'importe quelle autre de même longueur. Émile Borel[5] a appelé *suites normales* les suites de chiffres ou de lettres qui satisfont ce critère. Mais il est apparu par la suite qu'on pouvait engendrer algorithmiquement des suites infinies satisfaisant à cette condition. De nombreux critères plus restrictifs ont été proposés, dont il n'est pas toujours évident qu'ils soient équivalents les uns aux autres. Le critère le plus reconnu aujourd'hui est le critère algorithmique de Solomonov et Kolmogorov (années 1960) : une suite est aléatoire si pour tout n il est impossible d'en produire les n premiers éléments avec un algorithme plus court que la simple donnée de ces n premiers éléments. Ce critère présente toutefois des variantes et des subtilités du fait qu'il n'y a pas de mesure à la fois universelle et absolument exacte de la longueur des algorithmes[6].

Revenons à notre calcul basé sur la formule de Poincaré. Il a montré que si on lit une telle suite infinie, pour avoir une chance non nulle d'*atteindre* un texte donné de longueur k, il faut lire la suite jusqu'au m^k-ième caractère environ.

Cela implique que si nous lisons à la vitesse de 10^6 caractères par heure (vitesse de lecture excessive pour comprendre la *Critique de la Raison pure* mais suffisante si on cherche seulement à repérer des passages sensés dans un texte aléatoire), nous devrons lire pendant 10^{36} heures, soit environ 10^{32} années avant d'avoir une chance de tomber sur le passage "O temps tes pyramides". Si on est moins exigeant, on peut s'estimer heureux de tomber sur un passage sensé non choisi à l'avance. Une phrase semblable à "O temps tes pyramides" peut s'obtenir en mettant à la suite quatre mots de la langue française ; on peut estimer à 10^{12} le nombre de combinaisons de quatre mots susceptibles de produire un sens (éventuellement en forçant un peu l'herméneutique). Dans ce cas, l'événement A_j considéré avant (trouver la phrase "O temps tes pyramides" entre les rangs j et $j+20$ de la suite), dont la probabilité était 100^{-21}, est remplacé par l'événement A'_j (trouver n'importe quelle phrase sensée de 21 caractères entre les rangs j et $j+20$ de la suite), dont la probabilité est 10^{12} fois plus grande, soit 10^{-30}. Pour avoir une chance non infime de rencontrer un tel texte au cours de la lecture

(5) *Les probabilités dénombrables et leurs applications arithmétiques*, déjà cité.
(6) Voir le livre de Jean-Paul Delahaye *Information, complexité, et hasard*, déjà cité.

séquentielle, il faut alors parcourir 10^{30} caractères, opération qui prendra 10^{20} années.

Le critère algorithmique de Solomonov et Kolmogorov est radicalement restrictif; la notion de hasard qu'il sous-entend exclut absolument la possibilité d'une simulation algorithmique du hasard. Or nous avons vu dès le chapitre **I** que le hasard réel et pratique est généralement un effet de chaos déterministe; les fonctions **random** simulent le hasard par déroulement d'un algorithme; et la question de savoir si dans le monde réel il y a un hasard primordial (par exemple celui de la Mécanique quantique) est une question métaphysique: elle ne peut pas être tranchée par l'expérience. C'est dire que le critère de Solomonov et Kolmogorov, pris à la lettre, est purement théorique et ne concerne pas le monde réel. Par exemple, la suite des décimales de π, $\sqrt{2}$, e, etc, est engendrée par un algorithme; mais, bien qu'on ne sache pas le démontrer rigoureusement, ces suites sont vraisemblablement normales au sens de Borel et utilisables pratiquement comme simulation de l'aléatoire: si on convertit les décimales de π, $\sqrt{2}$, e, etc, en lettres, il faudra certainement aller aussi loin dans ces suites pour trouver le passage "O temps tes pyramides" que dans n'importe quelle suite "vraiment" aléatoire.

C'est que l'algorithme qui calcule les décimales est basé sur l'arithmétique et n'a aucune raison spéciale de favoriser ou défavoriser l'apparition de la phrase fatidique; il y a une sorte d'indépendance causale entre l'algorithme arithmétique et la phrase "O temps tes pyramides".

Un des projets utopiques formulés par quelques mathématiciens (notamment Peano, Hilbert, et Russel, mais contre une grande majorité de sceptiques) aux alentours de 1900 fut la formalisation intégrale de la mathématique. Dans cette conception, les démonstrations mathématiques devenaient l'application automatique d'un algorithme; un théorème aurait alors été par définition le résultat codé d'un tel algorithme. Mais il est bien évident que ces algorithmes de déduction logique formelle seraient tout aussi indépendants de notre perception de l'espace que les algorithmes arithmétiques peuvent l'être de la phrase "O temps tes pyramides". En sorte que si on avait automatisé la déduction des théorèmes de la géométrie selon ce principe, comme l'avait rêvé Peano, on aurait dû attendre aussi longtemps l'arrivée du premier théorème *sensé* que l'apparition de la phrase "O temps tes pyramides" dans le déroulement des décimales de π ou $\sqrt{2}$.

Si on veut trouver des passages plus longs, par exemple le texte intégral de *Madame Bovary*, qui comporte environ 980 000 caractères, il faudra lire au moins $n = 100^{980\,000}$ caractères de la suite, ce qui prendra de l'ordre de $10^{1\,960\,000}$ années.

Ces nombres prodigieusement grands (ou les probabilités prodigieusement petites qui leur correspondent) ne peuvent être calculés que par des raisonnements a priori et n'ont aucun sens empirique. C'est pourquoi Émile Borel (qui a beaucoup étudié ces probabilités extrêmes) a posé le principe "les événements dont la probabilité est infinitésimale ne se produisent jamais".

Les lois déterministes de la Physique macroscopique, comme la loi de Planck étudiée en **II.6**, sont déduites par des raisonnements a priori à partir de certaines hypothèses d'invariances. Pour la loi de Planck, l'hypothèse était l'équiprobabilité des modes d'occupation pour les photons du rayonnement. Ces raisonnements a priori conduisent toujours à des probabilités infinitésimales, parce que les formules de dénombrement contiennent des puissances ou des factorielles. Ainsi on peut calculer la probabilité a priori pour que la distribution des photons selon les intervalles de fréquences diffère de la loi de Planck. Nous avons vu en **II.6** que cette probabilité diminue en fonction de l'amplitude de l'écart selon un facteur gaussien (exponentielle du carré de l'amplitude), ce qui conduit très rapidement à des probabilités extrêmes. C'est pourquoi de tels écarts "ne ne produisent jamais" et que la loi est déterministe. Nous y reviendrons encore au chapitre **XIV**.