

MODÈLES DE HASARDS

Quand passent les cigognes...
La règle de saint Benoît, qui a annoncé le «Prélude aux systèmes» du Tome du Levant, est remise à l'ordre du jour, si l'on peut dire, dans «La Dynamique libre et les modèles du temps». Elle y montre, et c'est un facteur déterminant de la civilisation occidentale, que l'on doit réguler sa vie, en rythmer inexorablement le tempo, et limiter les actes à ce que commande la Règle bénédictine, ne laissant la place ni aux hasards, ni aux tentations que permet le désordre.

Mais la profession de moine, comme tant d'autres aussi ambitieuses, se féminise; certains comportements en deviennent quelques fois variables, volatiles, imprévisibles. Intuition féminine, inconstance primesautière des blondes?... Ou histoire des Sociétés où les hommes ont tenu les commandes et imposé les règles, alors que les femmes en ont si souvent subi les dysfonctionnements, les folies égoïstes et les mâles caprices des Chefs?... Aussi, quand elles prennent en mains les rênes, elles en font des noeuds serrés pour bien ficeler le processus de leur destin. Et voilà donc que se développent les couvents, les Sœurs Bénédictines, aussi peu hasardeuses et tentatrices que leurs mâles correspondants, mais tombées encore vivantes entre les mains de Dieu. Elles sont là, en communauté, comme il se doit, dans leur couvent au jardin carré près de la Meuse.

Puis est paru le merveilleux film soviétique «Quand passent les cigognes»; elles portent, accroché à leur long bec, un petit sac blanc contenant un bébé qui poussera dans les choux, là où elles le lâcheront sur le sol. L'autre jour, on voyait deux petites cigognes qui se marraient, elles crevaient de rire à se fendre la poire d'une oreille à l'autre: comme des avions en attente d'atterrir, elles avaient tourné au-dessus du jardin du couvent des Bénédictines. Et voilà les bonnes sœurs levant au ciel des yeux effarés; leur cœur se mit à battre, l'inquiétude les prit – sur laquelle ça va tomber? Et tandis que planaient les cigognes, sur leur foi et sa certitude il se mit à planer un doute: est-ce par hasard...?

On en vient alors à la question célèbre du Doyen de la faculté de Droit: «Comment plane un doute?» Et sa réponse: «Pour le savoir, il faudrait d'abord pouvoir définir le doute avec certitude! À cette fin, il faut regarder son ombre – car cela ne fait pas l'ombre d'un doute». Mais attention, ma soeur: toutes ces ombres de doutes peuvent vous faire sombrer dans les ténèbres...

MODÈLES DE HASARDS

Sommaire

1	Les domaines des hasards	5
1.1	Cinq espaces	5
1.2	Explorations spatiales	6
1.3	Les mondes transcendants	7
1.4	Les mondes de l'Univers	8
2	Les visions des mondes	10
2.1	Les deux versions	10
2.2	Interface mythes-univers	11
2.3	L'exploration par les investigateurs	12
3	Les sources des hasards	17
3.1	Stupeur et ébranlements cosmiques	17
3.2	Caprices de la nature	18
3.3	Les facteurs naturels	19
3.4	Les hasards de Mandelbrot	20
3.5	Les aléas de l'artificiel	21
4	Les voies du hasard	22
4.1	Les agents de l'aléatoire	22
4.2	Les outputs du hasard	24
4.3	Les ensembles d'éventualités	24
5	La voie statistique	27
5.1	L'investigation statistique	27
5.2	Les fonctions de distribution de fréquences	28
5.3	Les fonctions de répartition de fréquences	29
5.4	Les descriptions par intervalles	30
5.5	La densité de fréquence	31
5.6	La fonction de densité de fréquence (ou densité "discrète")	33
6	Le potentiel et ses descripteurs	34
6.1	Les heureux événements	34
6.2	Des mesures de "champ d'incertitude"?	36

6.3	La possibilité	39
6.4	Une mesure de nécessité	43
6.5	Incertitude et dispersion	44
6.6	L'apport mutuel des mesures d'incertitude	48
6.7	Variable aléatoire	48
6.8	Les fonctions génératrices	50
6.9	Les processus aléatoires	52
7	Les contrefaçons	53
7.1	L'aléa synthétique	53
7.2	Les concentrations	54
7.3	Les convergences	58
8	Les compositions aléatoires	64
8.1	Variables conjointes	64
8.2	Une différence de deux variables aléatoires	66
8.3	Distribution d'un produit et d'un rapport	67
9	La maîtrise du hasard	75
9.1	Se fier au hasard?	75
9.2	Les voies du hasard et leurs commandes	77

1 Les domaines des hasards

1.1 Cinq espaces

Le «Prélude aux systèmes» a averti les touristes et les aventuriers de la culture que la promenade au cours de laquelle il leur est loisible d'entendre ces exposés se situe dans une interface, un éther de l'intermédiaire, entre le monde de la pensée et les mondes réels où s'énervent les Ensembles d'Activités Humaines – les "EAH". Profitant de la liberté qu'offrent ces espaces non-bornés, on peut à présent errer dans les terrains vagues de l'esprit, ceux de l'incertain, des possibles et de l'aléatoire.

Chez les Lecteurs, la marche est plus lente que la pensée; aussi, lorsqu'ils font de longs trajets monotones, ils laissent leur esprit prendre quelque distance par rapport au réel, au terre-à-terre qui est sur le chemin devant soi. Comme le Bateau Ivre d'Arthur RIMBAUD, qui a quitté son canal pour déboucher sur l'océan, les Lecteurs ne se sentent plus guidés par les rêves des Classiques et de leurs "textbooks"; leur esprit peut alors s'ouvrir comme un vaste portique.

C'est par ce portique que, sans attaches académiques ni sciences pour les justifier, on présentera ici quelques aspects des hasards en hallucinant l'audience devant cinq espaces:

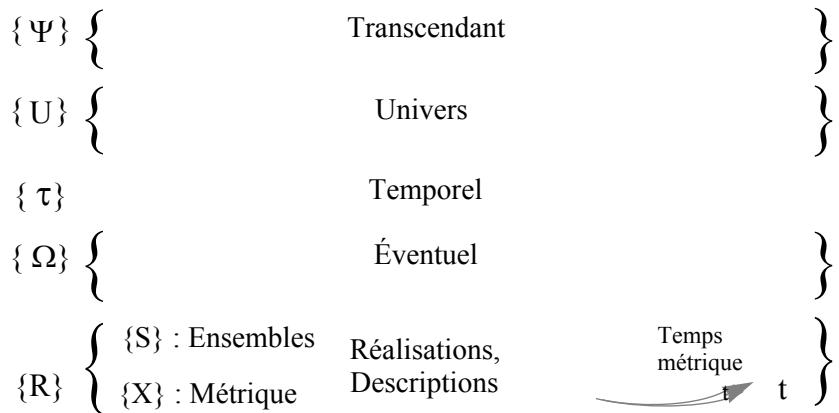
- $\{\Psi\}$ Le Transcendant: Formé des mythes, de l'imaginaire, du "psychocosme";
- $\{U\}$ L'Univers: Formé des mondes existants;
- $\{\tau\}$ Le Temporel: Le temps aurait son propre espace; il aura une "dimension" lorsqu'il sera appliqué sur l'espace métrique;
- $\{\Omega\}$ L'Éventuel: Formé de toutes les éventualités, les "possibles";
- $\{R\}$ Les Réalisations: Descriptions du réel, où les issues ont lieu, formé de:
 - l'espace $\{S\}$ des descriptions ensemblistes (dit "Sample Space");
 - l'espace $\{X\}$ des descriptions métriques, sous-tendu par des unités, des dimensions, des mesures, des valeurs. La "dimension" du temps (qui conduira à le porter sur un axe) est l'application métrique $\{t\}$ de l'espace temporel.

Comme un peu de systémique fait du bien, l'exposé va s'organiser de la façon suivante:

- Les espaces seront d'abord disposés selon un design commode pour la suite, tout simplement de haut en bas sur la Figure 1, laquelle sera développée ensuite;
- Les espaces seront partagés en ensembles et en spécifiant quelques éléments qui en font partie : les domaines des hasards;
- Des premières relations entre ces ensembles seront exprimées; certaines seront des applications d'éléments d'un espace sur un autre;
- Puis viendront les processus, les fonctions et les mesures "officielles" qui les concernent, et des formules comme on en trouve dans les livres des facultés.

Il n'y a bien sûr pas de grande invention, sinon un peu de fantasmagorie – le hasard en inspire de toutes natures – et une présentation qui jouera une Ouverture de musique romantique et puis quelques airs folkloriques servant d'entractes entre les classiques.

Figure 1. Espaces



1.2 Explorations spatiales

1.2.1 Les mondes visités

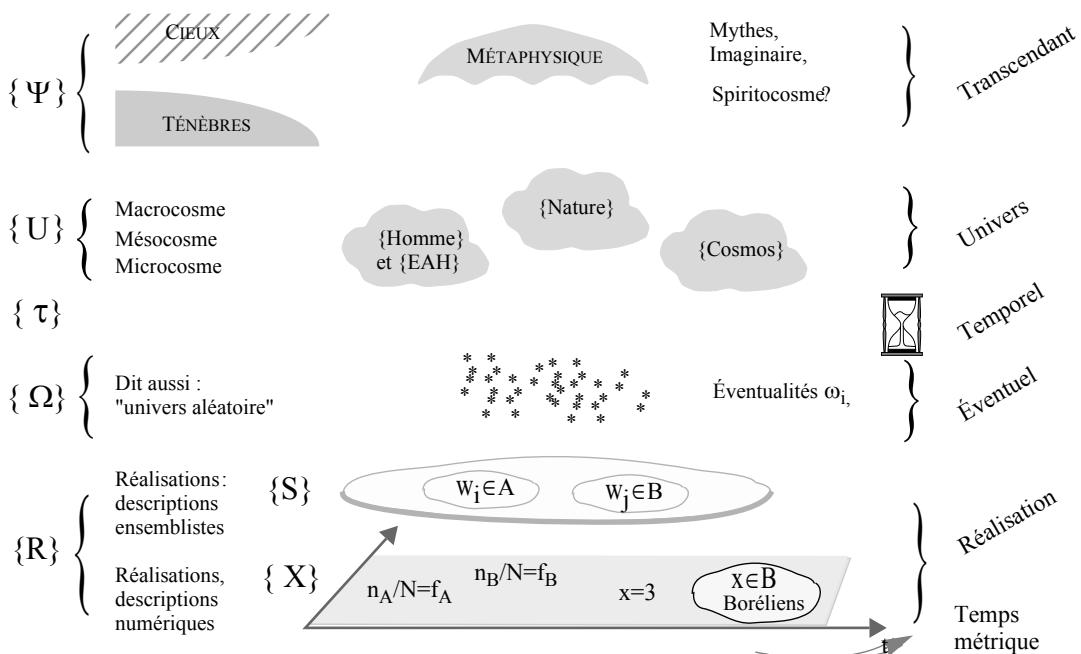
Les espaces proposés à la Figure 1 sont à présent développés ("unfolded" en anglais) pour y faire apparaître des mondes fascinants, les domaines où sont engendrés et où se produisent les effets des hasards.

- L'espace transcendant est hérité des mythes, de l'imaginaire, des religions, des habitants de l'Olympe et ses pavillons de banlieue. Il sera dans un instant exploré par des visions métaphysiques;
- Dans l'espace dit "Univers", on distingue (en y regardant de tout près):
 - le Cosmos,
 - la Nature,
 - l'Homme et ses Ensembles d'Activités Humaines (les "EAH");
 - Les recouvrements, les domaines communs aux précédents.
- L'Éventuel est formé des éventualités ; c'est tout ce qui pourrait avoir lieu, mais dans un ensemble donné (un thème, une situation, une problématique) simulé par l'esprit (sinon c'est tout et n'importe quoi);
- Les phénomènes du hasard, les possibles qu'ils engendrent, flottent dans un espace temporel;
- Les issues des hasards, ses "retombées", ce qui arrive lorsque des possibles se manifestent, se situent ici dans l'espace des Réalisations. Celui-ci est spécifié par ses descripteurs, déjà présents sur la Figure 1.

L'investigation, l'ensemble des processus mentaux par lesquels le saint esprit voyage dans ces espaces, en recueille l'information, et en fait des représentations des mondes possibles, donc produit de la connaissance.

Ces développements, commentés par la suite, sont représentés par la suite, sont représentés à la Figure 2.

Figure 2. Espaces en voie de développement



1.3 Les mondes transcendants

Les Anciens étaient victimes de forces supérieures inconnues, alors que les Nouveaux (nous-vous autres), sommes victimes de forces et de pouvoirs que nous connaissons, car ils sont des émanations maléfiques de peu confraternelles Sociétés Humaines.

Ces Anciens avaient déjà distingué plusieurs mondes, bien qu'ils n'en comprirent aucun. Le premier est situé sur les hauteurs, dans les cieux (il y a plusieurs ciels?), sur le Mont Olympe ou, en tout cas, dans des logements dignes des Êtres Suprêmes – qui squattent peut-être aussi le cœur des hommes, s'ils veulent bien se donner la peine d'y aller voir. C'est de là que soufflent des phénomènes qui dominent les humains, et dont ceux-ci se déclarent souvent victimes.

Ce monde-là ne se maîtrisant pas d'ici-bas, on ne peut que tenter de l'"amadouer"; des suggestions sont faites à cette fin dans la section 8, qui concerne "la maîtrise du hasard".

Alors que l'allégorie est la personification d'une idée abstraite, la mythologie personnifie des pouvoirs, des forces, des phénomènes et des comportements que l'homme ne contrôle pas ; elle leur fait jouer un rôle interactif, ce rôle étant une cause ou une interférence dans les affaires courantes de l'Univers. Il se pose donc depuis toujours un problème de "gestion du hasard" qu'on ne peut écarter lorsqu'ils s'agit des visions et stratégies de l'esprit humain, capable de simuler.

Comme c'est un monde de mythes (des fables, des récits), de mysticisme (relation avec une divinité), de "transcendant" et de métaphysique, il est désigné par Ψ comme Jupiter, l'Olympe et d'autres Chefs de droit divin, sans vexer personne. Noter que les mythes sont plus difficiles à maintenir qu'autrefois, car à présent ils doivent affronter la science.

Cette trinité est bien encadrée, comme il se doit, entre les cieux et l'Autre Monde, celui de l'au-delà, en général situé en bas – plus précisément, en-dessous, chez les Celtes et les Grecs en tout cas – et que l'on dit "dans les ténèbres", manifestement parce qu'on n'y voit pas clair. Ces domaines sont aussi placés sur la Figure 2, à partir de laquelle on pourra raconter quelques histoires aléatoires.

1.4 Les mondes de l'Univers

1.4.1 Les grands ensembles

Cet Univers sera ici partagé entre les mondes du Cosmos, de la Nature, de l'Homme (et ses Ensembles d'Activités Humaines, les EAH), et certains de leurs "recouvrements".

La distinction entre les différents "systèmes du hasard", et surtout les façons de les analyser et de les exprimer, demande aussi la distinction entre trois ordres de grandeur de l'univers, où les phénomènes ne diffèrent pas seulement par leur dimension, mais peuvent être qualitativement différents. Ils sont cités dans l'exposé «La Dynamique libre et les modèles du temps» comme le macrocosme, le microcosme, et le mésocosme :

- Le macrocosme et le microcosme sont les domaines qui, selon le critère principal de leur taille extrême et des limites des instruments, ne permettent pas l'observation directe ;
- La dimension du mésocosme ("moyen, intermédiaire") est qu'on peut percevoir, décrire et observer, partiellement en tout cas, avec des moyens existants et sans devoir y substituer des voies abstraites ou hypothétiques ; son degré d'observabilité en impliquera donc le contenu en information. En ayant l'esprit large, on peut proposer qu'il va de la molécule au système solaire, mais bien entendu la technologie, et les connaissances des observateurs, en déplacent constamment les bornes.

Pour faire des bulles bien nettes sur les bandes dessinées aux Figures 2 et suivantes, on fera ici le partage annoncé en 1.2 :

- Le Cosmos ;
- l'Homme (et sa population, ses sociétés, ses EAH) ;
- La Nature ;
- Les recouvrements, les domaines communs aux précédents.

1.4.2 Le Cosmos

Le cosmos, surtout le macrocosme, est resté longtemps le domaine privilégié de la physique déterministe ; le mot grec "cosmos" signifie d'ailleurs l'univers ordonné, celui des stoïciens. Il s'oppose au chaos, qui est la "masse confuse des éléments dans l'espace". De réputation, celui-ci aurait précédé le cosmos, et c'est un des mythes dont on a parlé en arrivant qui se serait chargé du passage de l'un à l'autre, donc de la création de mondes dits ordonnés.

Des versions plus récentes et positivistes ont montré cependant que le chaos et l'ordre pouvaient être engendrés l'un par l'autre. De plus, à présent, ce déterminisme-là se réduit seulement à une propriété valable seulement dans des cas-limites, ceux qui correspondent précisément aux systèmes dynamiques stables.

1.4.3 La Nature

Ce qui est désigné par la Nature sur la Figure 2 est tout le monde qui nous entoure, que ce soient des entités, phénomènes, processus ou relations qui existent sans avoir été synthétisés, rendus artificiels par l'homme ; à la limite, la Nature pourrait donc se passer de celui-ci, et il n'est d'ailleurs pas sûr qu'elle s'en porterait plus mal.

Dans le microcosme, le non-déterministe, le chaotique, était longtemps considéré comme relevant de l'imprécision, de l'ignorance : l'incertitude était dans le chef de l'analyste, de l'observateur et de ses instruments, et non dans l'objet de l'analyse. On propose plus récemment qu'au contraire l'incertitude, le chaotique et les bifurcations imprévisibles dans de nombreuses orientations improbables sont le fait de structures en instabilité dynamique et dès lors, comme l'a suggéré BOLTZMANN, la "flèche du temps" ne peut s'orienter que de façon probabiliste. Dans ce domaine, une des phrases-clefs (due à Ilya PRIGOGINE) est «Au plus loin de l'équilibre [de la stabilité dynamique], au plus grande est la multiplicité des états possibles».

Mais alors, si le système devient chaotique et probabiliste, il peut être amené à oublier son état initial, c'est-à-dire que son équation de dynamique des états ne peut inverser la flèche du temps, et donc admettre la réversibilité.

1.4.4 L'Homme

Et sa femme, ont des comportements autonomes et imprévisibles. Sont peu systémiques.

1.4.5 Les recouvrements

a L'astrologie

L'astrologie établit des connexions entre le cosmos dans sa partie mésocosme, c'est-à-dire les constellations, la nature et certaines des affaires des hommes. Établissant des relations entre la prédestination, le destin et les hasards de la vie, elle est particulièrement pertinente pour l'usage de la systémique en gestion, pour la stratégie et les grandes options de la politique.

b Le naturel

Le recouvrement du naturel et du domaine de l'Homme concerne :

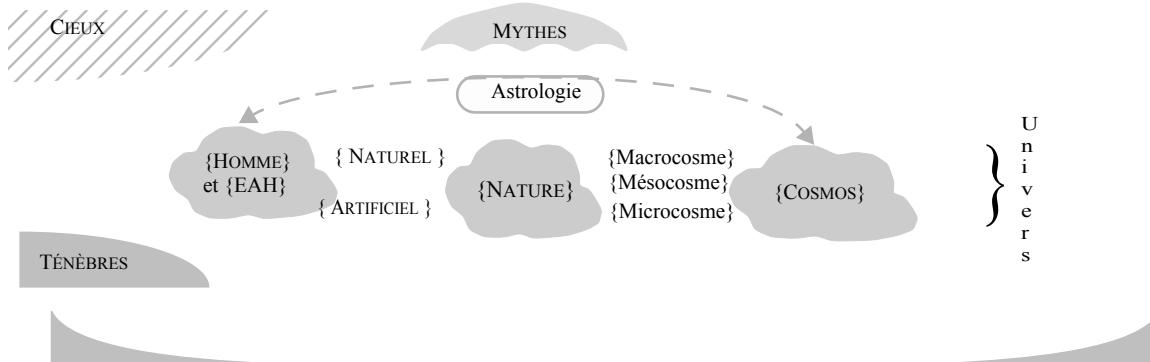
- Ce qui lui est "naturel", au sens de inné, non superposé par l'activité et la société humaine; on y mettrait aussi l'organicisme, la génétique, bien que ces domaines-là soient aussi progressivement pris en mains par l'artificiel;
- L'ensemble des états et phénomènes autonomes, situés dans la nature, que l'homme tente de découvrir et d'élucider mais qu'il n'y établit pas. Du point de vue du hasard, on y parlera d'"états de la nature", expression qui désignera ensuite les circonstances aléatoires en statistique dite "décisionnelle".

c L'artificiel

L'artificiel tente de maîtriser, remplacer, reproduire par synthèse certains aspects de la nature en relation avec l'homme. Ce domaine est un gros client de l'étude des hasards, précisément parce que c'est un contexte que l'on vise à maîtriser et prédire, et même que l'on fabrique dans ce but. Comme il se situe dans le mésocosme, ce sera aussi le domaine des expériences et de la recherche de lois et de prédictions qui y sont associées.

La Figure 3 présente ces meilleurs des mondes.

Figure 3. Les grands mondes des hasards



2 Les visions des mondes

2.1 Les deux versions

Le hasard des mondes a deux versions, l'ancienne et la nouvelle :

- L'ancienne version suppose l'existence du monde mythique, et dès lors des relations et influences avec l'univers, par exemple la providence, la fatalité, et les caprices des dieux : quelques cas notoires seront repris ci-après dans les "relations mondaines" ;

- La nouvelle version fait fi des mythes ou, en tout cas, il les remplace par d'autres moins hallucinatoires et mieux formulés. En simplifié (oh combien !), elle avance que les hasards sont engendrés par les phénomènes inhérents aux mondes de l'univers ; ils sont émergents au lieu d'être des bourrasques issues de l'éther céleste qui en balayaient les lois et en disloquaient les règles patiemment élaborées.

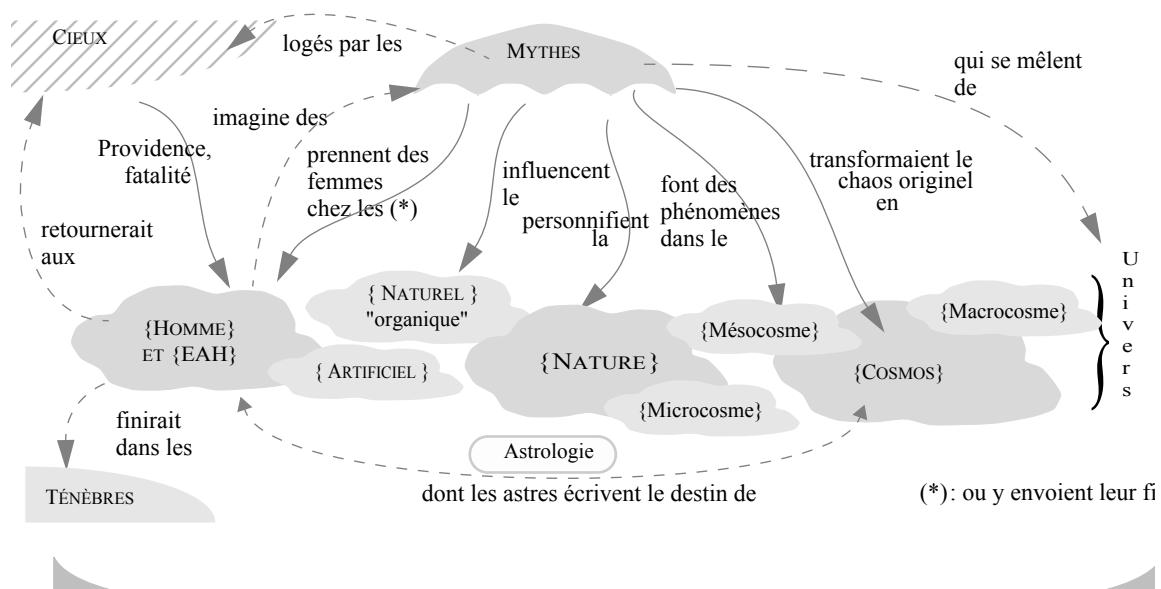
Cette vision des choses situe les hasards par exemple dans les facteurs d'onde et les perturbations du macrocosme, les innombrables états possibles dans le microcosme, les fantaisies de la nature, les variétés des organismes, les comportements imprévisibles des hommes, et enfin les facteurs de complexité des Ensembles d'Activités Humaines.

2.2 Interface mythes-univers

La Figure 4 ose à présent quelques-unes des relations mondaines spéculatives entre les mondes représentés sur la Figure 3. L'idée est de situer des histoires typiques, parmi la panoplie que présentent les histoires que se racontent les hommes au coin du feu et ce, depuis les brasiers des cavernes jusqu'aux holocaustes nucléaires.

La thèse représentée ici (correspondant à l'"ancienne version" ci-dessus) est que des hasards sont des caprices de relations entre le monde des mythes et ceux de l'Univers que ces mythes influencent. Ainsi, les cieux envoient des messagers pour transformer le Monde, des ZEUS frappent des coups de tonnerre où bon leur semble, NEPTUNE lève une tempête, VULCAIN forge une vengeance, on se permet d'enlever des petites femmes d'ici-bas pour les emmener au septième ciel, on transforme le Chaos en Cosmos, on écrit de là-haut le destin des hommes dans les astres, le tout orchestré par un formidable « Dies Irae Dies Illa », c'est-à-dire les colères divines dont l'Univers tremble encore et qui engendrent des séismes au tréfonds des âmes.

Figure 4. Interface mythes-univers



Ces relations demandent des interfaces pour les entretenir : des sacrifices incantatoires, des invocations de saintes personnes qui intercèdent, sont censées influencer ces hasards venus d'en-haut. Ces essais sont évoqués dans la section 8, «La maîtrise des hasards», située en fin de parcours pour ne pas troubler les esprits, ou biaiser les statistiques des malheurs.

Et le fin, le subtil, le raffiné probabiliste doit vivre en paix dans tout cela. Mais voici bien-tôt ses formules magiques, ses relations incantatoires qui font peu-à-peu donner à notre monde à la fois le hasard, la liberté et la... probabilité.

Il est temps que les hommes découvrent les hasards et cessent de croire qu'ils ont tout réglé, tout engendré par des lois qui le prédestinent si bien qu'il devient, devant ses saints préposés, un bureau de réclamation pour ceux dont le sort est défavorable. C'est vrai qu'il y a des tas de gens peu gâtés par le destin, et il y en a même fort peu qui sont nantis. Dans cette humanité défavorisée, tant d'entre eux sont dirigés par des pouvoirs cupides et orgueilleux qu'il vaudrait mieux qu'ils soient portés par les courants aléatoires, même entre des récifs, que pilotés par des pillards avides de butin et de captives non-bénévoles.

2.3 L'exploration par les investigateurs

2.3.1 Les fonctions de l'investigation

L'exposé sur «L'Investigation» apprend que c'est une fonction de production volontaire de connaissance, exercée pour l'élucidation d'une problématique. Il cite diverses approches de la connaissance, dont les Grecs anciens (donc toujours pas les nouveaux?) en distinguaient les formes suivantes :

- L'*epistêmê*, le savoir abstrait à vocation générale;
- La *technê*, le savoir abstrait à vocation pratique;
- La *mètis*, connaissance "rusée, tacite et furtive", disons obtenue par le renseignement;
- La *phronesis*, sagesse pratique et sociale.

Comme il s'agit pour l'instant d'explorer quelques imageries, on restera prudemment dans un vague *epistêmê*, laissant à l'exposé spécialisé – ou mieux, à des gens bien payés – le soin de faire des apports utiles. Il s'agit donc de visions naïves de l'univers en tant qu'objet sur lequel l'esprit humain – par ses processus et ses fonctions – peut faire des observations ou des hypothèses dont la vraisemblance est suffisante pour se prêter à l'analyse. Le processus en question exploite des fonctions de l'esprit (jugement, intuition, perception, émotion, rationalisation, selon JUNG et SINGER).

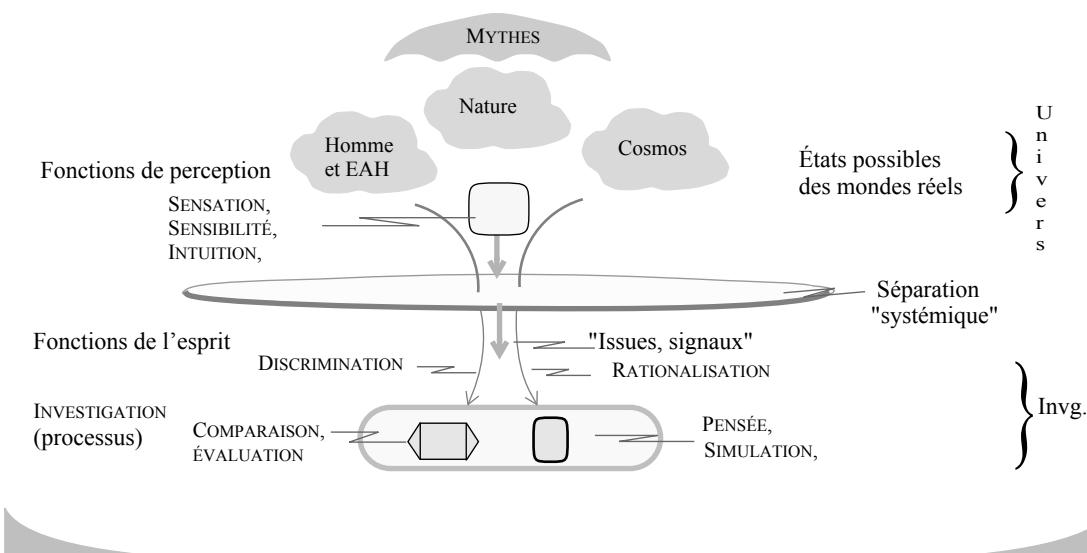
Cette distinction entre l'univers ("objet") et l'esprit humain ("sujet") qui en est le témoin, juge, concepteur et interprète permettra de séparer le monde réel de celui de son investigation, l'interface étant la systémique, comme il a été si élégamment exposé dans «La Genèse des systèmes» du Tome du Levant.

Ainsi on forme une trinité : Mythes, Univers et leur saint Esprit qui en est l'investigation. La Figure 5 montre une réunion de cette trinité, assise comme pour la Cène à la table de la systémique.

Un partitionnement des fonctions y est proposé (ce que ne font pas JUNG et SINGER), distinguant :

- Les fonctions de perception, impliquant la sensation, la sensibilité, dont les sens se forment des signaux transférés aux processus ;
- Les fonctions de discrimination, conduisant à l'évaluation et à la comparaison ;
- Les fonctions mentales comme penser, simuler. Une proposition est que les hasards seraient formés de l'ensemble des éventualités forgées par la fonction de penser, c'est-à-dire les simulations de potentialités, qu'elles soient réalisables ou imaginaires.

Figure 5. L'investigateur



2.3.2 Grands esprits

On a mis dans le cosmique le macrocosme et le microcosme, ensachant bien que chacun d'eux se "rétrécit" en faveur du mésocosme, puisque le domaine de l'univers observable par l'homme et ses instruments devient relativement de plus en plus couvrant. L'histoire des modèles déterministes dans le monde cosmique est un des grands, peut-être le plus grand, processus de l'histoire de la science, avec pour leaders l'astrologie et ses dérivés, mais avec des sous-produits intéressants, dont la physique, la mécanique céleste, la relativité, le modèle quantique – et on en passe, mais pas de meilleurs !

Au point où ils en sont, ces savants-là, on ne va pas les exhumer ici et les trahir, ou essayer de les rattraper dans leurs élucubrations ; on va seulement reconnaître que ce sont les premiers grands "systèmes", de pensée, de philosophie et de propositions connectées et ordonnées.

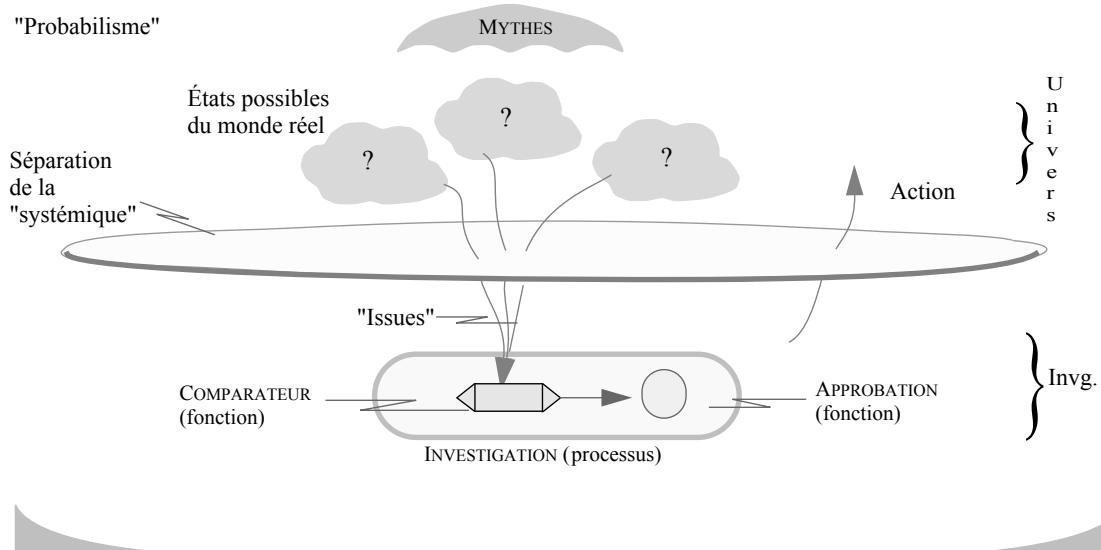
Ainsi en est-il de ce gentleman égyptien, décédé depuis longtemps dans les crues du Nil qui ne l'a pas cru, qui a fondé la thermodynamique en proposant que le soleil se chauffait en frottant sur la voûte céleste, ce qui n'est pas si bête à une époque où la majeure partie de l'humanité grelottait dans des huttes et des cavernes.

Un point intéressant est que, chez les peuples qui ont initié les thèses et les observations "scientifiques" en Occident, les propositions relatives au hasard ne sont pas antérieures à celles qui recherchent des règles ou des lois et ce, peut-être parce que l'inexplicable et le fortuit étaient déjà là, étaient le problème posé par la vie quotidienne. Un exemple en est l'opposition DÉMOCRITE-CARNEADES, laquelle est située bien après le "système" astronomique très régulier de PTOLÉMÉE.

2.3.3 Les maîtres d'âmes

CARNEADES est un Grec ancien (pas un nouveau, sinon on n'en parlerait pas), connu vers les vertes années de -150 avant J.C. comme brillant orateur – ce qui est normal, puisque c'est en causant si bien que les Grecs ont fait avancer tant de choses. De réputation, ce bel Héllène serait le fondateur du probabilisme, doctrine didactique de la morale par laquelle on peut suivre l'opinion la moins sûre, si néanmoins elle est probable – voilà donc un ancêtre de la théorie des possibilités, dont une imagerie est proposée à la Figure 6.

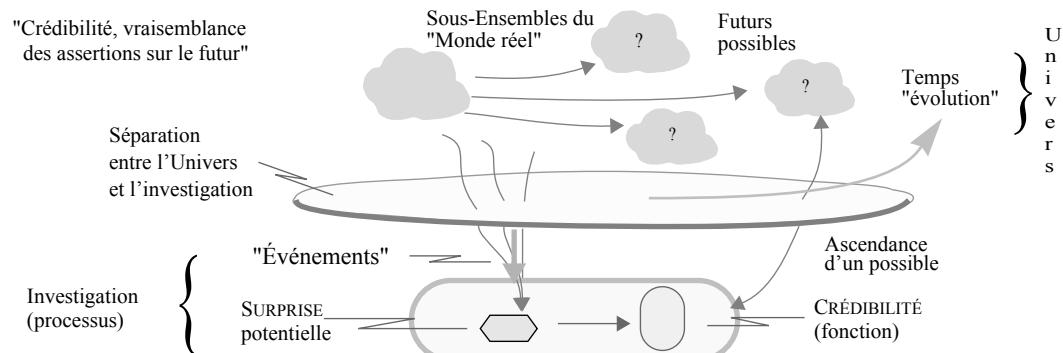
Figure 6. L'investigateur de Carneades



Cet investigateur de CARNEADES est talonné de près par celui de SHACKLE (1960) et POPPER (1959, dans L'univers irrésolu, Hermann, Paris, 1984), montrant 2150 ans de progrès sur le même thème. Les images faites ici, avec leurs philactères, ne sont bien sûr que des extraits de bandes dessinées (ici la Figure 7), qui n'impliquent pas les grands auteurs cités, décédés dans la Sous-France sans avoir pu s'en défendre.

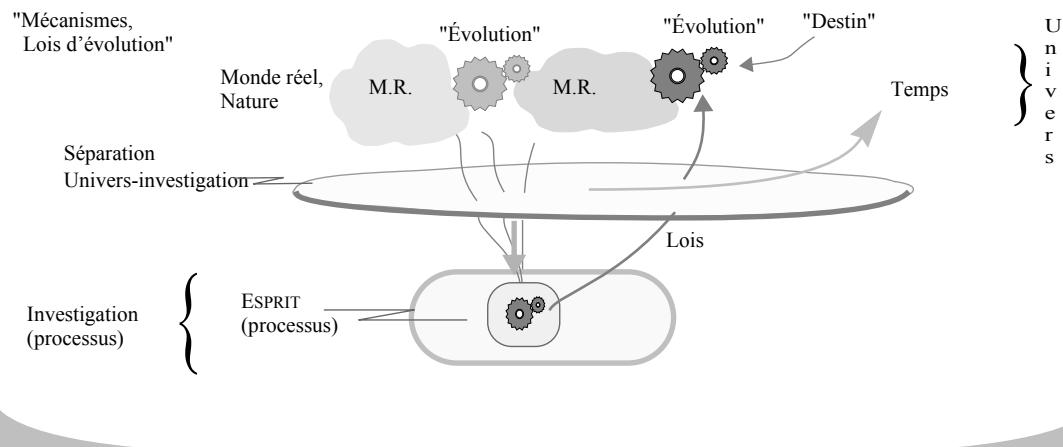
La portée de cette confrontation est d'avoir induit une (seule ?) certitude chez le présent auteur : *hn'yadansnotremondeaucunévénement,phénomèneoucomportementqu'aitqu'un seul facteur qu'il'aitengendré. Rien, ni maladie, ni guerre, ni chromosome n'est isolé, sans interaction, rien n'est déterminé par une cause unique.* Cette vision des choses, par la multiplicité des possibles qu'elle engendre, réconcilie les processus générateurs organisés et les aléas qui y sont associés.

Figure 7. L'investigateur de Schackle et Popper



Par provocation, on peut y opposer comme annoncé le paradigme mécaniciste de DÉMOCRITE, lequel utilise son "esprit" (c'est lui qui le dit) pour en proposer des lois concernant l'évolution. Les engrenages de la Figure 8 sont toutefois un peu anachroniques...

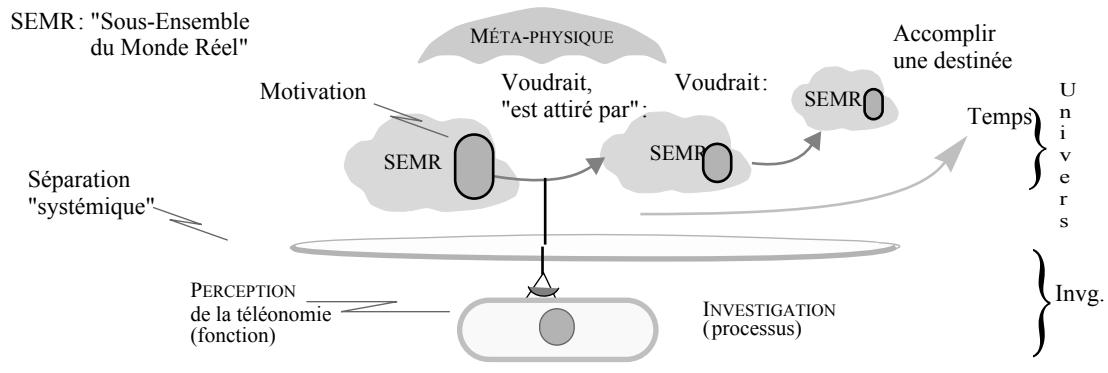
Figure 8. L'imagerie mécaniciste de Démocrite



2.3.4 Le Destin et la destinée

Par confrontation avec l'évolution mécaniciste, gérée par des lois, et un destin offert par Monsieur DÉMOCRITE, on peut présenter l'aspect téléologique proposé par ARISTOTE. Celui-ci implique une certaine finalité, qui n'est pas hasardeuse mais qui n'est pas non plus définie par le monde réel, physique. Elle est issue d'une certaine métaphysique qui le coiffe et prend le relais des "mythes" un peu trop naïfs pour le grand homme. Le "télos" implique un but qui est, pour les sous-ensembles du monde réel et leurs agents, d'accomplir leur destinée ; c'est la motivation en ce sens qui mènerait leur évolution.

Figure 9. Téléologie et Destinée d'Aristote

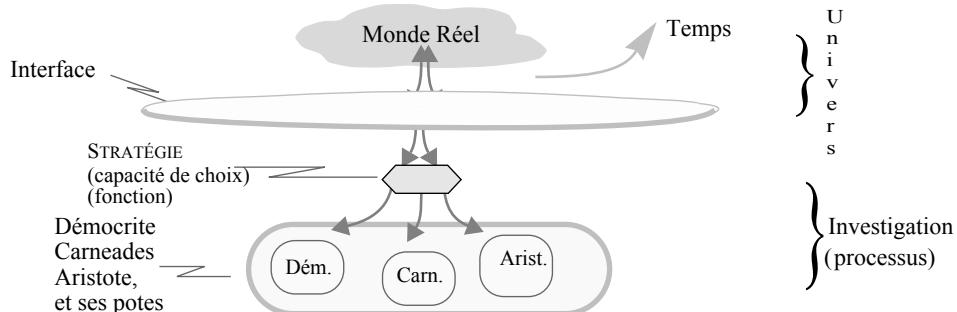


2.3.5 Un stratège

Certains penseurs plus récents, à peine décédés, comme SINGER, présentent non pas un paradigme spécifique de relation avec les mondes incertains, mais un paradigme "englobant", appelé stratégie. Celle-ci implique une capacité de choix entre différentes façons d'appréhender ce qui est mal connu, dont aux premières loges les versions mécaniciste, probabiliste ou téléologique qui viennent d'être mentionnées. SINGER montre donc un aspect de la systémique qui n'impose pas "la" voie d'investigation, mais laisse choisir celle qui est, aux yeux de l'investigateur, la plus cohérente avec celle de sa "vision du monde", la fameuse "Weltanschauung" de la philosophie allemande.

Cette "façon de voir" de SINGER est sur la planche 10 de cette bande dessinée.

Figure 10. La "stratégie" d'investigation de Singer



3 Les sources des hasards

3.1 Stupeur et ébranlements cosmiques

Il relève de CARNEADES que "tout est hasard"; mais selon ce gars-là, cela n'est pas dû à une totale aberration des comportements de la nature: c'est parce que l'esprit humain ne peut parvenir qu'à des propositions probables, et pas à des certitudes. Amusant: c'est le gros problème de la théorie quantique de notre époque, 2000 ans après, ce qui conduit Albert EINSTEIN, pensant que le monde est "intelligible", et toujours mal à l'aise devant les aspects probabilistes de la théorie quantique qu'il a néanmoins fondée par ses propositions sur les quanta (d'énergie), à écrire:

«Tu crois au dieu qui joue aux dés, et moi à la seule valeur des lois».
EINSTEIN et BORN, Correspondance 1916-1955, Paris, Seuil, 1972, p.165.

Voilà donc qu'en sciences physiques bien des constructions déterministes, dont la "mécanique" céleste de NEWTON et autres abatteurs de mythologies, ont été prises à partie par du non-déterminisme. Parmi les propositions les plus célèbres (et leurs auteurs-clés) qui y ont joué un rôle on peut citer:

- Le principe d'incertitude de HEISENBERG, impliquant la non-neutralité de l'objet relativement à son observation;
- Les aspects statistiques des états de la matière (avec BOLTZMANN). Cet aspect est expliqué gentiment dans l'exposé «L'Information et sa thermodynamique», qui passionne les visiteurs de ce Tome Nord;
- Les aspects probabilistes de la mécanique quantique (N. BOHR, L. DE BROGLIE);
- Les théories du chaos, des fractals (B. MANDELBROT);
- Les discontinuités dites catastrophes (R. THOM);
- Les bifurcations dues à de faibles perturbations (I. PRIGOGINE);
- La prolifération des jeux de loto (depuis l'Italie), et surtout le scepticisme dubitatif et incrédule des Lecteurs des présents exposés.

Le déterminisme est donc pris à partie, et il en reste si peu de nos jours qu'il est peut-être plus sage d'essayer de commander aux aléas qu'aux processus régis par des "lois".

Lorsque les progrès ont été dans le sens de la description des aspects physiques de l'univers par des formulations, celles-ci "mangent" ou prennent en charge des parties de cet univers, lesquelles deviennent de moins dues au hasard et doivent plus à des lois. Si l'espace qui entoure les EAH est mieux connu, certaines craintes dues à l'incertitude s'évanouissent alors, et les contributions des auteurs des lois déterministes sont les bienvenues, car ils sont plus rassurants que les demi-dieux capricieux.

On n'a pas dit, cependant, que l'élaboration d'un modèle déterministe donnait toujours de bonnes nouvelles: on pourrait prédire avec une excellente précision actuellement que dans quelques semaines une très grosse comète va nous péter toutes les dents de lait.

Les déviations par rapport à ces lois – où la périodicité et le degré de répétition ont joué un rôle primordial pour engendrer la vision déterministe – sont considérées comme des perturbations "surajoutées" au phénomène fondamental, vision qui prévaut encore aujourd’hui, à juste titre, en théorie des signaux.

Plus récemment, les non-conformités à des lois trop rigides (citées ci-dessus) ne sont pas nécessairement des rejets des théories élaborées, lesquelles sont bien montrées et compatibles avec les observations réalisables en leurs temps, mais plutôt des élargissements des champs d’investigation, à la fois des objets et des approches pour les appréhender, c’est-à-dire les décrire et les expliquer, conditions pour ensuite les modéliser.

Il est vrai aussi que la mathématique d’une part, et la théorie des probabilités et la statistique d’autre part, ont fait un long parcours parallèle, c’est-à-dire sans s’épauler mutuellement pour faire face à des problématiques communes. Les remises en cause pour les raisons évoquées ci-dessus ont contribué à en faire des partenaires, bien que le probabiliste distingué soit parfois surpris de la naïveté statistique dans des théories de la physique de haut niveau...

3.2 Caprices de la nature

On dit que tous les goûts sont dans la nature. On ne dit pas où est le mauvais goût, mais on exclut qu’on l’ait en soi.

Il en va de même pour les hasards : beaucoup de gens se plaignent de leur sort, ou des sorts qui leur sont jetés, peu se plaignent de leur propre hébétitude. Beaucoup s’attendent à ce que ce soit leur tour – dès que possible, s’il vous plaît ! – d’être l’heureux élu de providences qui en ont comblé d’autres. Ils oublient que la plupart des heureux événements arrivent là où on leur a préparé des circonstances favorables plutôt que d’attendre que leurs fruits tombent mûrs dans la bouche. Ils oublient aussi que les hasards font des tas de choses, mais certainement pas la justice.

La plupart des gens voient encore les événements fortuits comme ponctuels, produits par une seule éclosion d’un hasard. Cela ressemble aux cas des "malchances", aux "tuiles" – parfois un coup de bol – alors que même un systémicien de série C y subodore un processus d’interactions sous-jacent. Évidemment, pour ces naïfs, la pluie qui tombe pile au moment où ils mettent l’allumette au barbecue renforce leur point de vue.

Les écritures disent qu’ARISTOTE plaçait déjà la naissance de l’aléatoire non pas dans les phénomènes et événements spécifiques – qu’on envisageait déjà de comprendre et expliquer un jour – mais dans les conjonctions d’événements, les interactions de phénomènes. Dans les modélisations ultérieures, on y retrouve la théorie de la complexité par interactions, ce qui est évidemment légitime, ainsi que l’explication d’événements par le fait qu’ils sont en réalité l’aboutissement de scénarios (malencontreux ou non), ou de "fautes" résultant de défauts de coordination ou d’information mutuelle.

Une version moderne n’est plus celle des lois d’une part et des perturbations, des parasites, d’autre part, mais le fait que des processus qui ont toutes les apparences du hasard peuvent être engendrés dans tous les milieux, même les plus déterministes à première vue.

Des exemples en sont les suivants :

- Des équations engendrant des fractals et le chaos ;
- Des bifurcations "quasi" imprévisibles, des changements de phases qualitatifs en raison de faibles perturbations, et bien d'autres aventures associées aux non-linéarités, aux interactions, aux discontinuités "naturelles" ;
- Des hasards apparents dûs à des actions volontaires des EAH dont les conséquences sont inattendues et trop dispersées pour être sans surprise.

3.3 Les facteurs naturels

Pour l'observateur, les facteurs qui lui font penser qu'il subit un hasard sont :

- Le mode de génération est un processus indéterminé : il est de nature stochastique, mot issu de "stochastikos" : "habile à viser", puis "le devin", en grec ancien (le neuf ne voulantsansdouteriendire,puisque'ennes'yréfèrejamais). Ces processus stochastiques sont des suites d'événements engendrés sous l'influence du hasard ; ils se sont développés parallèlement aux mathématiques déterministes. Les systèmes du hasards devaient donc être issus de la rencontre des deux ;
- Il y a une complexité qui dépasse l'entendement, donc la description immédiatement observable. La complexité peut être simplement due au nombre et à la variété ;
- L'objet peut être non-observable, indescriptible. C'est une raison pour laquelle l'association s'est faite la première entre le chaos et les gaz. Le mot "gaz" est, comme tant de choses belges aujourd'hui, un mot flamand, introduit par le médecin-chimiste hollandais J.B. VAN HELMONT en 1632 dans son *Ortus Medicinae*. Il aurait eu le génie de l'extraire du grec ancien (toujours encore jamais du nouveau) "chaos".
- Le fait de "ne pas pouvoir être résumé" – c'est l'acception courante de "chaotique" ;
- La présence de variables absentes (?) ou cachées.

Cette proposition avance qu'une composition aléatoire affecte apparemment une description ou sa capacité prédictive parce que des variables sont cachées, ne sont pas repérées et donc pas reprises dans la formulation. C'est une des thèses citées par A. EINSTEIN à propos de la théorie quantique, à savoir que sa formulation probabiliste n'est pas due aux facteurs hasardeux de la nature, mais à sa description incomplète.

Ce dernier point important mérite un commentaire. Bien que son argument soit valide dans beaucoup de contextes, il est explicite dans les modèles économétriques, spécifiquement dans le modèle linéaire de régression multiple :

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n + u$$

Le vecteur aléatoire u , contenant les écarts par rapport à la "loi" $b'x$, est censé prendre en charge les effets de facteurs, non décrits par les x_i , qui sont trop multiples et volatils pour être représentés et, de plus, ils doivent en être indépendants ; c'est donc aussi un cas de "description incomplète".

Dans ce cas de formulation économétrique, au plus on introduit de variables explicatives non-corrélées à 100%, ou plus on "mange" la variabilité aléatoire, ce qui soutient formellement la proposition de facteurs absents.

Cependant, ce faisant, on peut ajouter des facteurs non-pertinents ou redondants. On peut alors expliquer un ensemble spécifique d'observations mais, par manque de parcimonie, le modèle ne pourra pas être accepté comme loi de comportement; ce ne sera qu'une description ponctuelle, un bon ajustement.

3.4 Les hasards de Mandelbrot

Les propositions de Benoît MANDELBROT (résumées dans *Fractals, hasard et finance*, Flammarion, 1997) secouent les tenants de la croyance que la maîtrise des phénomènes peut s'obtenir malgré des perturbations et des influences aléatoires et, d'une façon générale, en dépit des coups de semonce du hasard.

Dans son ouvrage cité, les fractals sont définis comme suit : «Les fractals sont des objets – qu'ils soient mathématiques, dus à la nature ou dus à l'homme – qu'on appelle irréguliers rugueux, poreux ou fragmentés et qui, de plus, possèdent ces propriétés à toutes les échelles». Visuellement, les objets fractals ont donc la même forme qu'ils soient vus de près ou de loin. Il en donne pour exemples les côtes de la Bretagne, un buisson de choux comestibles appelés "brocoli" et quelques autres.

MANDELBROT semble avoir combiné l'analogie physique, la mathématique, son intuition et un sens du "spectacle écrit" (ses exposés) pour proposer que le hasard ait trois acceptations, qu'il qualifie lui-même d'"états" du hasard, par analogie avec le solide, le liquide et le gazeux pour les corps. Ces états sont le hasard bénin, sauvage et lent. Il considère alors la géométrie fractale comme la "géométrie du hasard sauvage", et du "déterminisme chaotique". En effet, son propos est alors le suivant : les "lois du hasard", essentiellement décrites par les fonctions de densité appliquée à la probabilité (qui ne serait qu'un cas particulier de la théorie de la mesure), ne sont valables que pour les processus stochastiques reconnus, acceptant les comportements moyens et les convergences.

Les appuis en faveur de cette dernière assertion sont notamment, selon le présent auteur :

- Des régularités fondamentales supposées de la part de la Nature (et son univers), dont les comportements peuvent alors être exprimés par des "lois";
- Des "nécessaires" auto-rééquilibrages, et l'homéostasie des processus subissant des impulsions. L'homéostasie est définie ici (donc hors de son domaine, l'organisme) par la capacité de maintenir, ou recouvrer, un état d'équilibre en dépit d'interférences pouvant mener à une discontinuité ou à une catastrophe – un changement d'état qualitatif, au-delà du quantitatif;
- Les convergences statistiques qui sont issues des expériences répétitives dans des conditions homogènes, ce qui implique une stabilité plus grande de la part des agrégats que de la part des éléments individuels.

La théorie des probabilités adaptée au "hasard sauvage" est chez MANDELBROT moins développée que sa mathématique. Son exemple dominant est situé dans le cadre de la bourse, et exploite la fonction de CAUCHY, qui effectivement ne possède pas de propriété de convergence de l'espérance mathématique, ce que l'on exposera à la section 6.

3.5 Les aléas de l'artificiel

Le hasard est un générateur de phénomènes et processus à la fois fortuits et inexpliquables, qui engendrent des éventualités qui, aux yeux de l'observateur, non seulement sont indéterminées, imprévisibles, mais peuvent même ne pas figurer sur une liste des attentes possibles. Ces éventualités livrent des issues, des "occurrences", lesquelles sont organisées dans des sous-ensembles pour devenir formellement des événements.

L'origine du hasard est modeste, venant discrètement de l'arabe "al-zahr", le jeux de dés. Sublimée par les mythes des (vieux) grecs, cette belle cause est souvent personnifiée ; les événements non maîtrisables, de la foudre venant du ciel aux coups de foudre venant du cœur, sont des caprices des dieux. Ainsi "hazardous" veut dire "pouvant présenter un danger fortuit" en anglais, ce qui est encore le cas en français pour l'adjectif "hasardeux". C'est pourquoi on a dit (P. VALÉRY ?) que «le hasard est le pseudonyme que Dieu se donne quand il ne veut pas être reconnu pour responsable de ses actes».

La notion d'aléa ("coup de dés" en latin, plus prosaïque) est en revanche plus proche de la terre et de l'artificiel. Initialement c'est celle d'un risque, d'un incident qui peut avoir des conséquences défavorables, et qui est engendré par des causes fortuites. En français, cette connotation de risque et de pénalité s'est cependant affaiblie et les aléas peuvent être attribués à une influence du hasard dans des processus stochastiques généralement pré-rés, qui vont se pointer un peu plus loin dans cet exposé.

Dès lors, on peut parler de l'aléatoire en termes de trois caractéristiques : le court terme, la prédiction, la complexité, commentés ci-après.

a Le court terme

Il faut convenir de ce que l'aléatoire, surtout lorsqu'il est confronté au déterministe, est "à très court terme". Plus précisément, il ne qualifie formellement qu'une seule ou très peu de transitions d'état.

Soit donc un processus dans un état donné – qui pour être défini comme tel doit être reconnaissable s'il se reproduit. Un processus est déterministe si, placé dans les mêmes conditions initiales, il aboutit au même état lors de répétitions. Le "même état" signifie que si des écarts sont constatés, ceux-ci ne sont pas significatifs pour le phénomène considéré, et ne sont dus qu'à des faits marginaux tels que la difficulté d'observation ou à des signaux perturbés.

b La prédiction

Un processus peut être déterministe – et même bien connu – mais de prédiction difficile, notamment en raison de sa complexité. Ainsi certains phénomènes comme l'attraction de deux corps est très claire depuis les travaux de NEWTON, l'exprimant le rapport du produit des masses au carré de la distance ; cependant la dynamique d'attraction de trois corps de masse inégale est un challenge de dynamiciens, et de mathématiciens (traité par H. POINCARÉ en topologie).

Or des corps, il y en a nettement plus que trois... et certains exultent !

Il en est ainsi aussi pour des processus engendrant des états chaotiques : de très faibles différences paramétriques, ou des conditions initiales numériques très voisines, peuvent faire passer le processus dans des états tellement chaotiques (visuellement) qu'ils sont totalement inattendus compte tenu des précédents. Ils donnent l'impression qu'un vent de folie a soufflé sur la procédure itérative, laquelle aurait perdu la "raison". L'issue peut donc être "apparemment" aléatoire, parce qu'elle est difficile à décrire ou à distinguer.

L'absence de capacité prédictive est donc un critère très courant, mais certes pas inexorable, du caractère déterministe. De plus, un processus qui donne des résultats inattendus ou incohérents à cause de phénomènes incontrôlés qui l'influencent n'est pas aléatoire "per se" pour autant; cependant, il serait plus pertinent alors de considérer à la fois le processus-clef et, disons, cet "environnement turbulent".

Dans ce contexte des processus, l'aléatoire qualifie donc la non-reproduction des transitions d'état sous les mêmes conditions initiales. Cela devrait déjà être vrai d'un état à l'autre, et il n'est pas nécessaire d'exiger que des processus se mettent à errer, ou être agités de tremblements qui les déstabilisent, pour les qualifier d'aléatoires.

c La complexité

Il n'est point besoin non plus de grande complexité, de variables cachées ou d'influences fabuleuses pour faire de l'aléatoire : il suffit que la transition d'état soit incontrôlable; ainsi personne sans doute ne peut gérer l'issue d'un jet de dés normaux fait à bonne distance de leur terrain d'atterrissement. Ceci dit, le responsable de la transition d'état peut être décrit, le processus peut être résumé par des conventions et mesures, ainsi que par des bornes de ce qu'il peut faire (par exemple de 2 à 12 pour deux dés). Toutefois, ces issues restent dispersées sur des réalisations qualitatives, ou sur des valeurs, sur des intervalles numériques, ou encore sur une variété théoriquement infinie.

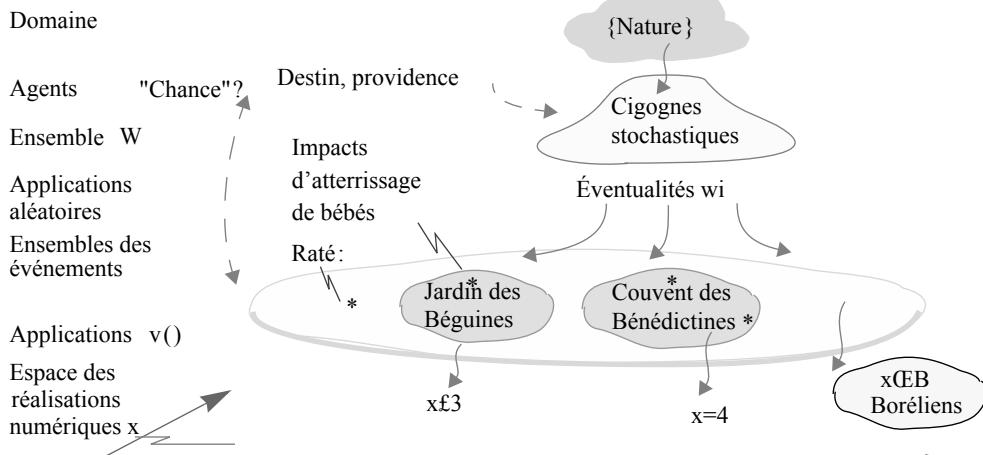
4 Les voies du hasard

4.1 Les agents de l'aléatoire

Les agents des hasards de l'Univers engendrent des éventualités (les ω_i) distinguables par un observateur, situés dans l'espace des événements. Des fonctions vont en faire une application sur un espace plus terre-à-terre, celui des ensembles réels. La Figure 11 en montre un exemple très typique : des cigognes, désirant se distraire un peu, se promènent avec un bébé au bec au-dessus d'un espace de réalisation formé du jardin des Béguines et du couvent des Bénédictines.

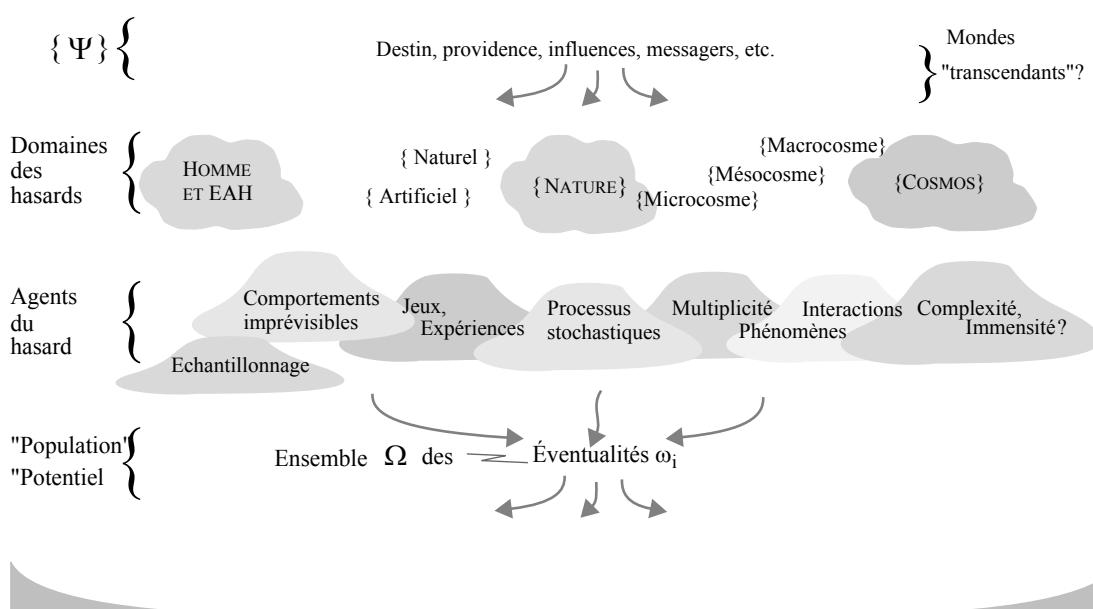
Il y a aussi dans le voisinage d'autres Congrégations faisant partie de l'ordre des Boréliens. En effet, les investigateurs (de l'évêché ?) intéressés à la chose peuvent y installer des capteurs qui font le transfert sur l'espace métrique des réalisations numériques. Là, tout sera compté et on établira des dossiers qui seront transmis à la secte des statisticiens.

Figure 11. Retombées du hasard



La Figure 12 généralise cette imagerie par laquelle des "agents" forment des ensembles aléatoires (des éventualités) à partir des ensembles-sources du hasard, dit l'"Univers" pour ne pas être mesquin.

Figure 12. Ensembles hasardeux



4.2 Les outputs du hasard

Pour ce qui est des sources, c'est fort bien d'être mythomateux, co(s)mique ou naturel, mais quelle que soit leurs ascendents et leur génétique, les hasards doivent bien de temps à autre arriver quelque part ou faire quelque chose qui les manifestent. On n'imagine pas des hasards qui errent deci-delà, dans un univers indifférent, dans des cieux sans dieu, parmi des hommes extra et terrestres qui n'y croiraient pas. En termes de systémique, on dira que, dans les différents domaines que l'on vient de citer, les hasards engendrent ou influencent des processus, donc des outputs. Ces outputs seront ici appelés des éventualités, désignées par ω_i à partir de dès à présent, dans la suite et sur la Figure 12.

Le mot "éventualité" désigne ici "tout ce qui peut arriver" de façon fortuite et non maîtrisée pour toute investigation dans un contexte de problématique donné. L'ensemble global des éventualités a été désigné par l'Éventuel, de symbole Ω .

Les éventualités se particularisent selon les contextes, les habitudes d'auteurs et de milieux douteux. Par exemple, lorsqu'elles sont issues du monde humain, leur ensemble est souvent appelé "population" en raison de travaux de la statistique concernant ce domaine pionnier. Dans ce contexte, la population a la connotation d'un ensemble dénombrable, mais pas exhaustivement observable; le nombre total N de ses éléments est alors l'effectif total, qui est le cardinal de l'ensemble, c'est-à-dire la population.

On parle de membre pour désigner tout élément de cet ensemble, et tout sous-ensemble observable issu de cette population est un échantillon, avec l'idée de s'en servir comme son représentant. Ce vocable est souvent étendu aux autres contextes, et le mot population désigne couramment tout ensemble qu'étudie la statistique; par exemple la population des pièces fabriquées, ou des échecs scolaires etc. De façon plus générale, on parle des "éléments de la population" dans le cas du comptage.

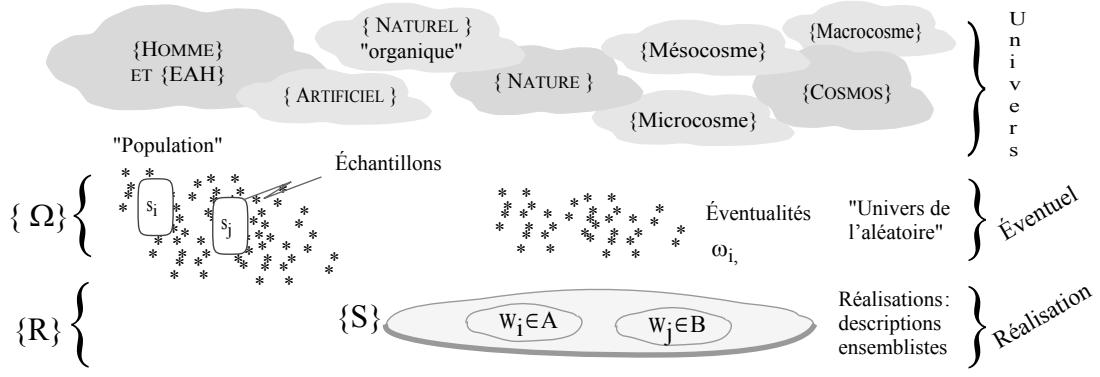
Dans le domaine probabiliste issu des jeux de hasard, on parle plus facilement de "cas possibles", lesquels seront ensuite caractérisés en "favorables" ou "succès", mais on continue aussi avec des effectifs, des fréquences et autres conventions habituelles.

En théorie des probabilités, les expressions sont plus générales que celles de la statistique, qui en est un bon client lorsqu'elle a besoin de formulations. L'ensemble des éventualités n'y est pas nécessairement dénombrable, et la théorie n'est pas dépendante d'un contexte. Cet ensemble de tout ce qui peut se manifester dans une problématique donnée, est appelé l'"Univers", désigné par Ω . Dans cet espace des éventualités planent, en attente de "réalisation", toutes les simulations possibles de l'univers réel de référence. La Figure 13, allégée du "transcendant" et des rêveries, montre ce peut arriver dans cet Ω .

4.3 Les ensembles d'éventualités

Soit une population dénombrable Ω , de cardinal N , formée des éléments ω_i , $i=1, \dots, N$. Tout sous-ensemble de Ω est appelé un échantillon. Les éléments ω_i de cette population ont des attributs qui intéressent l'investigateur. Les attributs intéressants sont portés sur des ensembles descripteurs par une application $S(\omega)$.

Figure 13. Espace des éventualités



Les ensembles de description donnent un nom et un degré d'appartenance aux ensembles représentant ces propriétés; soit A un tel ensemble: l'écriture $S_A(\omega)$ donne le degré d'appartenance de ω à l'ensemble A.

Par exemple, il peut s'agir d'un échantillon de personnes dont l'attribut intéressant est la corpulence. La propriété peut être "gros", "maigre" (dont le féminin est "fine"), "malingre" (dit "mince" au féminin) etc... La description ensembliste serait "A est l'ensemble des gros", "B l'ensemble des maigres", etc.

Plus généralement, soit l'ensemble Ω de toutes les éventualités, pas nécessairement fini à présent. La convention est d'exprimer le fait qu'il se passe quelque chose dans Ω par l'appartenance d'éléments ω_i à l'un ou l'autre ensemble. Pour cela, il faut définir les ensembles de Ω qui conviennent à la poursuite formelle de l'élaboration de systèmes et processus dans les milieux du hasard: ce sont des familles de parties de Ω , c'est-à-dire des tribus (ce qui rappelle d'ailleurs l'origine "population").

Ces tribus sont désignées par $\{S\}$ sur les figures précédentes. Ce sont des êtres mathématiques qui doivent satisfaire aux conditions suivantes:

- S contient Ω ; donc $\Omega \in S$;
- Pour tout ensemble $A \in S$, le complément de A appartient aussi à S (c'est la stabilité par rapport au complément);
- S est stable pour la réunion dénombrable: pour toute suite (A_i) d'éléments de S:

$$\bigcup_{i \in \mathbb{N}} A_i \in S$$

Il s'ensuit que la tribu contient:

- L'ensemble vide \emptyset , puisqu'il est le complément de Ω ;
- Toute réunion finie d'éléments de Ω ;
- Toute intersection finie ou dénombrable d'éléments de Ω .

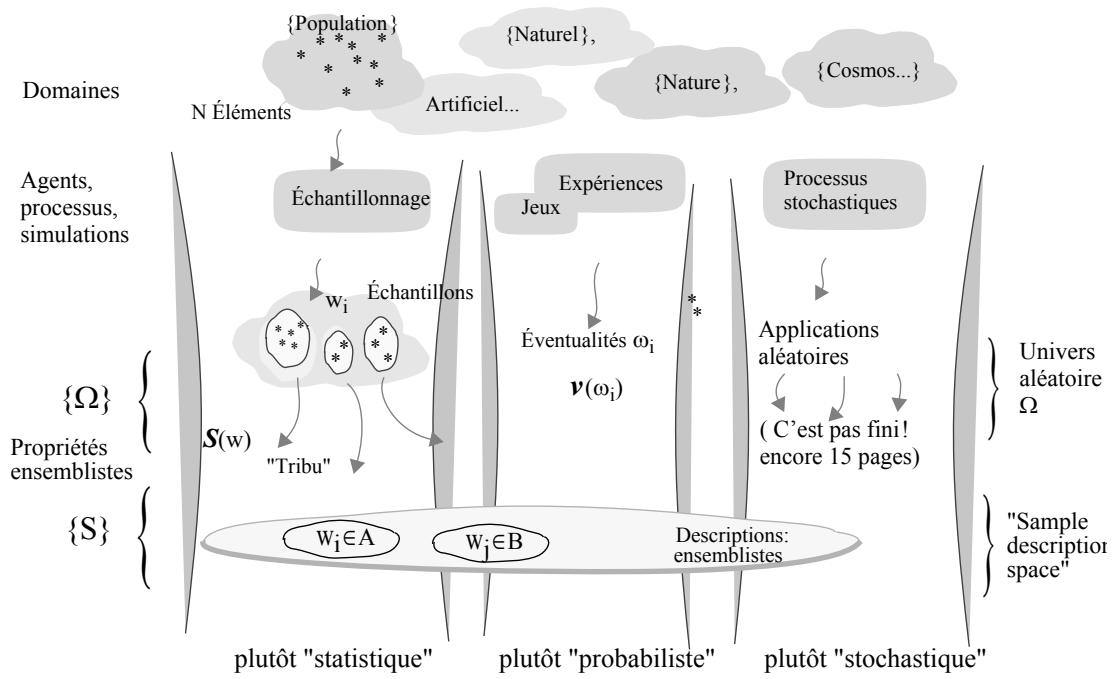
Dès lors, un espace des éventualités est un couple formé d'un univers et une tribu de cet univers: $\{\Omega, S\}$. Un événement est alors défini comme un élément de la tribu S .

Lorsque l'ensemble Ω est fini, les mêmes conditions définissent un "clan", qui est donc aussi une famille de parties de Ω , plus restrictive dans ce cas. La tribu est un être mathématique défini comme il vient d'être dit, mais une autre expression officielle qui y correspond est celle de " σ -algèbre". Celle-ci est aussi rigoureuse et adéquate, mais il n'est pas facile d'accepter par des Lectrices impatientes, sinon frétilantes, la relation entre cette expression et la définition, plus intuitive, des ensembles formés par les éventualités "qui pourraient se produire".

La justification du choix, fait ici, de $\{S\}$ comme symbole de cette famille d'ensembles est inspirée par son expression de "Sample description space", en anglais, où "sample" signifie "échantillon".

La Figure 14 tente de placer visuellement ces objets et concepts. Elle a aussi l'audace de placer des zones de "voies dominantes" du hasard, à savoir la statistique, le calcul des probabilités et les processus stochastiques ; on y a associé leurs "agents" le plus couramment producteur de varité aléatoire. Cette trinité n'est surtout pas une "distinction" de voies différentes ; elle n'est que didactique et historique de par le contexte, car les recouvrements sont importants et ces voies sont complémentaires.

Figure 14. Dominances des voies



Un cas particulier important est la tribu de Bernouilli, qui porte bien son nom du fait que les BERNOUILLI forment une grande famille de savants et de leurs épouses et belle-mères.

Ils ont inventé l'appartenance dichotomique, à savoir que la réalisation de ω est dans l'ensemble A ou n'y est pas. Donc $S(\omega) \in A$ ou $\in \bar{A}$, le complément de A . La tribu est donc formée de Ω , A , \bar{A} , et \emptyset , l'ensemble vide. Une bonne illustration en fut donnée par une étudiante, à savoir le jet d'un pièce de monnaie non biaisée : la réalisation est A si la pièce retombe pile, \bar{A} si ce n'est pas pile, Ω si elle retombe sur les deux côtés à la fois, et \emptyset si elle ne retombe pas du tout.

Une tribu définie sur un ensemble fini est restreinte au clan. On peut définir alors le clan de BERNOUILLI comme étant une alternative. Avec ce petit engin, on peut développer toute la théorie classique des distributions de probabilité, et ses fonctions continues si on peut faire appel à des convergences. Quant au "non-classique", plus récent, on parle de possibilité, nécessité et sous-ensembles flous, auxquels une section est consacrée.

5 La voie statistique

5.1 L'investigation statistique

L'investigateur "statistique" s'intéresse principalement au repérage et au comptage des attributs d'une "population" dénombrable (ce qui ne veut pas dire observable exhaustivement) via des sous-ensembles de celle-ci, appelés généralement échantillons. Il considère une ou plusieurs propriétés, qui peuvent être de nature soit qualitative, soit quantitative :

- Le fait pour un élément de posséder (ou non) une propriété qualitative peut s'exprimer par l'appartenance (ou non) de cet élément à un ensemble défini par cette propriété. De la sorte, conformément à sa définition, il s'agit d'un événement. Par exemple, le fait pour un produit d'être défectueux, ou d'une personne d'avoir telle affection ou de relever de telle profession.

On aura alors $S(\omega) = s_i, \dots, s_j, \dots, s_m$ selon la variété de propriétés relevées.

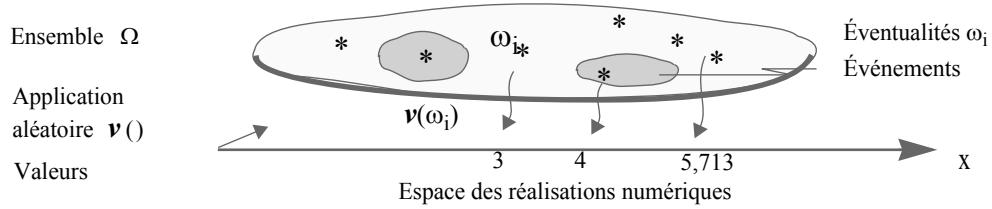
- Si la propriété est de nature quantitative, le même raisonnement tient, mais l'ensemble d'appartenance doit alors être défini dans l'espace numérique ; dans ce cas, les familles d'ensemble seront définies par un borélien, ce qui ressemble à une tribu définie par des intervalles.

On aura alors $X(\omega) = x_i, \dots, x_j, \dots, x_m$ selon la variété d'états numériques considérés.

Si les éléments de la population ne possèdent qu'une seule propriété quantitative intéressant l'investigation, on peut associer à cette propriété un nombre réel que l'on appelle la valeur de l'élément. Cette application se fait donc de l'ensemble Ω sur l'ensemble des réels, et porte le nom très logique de valorisation, dont le foncteur est $V(\omega)$, situé sur la Figure 15.

Une application numérique univoque de ce type est appelée série statistique.

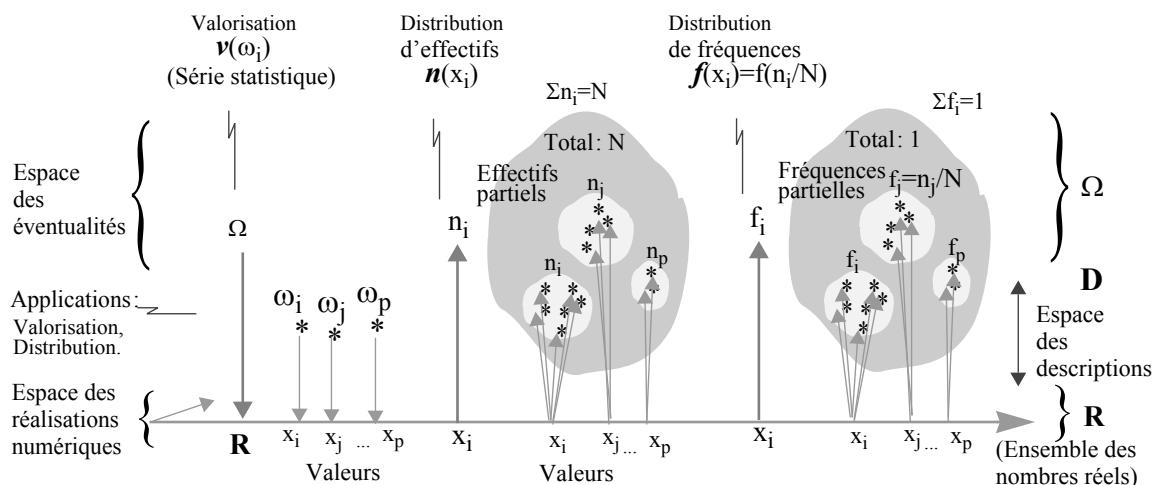
Figure 15. La fonction de valorisation



5.2 Les fonctions de distribution de fréquences

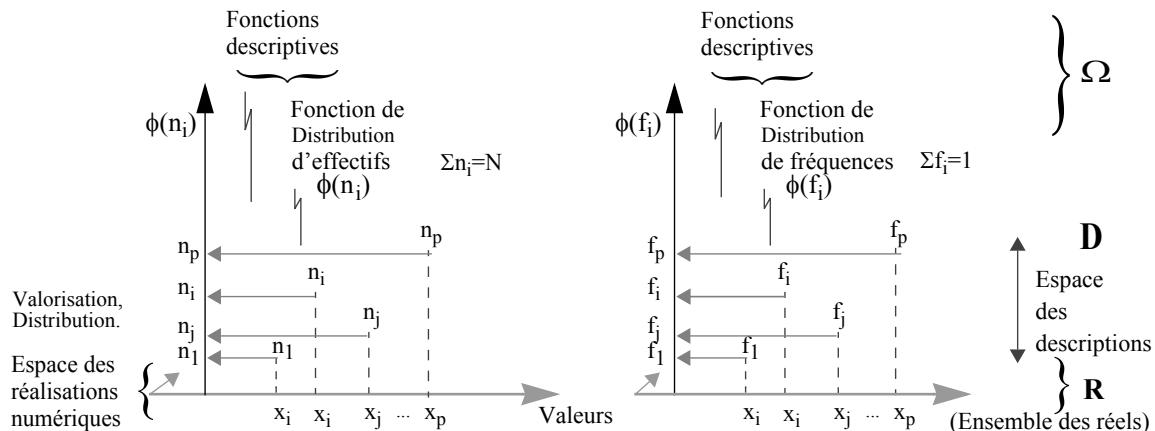
- Le nombre d'éléments d'une population finie est appelé son effectif total – disons N . La fonction de valorisation $V(\omega)$ peut fort bien appliquer plusieurs membres de la tribu sur la même valeur x_i ;
- Le nombre de membres pour une valeur donnée de x_i , s'appelle l'effectif partiel de x_i ;
- L'application de l'ensemble des valeurs x_i sur leur effectif partiel est la distribution (d'effectifs);
- Si l'expression des effectifs partiels est le rapport au total (donc n_i/N), elle s'appelle la fréquence, et l'application des valeurs de x_i sur leur fréquence est la distribution de fréquence. Ces applications descriptives sont imaginées sur la Figure 16.

Figure 16. Les applications descriptives



La Figure 17 montre l'application sur l'espace des descripteurs ; le domaine étant discret, ce sont pour l'instant les fonctions de répartition des effectifs et des fréquences.

Figure 17. Les fonctions de distribution



5.3 Les fonctions de répartition de fréquences

a Effectifs cumulés

Soit une série statistique jusqu'à l'indice α . L'effectif cumulé est la somme des effectifs des valeurs x_i jusqu'à cet indice α (donc $n_1 + n_2 + \dots + n_\alpha$), soit:

$$C_x(n_\alpha) = \sum_{i=1}^{\alpha} n_i$$

b Fonction de répartition d'effectifs

Une fonction de répartition d'effectifs est une application des valeurs x_α sur leur effectif cumulé, soit:

$$x_\alpha \rightarrow \sum_{i=1}^{\alpha} n_i = C(x_\alpha)$$

c Fonction de répartition de fréquences

La fonction de répartition de fréquences est une application qui se définit de la même façon (via $f_1 + f_2 + \dots + f_\alpha$) en faisant le rapport des effectifs au total N :

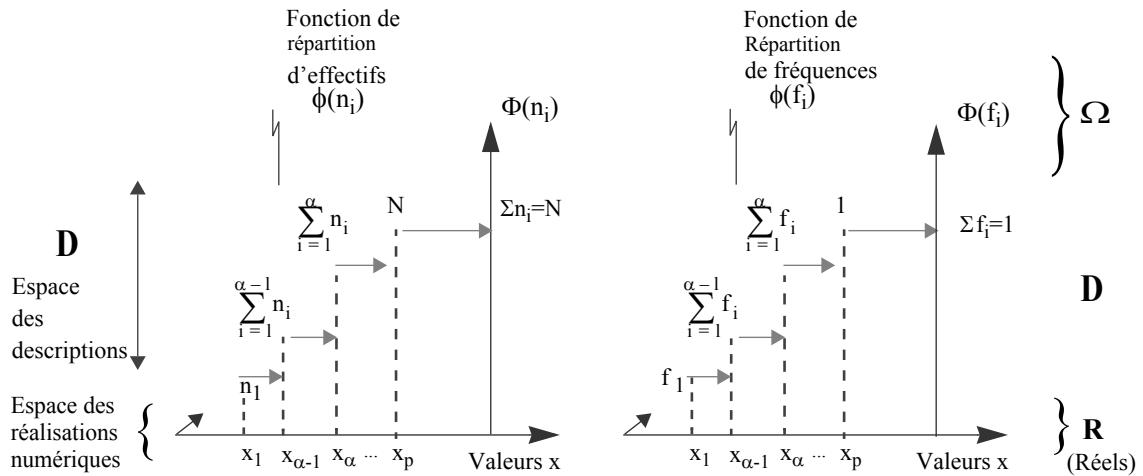
$$x_\alpha \rightarrow \sum_{i=1}^{\alpha} f_i = F(x_\alpha)$$

Il va de soi que, comme $\sum f_i = 1$,

$$0 \leq F(x_\alpha) \leq 1$$

Ces fonctions de répartition sont présentées à la Figure 18. La représentation est simplement celle des cumuls des effectifs et fréquences de la Figure 17, l'application se faisant sur l'axe (vertical) où se projette l'espace de description l'espace de description.

Figure 18. Les fonctions de répartition



5.4 Les descriptions par intervalles

Outre ses qualités descriptives propres, et son emploi fréquent en statistique, la description par classe permet d'aborder de façon simple et familière les notions de masse et de densité. Les séries statistiques présentant de nombreux caractères distincts sont souvent décrites en se servant d'une suite croissante de valeurs $(b_1, \dots, b_j, \dots, b_p, \dots)$ formant des intervalles ouverts à gauche et fermés à droite.

Toutes les valeurs x du caractère appartiennent à un tel intervalle, disons de $[b_{j-1} \text{ à } b_j]$, $j=1, \dots, p$.

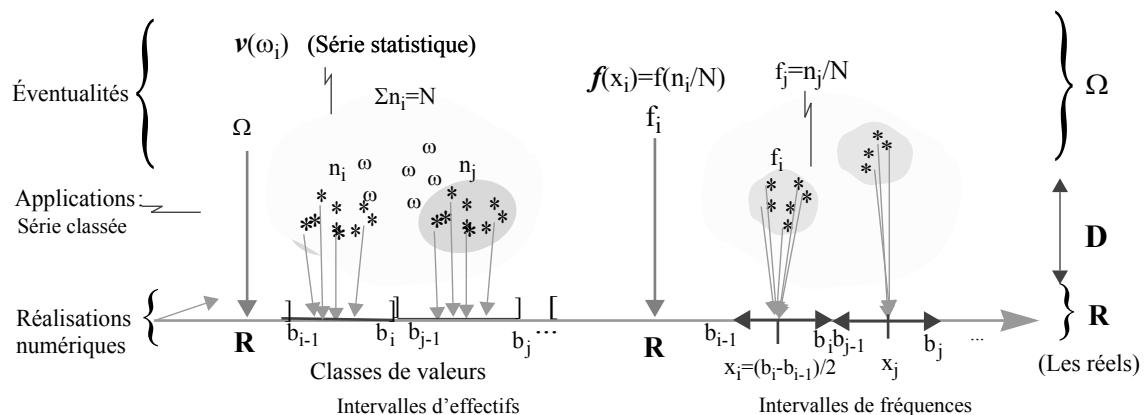
Dans le cas de statistiques de populations dénombrables, ces intervalles sont communément appelés des classes, lesquelles sont nécessairement contiguës et forment une partition complète de \mathbf{R} .

Les diverses façons de constituer des classes dépendent du propos de l'investigateur, sinon des options "par défaut" des programmathèques de statistique. Des critères en sont l'homogénéité des effectifs, ou des intervalles de même étendue. Tous les éléments de la population se projettent sur l'une ou l'autre de ces classes, mais dans ce dernier cas, leur répartition peut être très inégale, de sorte que certaines sont plus "habitées" que d'autres.

Une concentration des classes est une application par laquelle tous les éléments relevant d'une classe sont affectés au "centre de classe". Celui-ci est défini arbitrairement, "au milieu" mais le plus courant est de le définir par la moyenne arithmétique des bornes. Ce faisant, on est ramené au cas précédent pour les effectifs et les fréquences, mais on garde l'information du "découpage" de l'espace des valeurs – ici la droite réelle pour l'instant.

Le Figure 19 montre ces projections de l'espace de description sur l'axe vertical.

Figure 19. Descriptions par intervalles



5.5 La densité de fréquence

La fréquence a l'interprétation de "nombre de fois par unité ou intervalle de temps"; donc, comme chacun sait, plus souvent cela arrive, plus c'est fréquent. Dans le contexte de la statistique, la fréquence est aussi un nombre de fois par intervalle, mais ces intervalles sont tout type de valeurs que peut porter la droite réelle, par exemple la taille de personnes, des revenus, etc. Ainsi, on dira que la fréquence des tailles de l'intervalle $[170 \text{ à } 175] \text{ cm}$ est élevée – peut-être plus que les autres.

Ces intervalles peuvent aussi être temporels, auquel cas on retrouve donc cet aspect temporel de la fréquence, qui est d'ailleurs celle de toute la théorie des signaux.

Les classes sont des sous-ensembles de la droite réelle, de sorte qu'on retrouve l'espace de réalisation en tant que famille de sous-ensembles, celle-ci étant la plus simple. Mais la fréquence n'a de sens qu'en relation avec l'étendue de l'intervalle. C'est pourquoi on définit la densité de fréquence d'une classe $c(x_\alpha)$ par le rapport de la fréquence de cette classe à son étendue :

$$d(x_\alpha) = \frac{f_\alpha}{b_\alpha - b_{\alpha-1}} = \frac{f_\alpha}{\Delta_\alpha}$$

La notion familière de densité est celle du rapport d'une masse à un volume. Cela peut se transposer ici en considérant l'intervalle de valeur comme un "volume" sur cette seule dimension, et les f_i comme des parts d'une masse totale de 1 unité. Lorsque le lieu des réalisations est temporel, comme dans les processus stochastiques, au lieu de densité on parle plutôt d'intensité, laquelle augmente si l'on a plus de réalisations par intervalle de temps.

C'est typiquement le cas des processus dits "de Poisson", qui s'intéressent au nombre de réalisations d'éventualités par intervalle temps ; une plus haute fréquence donne une plus grande intensité du processus, par exemple l'intensité du trafic automobile, l'intensité d'un écran de télévision frappé par un canon à électrons. Notons que l'égale longueur des classes n'est pas imposée ; elle n'est souhaitable qu'à des fins de comparaison des intervalles, mais justement cette relativité de la densité lui donne une vertu descriptive très générale, qui va se prolonger sur les espaces continus.

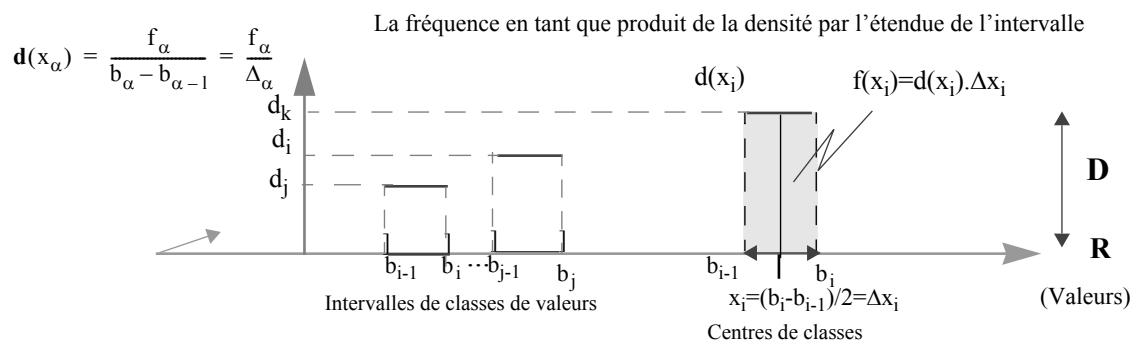
La fonction de densité de fréquence $\phi(x_\alpha)$ est définie alors pour tout x appartenant à l'intervalle d'indice α :

$$\varphi(x_\alpha) = \left(\frac{f_\alpha}{b_\alpha - b_{\alpha-1}} = \frac{f_\alpha}{\Delta_\alpha} \right) \quad \forall x \in [b_{\alpha-1}, b_\alpha], \text{ et } 0 \text{ hors de l'intervalle}$$

Les coutumes primitives font dessiner les fréquences associées aux intervalles selon des "histogrammes". Effectivement cela vient du grec (ancien ! Du nouveau, il ne vient rien) *histos* qui est un "tissu debout" et *gramme* signifie des traits graphiques, on devine ce que sera la représentation. La Figure 20 montre un diagramme de densités de fréquence de classes, où les deux premières sont sur l'étendue et la troisième sur le centre de classe.

Dans ce cas, la fréquence peut s'exprimer par le produit de la densité par la longueur de l'intervalle, ce qui donne bien la surface portée par l'intervalle. Le raisonnement peut alors être le suivant: si l'on veut faire un pronostic sur les appartenances d'éléments à tel ou tel sous-ensemble (un intervalle en est un), le pronostic de haute fréquence sera celui de plus haute densité. Dès lors, c'est l'information sur les densités relatives des sous-ensembles qui est celle qui sous-tend le pronostic d'appartenance. C'est précisément ce qui se fera avec les probabilités de réalisation d'éventualités.

Figure 20. Densités et fréquences



5.6 La fonction de densité de fréquence (ou densité "discrète")

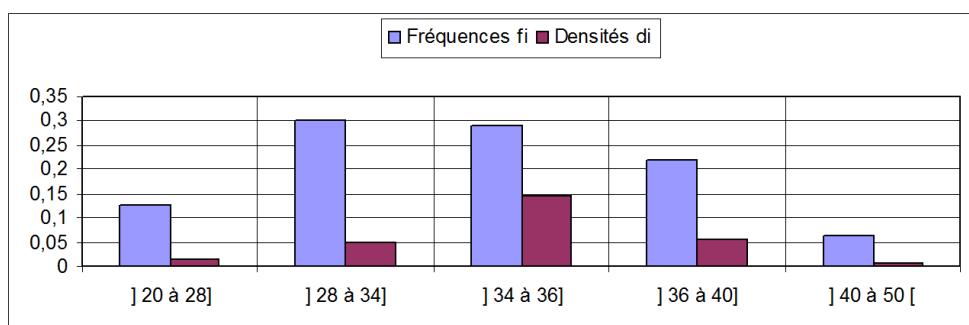
Le Tableau 1 est un exemple "concret", comme le veulent les étudiants en gestion. Il s'agit des chaussures de sport de la marque "Adadas" qui doit équiper les nonnes de l'ASLJSIGASRH (Association de Sport et Loisirs "Joie et Santé" de l'Institut de Gériatrie d'Anguille-sous-Roche). Le tableau donne une série d'observations des longueurs de pieds de 200 nonnettes par intervalles de valeurs d'inégales étendues; les centres de classes donnent les "pointures". Les lignes suivantes donnent les fréquences ainsi que les fréquences cumulées (donc jusque "1") pour montrer la fonction de répartition F.

Tableau 1. Tailles des pieds de 200 nonnes

Tailles des pieds de 200 nonnes					
Intervalles] 20 à 28]] 28 à 34]] 34 à 36]] 36 à 40]] 40 à 50 [
Etendues $b_i - b_{i-1}$	8	6	2	4	10
Effectifs n_i	25	60	58	44	13
Fréquences f_i	0,125	0,3	0,29	0,22	0,065
Fréquences cumulées F	0,125	0,425	0,715	0,935	1
Densités d_i	0,016	0,050	0,145	0,055	0,007
Pointures	A-24	B-31	C-35	D-38	E-45

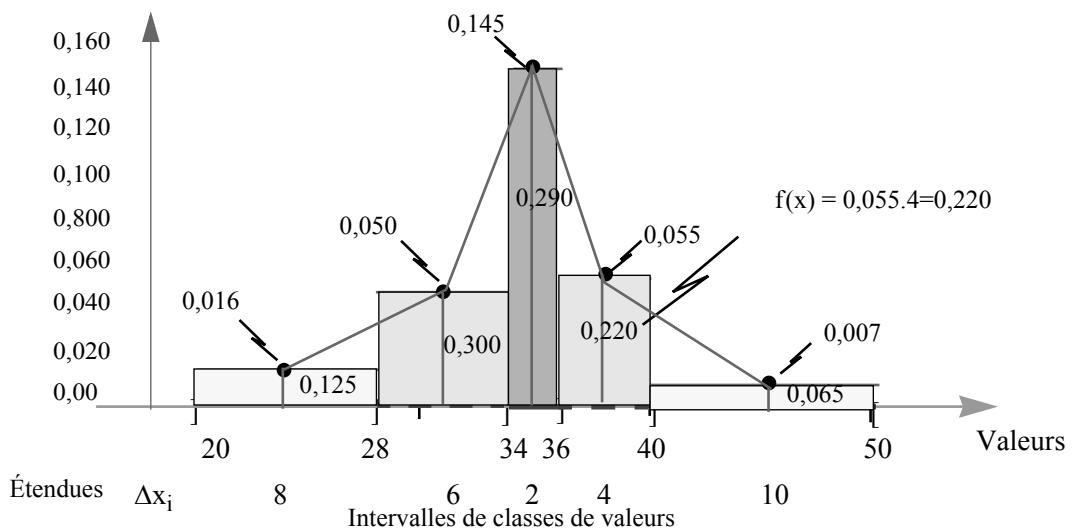
La Figure 21 correspond à ce petit exemple numérique, fait "automatiquement" sur un tableur sans exprimer les différentes étendues des classes. Ce graphique ne montre pas les relations fréquences-densités, de sorte que les utilisateurs de ce programme standard ne ressentent pas ce délicieux feeling de la statistique, l'arôme de la Figure 22.

Figure 21. Fréquences et densités de chaussures



La Figure 22, en revanche, est dessinée spécifiquement, c'est-à-dire en préservant les espacements des intervalles. Elle tente aussi d'imager les différentes densités (qui sont en ordonnée) en leur donnant des ombrages relatifs. Enfin, les points indiquant les niveaux de la densité ont été joints pour montrer l'allure de la fonction $\phi(x)$ de densité discrète.

Figure 22. Fonction de densité et aires des fréquences



Cette notion fondamentale de densité sera reprise dans cet exposé lors des expressions du potentiel, en particulier celles auxquelles sont associées des mesures de probabilité.

6 Le potentiel et ses descripteurs

6.1 Les heureux événements

Une éventualité a une énergie potentielle, qui est la propension à entrer dans le monde réel, en quittant l'espace virtuel, celui des simulations par la nature ou par la fonction de penser. L'ensemble des éventualités associées à un thème donné de la nature ou de la pensée forme alors un potentiel, désigné ici par Ω .

Du point de vue subjectif, donc du sujet confronté à l'incertitude, on ne peut confondre l'inconnu d'une part et d'autre part l'incertain tel qu'il est traité par les descriptions du hasard. L'inconnu est ici ce qu'on ne sait pas, ce dont on ne connaît pas l'existence; il n'est donc pas question que l'esprit en appréhende la potentialité, puisqu'il ne peut en simuler des réalisations et leur vraisemblance.

De plus, l'imaginaire ne fait pas non plus partie de l'incertain, car on ne s'intéresse pas à sa réalisation : ce n'est pas le propos, et ses objets ne sont pas candidats à entrer dans un "système", dans un domaine de la gestion.

Les descripteurs du potentiel ont pour mission d'exprimer les chances que certaines éventualités nous tombent dessus. Pour pouvoir exprimer, comparer, associer ces chances, il convient d'effectuer des applications sur des espaces métriques. Celles-ci définissent des mesures des propensions des éventualités à se projeter sur l'espace des réalisations.

Selon les multiples thèmes et contextes – naturels, artificiels, cosmiques, et j'en passe (à côté) – les éventualités peuvent se réunir individuelles (elles restent des singletons) ou s'assembler pour former des événements. Un exemple simple de singleton est l'une des faces d'un dé, disons le "2" ; un regroupement d'éventualités élémentaires peut être le fait d'être un trèfle (il y en a 13) dans un jeu de cartes normal.

Plus précisément, en 4.1.2, un espace des éventualités a été défini par un couple formé d'un univers Ω et une tribu S de cet univers : $\{\Omega, S\}$. Un événement (tel A_i) est donc un ensemble pouvant être au minimum une seule éventualité, au maximum tout l'univers aléatoire, et défini tout simplement comme un élément de la tribu S .

Quand un événement transforme son énergie potentielle en énergie cinétique, il se projette dans un espace spécialement conçu à cet effet, l'espace des réalisations : il est arrivé quelque chose.

Cet espace des réalisations présente évidemment une grande variété, encore un fois selon les contextes, les thèmes, les situations et tout ce qui peut se passer. Certains contextes de réalisations ont leur vocabulaire spécialisé, comme "résultat d'échantillonnage", "issue d'expérience". L'expression "expérience aléatoire" a d'ailleurs été étendue aux processus dont l'issue n'est pas maîtrisée (dépend d'un hasard), que ce processus ait été volontairement construit ou qu'il se passe naturellement.

Un premier parcours – celui que l'on suivra ici – concerne donc les applications de l'espace des événements sur celui de ses réalisations, et des formulations y prendront place sous les expressions de vraisemblance, possibilité, nécessité, probabilité. Pour pouvoir utiliser les mathématiques appliquées, les méthodes numériques, on va d'abord définir un espace qualifié de "probabilisable", à savoir celui des réalisations défini par une tribu borélienne sur \mathbf{R} , désignée ici par \mathbf{B} .

On appelle tribu borélienne sur \mathbf{R} la plus petite tribu contenant tous les intervalles ouverts de l'ensemble \mathbf{R} des nombres réels. Ceci implique que \mathbf{B} contient :

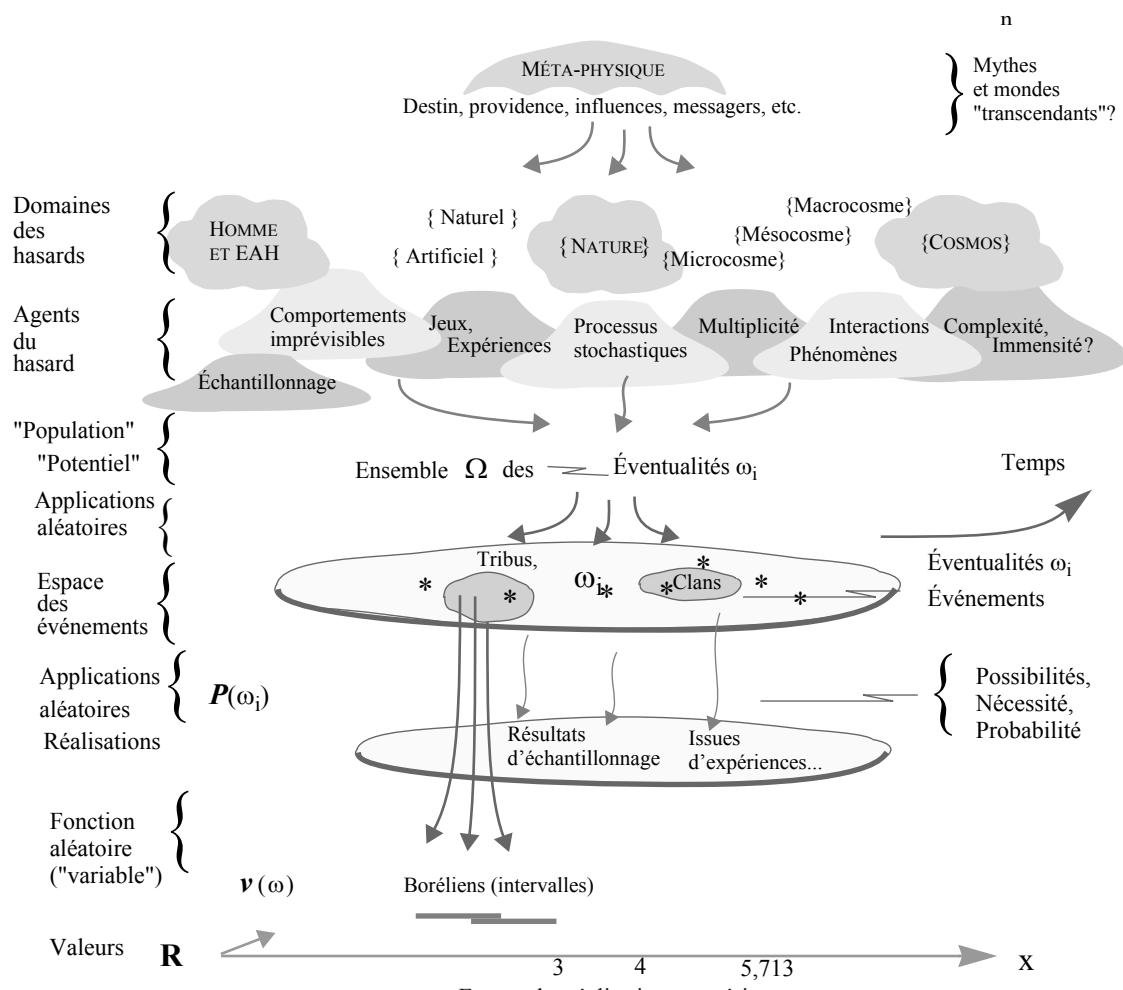
- Tous les intervalles de \mathbf{R} , y compris les intervalles réduits à un point,
- Toute réunion dénombrable d'intervalles.

La Figure 23, d'une richesse et d'une qualité exceptionnelles, représente les espaces considérés (avec un peu de recul) jusqu'à présent ; on y retrouve avec plaisir des trucs appris à l'école et oubliés, cela va de soi, depuis longtemps. Non : on devrait plutôt parler de trucs enseignés à l'école, ce qui n'implique nullement qu'ils aient été appris. Non, ce n'est pas encore juste : on devrait parler de trucs qu'on aurait dû enseigner à l'école au cas où une ministre aurait conçu d'en envisager l'opportunité politique, et non pas de choses intelligentes ou utiles.

De toute façon, faire œuvre de bienfaisance, en politique, c'est faire ce qui est bienfaisant pour sa carrière et sa réélection.

Donc, on disait, la Figure 23, une architecture classique dans son élégance, et moderne dans ses fonctionnalités.

Figure 23. Les ensembles hasardeux



6.2 Des mesures de "champ d'incertitude"?

L'information est définie ici comme absolue lorsqu'elle conduit à une connaissance non-reversible ; l'état qui en résulte pour le sujet est la certitude. Cette issue peut s'obtenir d'un seul coup, par exemple par une constatation incontestable, ou au terme d'un processus d'apprentissage convergent in fine vers un attracteur stable.

Une mesure monotone devrait pouvoir accompagner ce processus, et montrer le sens de la progression de l'apprentissage s'il y en a un, et rendre compte de son aboutissement. Ainsi, on parle de l'entropie pour suivre l'incertitude, et atteindre une probabilité égale à un est un exemple de mesure d'aboutissement.

Ce processus d'apprentissage se déroule dans le champ d'incertitude du sujet ; celui-ci est ouvert à des entrées d'information issue de ses "environnements".

Le sujet dispose d'un processeur des signaux pour modifier les relations dans ce champ ; tandis qu'il apprend, qu'il change ses degrés de croyance, le champ lui-même se modifie, son intensité varie dans les différentes directions de l'incertitude. Un champ est en effet un objet épistémique, c'est-à-dire relatif à la configuration du savoir à une période donnée et à un contexte.

Comme l'univers des possibles est personnel, propre au sujet en raison de ses capacités de simulation, on peut lui associer un apport et une évolution de l'information subjective. Les différentes descriptions de l'incertitude du sujet seront dès lors des descriptions de son état du point de vue de la connaissance.

Celui-ci peut-être naturel (par exemple la crédulité), ou aidé par l'artificiel (par exemple une logique des prédictats). Lorsque la progression de connaissance du sujet est un processus sous contrôle, qui le conduit en quelque sorte vers une éventuelle convergence, on parle de système d'apprentissage. Pour qu'il soit encore plus systémique, on devrait pouvoir en exprimer un critère de performance ; dans le cadre de la gestion, un critère valable à cet égard est la probabilité d'inversion de la décision, explicitée dans l'exposé « Vers le Décideur artificiel » situé dans le Tome d'Ouessant.

L'expression même de l'incertitude peut prendre différentes formes, mais elles relèvent de cette même idée que, vis-à-vis d'un sous-ensemble (un élément de la tribu) S de l'univers Ω des possibles, le processeur de l'agent exerce une sorte de fonction $v(S)$ qui exprime le degré de confiance qu'il a dans le fait que l'événement S est celui qui prévaut, ou qui va se réaliser si la situation est prédictive. Ce $v(S)$ est écrit pour vraisemblance, concept qui peut envelopper différentes expressions de degrés de croyance.

Mais peut-on imaginer une mesure de ce "champ", où régnerait, si l'on en croit les troubadours du hasard plutôt que les compositeurs classiques, un potentiel et des ondes ? S'il est vrai que le hasard et l'incertitude ont semé des terreurs et des fantasmes, cela ne justifie pas d'accepter des expressions qui sentent l'analogie à plein nez. En effet, les mots de potentiel etc., inspirent de chercher des unités et des instruments de mesure en physique électromagnétique. Ce sera encore une illusion perdue de la métaphore, de l'analogie.

Les unités candidates à cette fin sont celles de l'induction, de flux d'induction, d'intensité de champ magnétique ; elles demandent toutefois toujours qu'une au moins des unités spécifiques de l'électro-magnétisme soit présente. En effet, on pourrait être tenté de faire appel à l'un des deux groupes de mesures électriques, caractérisés comme suit :

- La représentation de l'électricité et du magnétisme fondée sur l'approche de l'action à distance ; les unités sont alors classiques du système LMT, c'est-à-dire Longueur, Masse, Temps. Elles se rapportent à l'époque où l'on ne disposait que de procédés mécaniques pour mesurer les grandeurs électriques et électromécaniques ;

- La représentation fondée sur la théorie des champs est associé au développement de la théorie électrodynamique des champs et des mesures spécifiques à l'électricité et au magnétisme: il lui faut alors quatre dimensions.

Le nombre de "dimensions" nécessaires à la modélisation – donc à la représentation avec des mesures – d'un système physique varie avec la nature de celui-ci:

- En cinématique: 2;
- En mécanique: 3;
- En thermodynamique, photométrie, électrodynamique: 4 – nous y voilà donc.

Pour la représentation fondée sur la théorie des champs, il faut par exemple compléter le LMT par une des unités suivantes si on veut rester en "SI", le Système International d'unités adopté en 1962:

- L'ampère: intensité de courant électrique;
- Le degré Kelvin: température thermodynamique;
- Le candela: intensité lumineuse.

La nature d'un système physique est donc définie comme différente d'une autre si le nombre de dimensions de la représentation est différent. Ces dimensions sont typifiées comme des qualités fondamentales du système; or, à celles-ci sont associées des unités fondamentales, lesquelles sont qualitativement spécifiques à ces dimensions.

Le passage d'un système d'une nature à l'autre se fait en substituant à une dimension (donc une qualité fondamentale) du système-source la formule dimensionnelle de la qualité dérivée correspondante dans le système "destination". Les unités dérivées peuvent alors acquérir un nom propre dans le nouveau système: ainsi la force LMT a pour unité la dyne en cgs (cm g s^{-2}) ou le newton en (m kg s^{-2}) en système MKS etc.

Si le passage d'unités est possible d'un domaine (par exemple l'électro-magnétisme) à un autre qui en est privé (par exemple l'incertitude), alors les concepts utilisés deviennent mieux qu'analogues, ils deviennent transposables. Ici on a osé parler en "intensité de champ" d'incertitude (et d'information), et en "potentiel" (différencié) de réalisation d'événements.

Voyons ces deux seuls exemples, sinon cela va chercher loin. On aura besoin de:

- R : la résistance électrique (en ohms); I : Intensité ("courant", en ampères);
- Ω : la résistance en Ohms; U : tension en volts;
- LMT: la Longueur-Masse-Temps (cm gr sec);
- ϵ_0 : la constante de champ électrique, soit $\text{L}^{-1}\text{TU}^{-1}\text{I}$, en farad/mètre;
- μ_0 : la constante du champ magnétique; pour les amateurs, elle est définie conventionnellement par: $\mu_0 = 10^{-7} 4 \pi \Omega \text{ s m}^{-1}$

Alors, l'unité d'intensité de champ électrique (dans le système LMT μ_0) est:

$$E(\text{UEM}) \sim \text{L}^{1/2} \text{ M}^{1/2} \text{ T}^{-2} \mu_0^{1/2}$$

Dans le même système d'unités (LMT μ_0), on peut mesurer à présent ce fameux potentiel, à savoir ici le potentiel vecteur magnétique A :

$$A(UEM) \sim L^{1/2} M^{1/2} T^{-1} \mu_0^{1/2}$$

Il correspond au weber par mètre, (Wb m⁻¹) où le weber est l'unité de flux d'induction magnétique.

Dans ces deux cas, on a effectivement eu besoin d'unités spécifiques au domaine de l'électromagnétisme pour construire les unités adéquates. De plus, pour effectuer les mesures électriques, il est besoin de combiner deux procédés, impliquant le potentiel ou la charge, le courant ou l'induction; comme on n'a pas de "qualité dérivée" correspondante dans le domaine du hasard et des éventualités qui entourent les EAH, on n'a pas de possibilité rigoureuse de transposition.

Cela a pour implication que l'on ne pourra pas construire une formulation de champ de l'incertitude, du temps ou de l'information en tant que champs de tension; on ne pourra pas non plus formuler la propension à se réaliser par la mesure d'un potentiel de réalisation, en utilisant les unités fondamentales issues du domaine où elles sont valides pour mesurer ces notions, à savoir la physique et l'électromagnétisme. En conclusion, l'analogie formelle est abusive, car elle n'est pas digne de transposition dans les formalisations; elle doit se limiter, comme l'a expliqué l'exposé sur «La Systémographie», à aider le discours, c'est-à-dire à être discursive.

Il faut alors se résoudre à faire comme Monsieur Tout-le-monde, et faire appel à ce que les auteurs raisonnables ont construit comme mesures qui permettent des formulations et des applications numérisées de certaines appréhensions de hasards.

6.3 La possibilité

6.3.1 Le doute?

Le doute d'un agent augmente avec son potentiel de perception et de traitement de signaux... son intelligence? Un exemple de taille est celui de René DESCARTES, qui a même pris le doute comme point de départ pour devenir intelligent. Dans la vie et le langage courant, on entend d'ailleurs beaucoup plus d'affirmations péremptoires, non-conditionnelles, chez les gens peu cultivés: les richesses du langage du doute et de la pensée sont souvent parallèles. Il résulte de cet argument que l'agent plus lucide entrevoit plus de possibilités, s'adresse à un champ plus complexe, et de plus de potentiel d'incertitude.

L'assertion formelle de probabilité est arbitraire, ou subjective, lorsque son évaluation n'a pas de référence et ce, pour plusieurs raisons:

- Une première raison est qu'il n'y a pas de base de fréquence.

Supposons qu'on veuille assurer la bourgade d'ANGUILLE-SOUS-ROCHE contre les tremblements de terre – qu'on y a jamais ressenti de mémoire d'âne: quelle prime verser, sur base de quelle probabilité?

Soit à présent qu'on assure la ville de POMPÉI, déjà toute cassée mais sans tremblette depuis deux mille ans : la dernière éruption en fut cependant tellement spectaculaire que l'on peut s'attendre à une assertion de probabilité subjective plus élevée ;

- Une autre bonne question est : dans quelle mesure une événement est-il "aléatoire" ? Peut-on le dire loin du déterminisme, ou hors de la maîtrise d'un opérateur ? On devrait imaginer une distance par rapport au déterminé et certain, comme on mesure le degré de flou des sous-ensembles par leur distance de HAMMING par rapport aux ensembles vulgaires.

6.3.2 Une mesure de possibilité

Quand une épouse dit qu'elle est sûre d'elle, c'est qu'elle est sûre d'elle ; quand elle dit qu'elle est sûre de son mari, c'est qu'elle est bien sûre d'elle. De telles situations sont parmi les clientes de la théorie des possibilités.

Soit Ω un ensemble de référence, et soit S l'ensemble des sous-ensembles ordinaires de Ω ; un événement est un élément de S . Dans le chef d'un agent, un événement est considéré comme impossible s'il est totalement exclu qu'il puisse avoir lieu ; sinon, il est simplement possible, et dès lors il peut prendre place parmi ceux qui peuvent agiter un processus de conception des choix puis de décision. Dans cette perspective, il est fort utile de disposer d'un indicateur du degré auquel tel événement est considéré comme possible, ce qui est l'objet d'une mesure de possibilité.

Un indicateur qui complète cette évaluation subjective devrait renseigner sur la mesure dans laquelle un événement est certain. Ce deuxième aspect est pris en charge par les mesures de nécessité.

Une mesure de possibilité Π est une fonction définie sur l'ensemble S , prenant ses valeurs dans l'intervalle fermé $[0; 1]$, et spécifiée par les propriétés suivantes :

$$(1) \quad \Pi(\emptyset) = 0; \quad \Pi(\Omega) = 1$$

Pour toute famille de sous-ensembles $S_1 \in S, \dots, S_j \in S, \dots$:

$$(2) \quad \Pi(\cup_{j=1,2,\dots} S_j) = \vee_{j=1,2,\dots} \Pi(S_j)$$

La notation \vee_j indique le supremum des valeurs figurant dans la liste. En l'occurrence, la propriété (2) signifie que la mesure de possibilité associée à une union d'ensembles est la plus élevée des possibilités de chacun d'entre eux.

Bienentendu, la valeur $\Pi(S_j) = 0$ implique que S_j est impossible, et $\Pi(S_j) = 1$ implique que S_j est "tout-à-fait possible" ce qui est une expression peu heureuse, disant peut-être que rien ne paraît interdire d'office qu'il se réalise.

Reprenant le cas angoissant de la course, supposons que Ω soit l'ensemble des moments d'arrivée du coureur BINAMÉ par rapport au peloton, les sous-ensembles S_j d'événements considérés étant A : "devant", B : "dedans", C : "derrière". Le A serait obtenu par exemple avec le dopage, B avec le placebo (qui est neutre quant à la performance), et C par un crash dans les betteraves.

Envisageons d'abord deux des S_j , disons A et B. Dans ce cas, la règle (2) ci-dessus s'écrit:

$$(3) \quad \forall (A, B) \in S^2, \quad \Pi(A \cup B) = \vee \{ \Pi(A), \Pi(B) \}$$

La lecture de (3) est que la possibilité de réalisation de l'un ou l'autre des deux événements (l'union) est le maximum (\vee) des mesures de possibilité de chacun d'eux. Dès lors :

Les mesures de possibilité ne sont pas additives.

Soient les évaluations respectives $\Pi(A)=0,4$ et $\Pi(B)=0,75$. Clairement, la réunion des événements "devant" et "dedans" a une possibilité de 0,75. On pourrait même évaluer que l'arrivée B dans le peloton soit "tout-à-fait possible", soit $\Pi(B)=1$, et les deux autres 0,75 et 0,4 ou ce qu'on veut de positif: il est clair que, même en ce cas d'exclusivité, il n'y a pas additivité (et d'ailleurs pas du tout d'indépendance).

Les propriétés (1) et (2) n'informent pas sur l'intersection, laquelle est définie par (4):

$$(4) \quad \forall (A, B) \in S^2, \quad \Pi(A \cap B) \leq \wedge \{ \Pi(A), \Pi(B) \}$$

La possibilité de l'intersection de parties de S est donc majorée par le plus petit des coefficients de possibilité associés à chacune des parties (" \wedge " indique le minimum).

(5) Un corollaire de cette propriété est le suivant:

Deux (ou plusieurs) événements peuvent être possibles ($\Pi(S_j) > 0$) mais leur occurrence conjointe impossible: $\Pi(S_i \cap S_j) = 0$.

Ainsi, il est possible que le coureur arrive devant, ou derrière, mais pas les deux simultanément, ce dont on se doutait un peu.

A présent, la propriété (5) donne à la mesure de possibilité la propriété de mesure floue :

$$(6) \quad \text{Si } A \text{ et } B \in S, \text{ et } A \supseteq B, \quad \text{alors } \Pi(A) \geq \Pi(B)$$

Les mesures de possibilité sont monotones par rapport à l'inclusion.

La possibilité de l'ensemble englobant est donc toujours supérieure ou égale à celle de toute partie. Il va de soi, dès lors, que $\Pi(A \cup B) \geq \Pi(A)$.

Enfin, les relations avec les compléments sont fascinantes. Qualifions de "contraire" l'événement A^c qui est le complément de A dans S. Dans ce cas :

$$(7) \quad \forall (A) \in S, \quad \text{soit } \Pi(A) = 1, \text{ soit } \Pi(A^c) = 1$$

Dans un ensemble S de possibilité, un événement ou son contraire est tout-à-fait possible.

La relation implique qu'une assertion de possibilité inférieure à 1 sur un événement implique que son contraire est "tout-à-fait possible".

Ainsi, si l'assertion "le coureur Z va gagner la course", soit $\Pi(A)$, est <1, alors l'assertion "Z perd la course" (c'est-à-dire "ne la gagne pas", donc B ou C) est "tout-à-fait possible".

6.3.3 Distribution de possibilités

Le client de l'information est un être hésitant qui souhaiterait associer une expression de possibilité et vraisemblance aux différents états d'une liste gérée par la Providence – c'est-à-dire le processus supérieur qui gère les circonstances dans lesquelles se déroulent les processus réels de la vie. Les mathématiciens jouent tout le temps avec leurs parties, leurs boréliens et autres collections inextricables de sous-ensembles. Le client, quant à soi, préfère se donner une ensemble-liste d'éléments disons S , de cardinal M et disposer d'une mesure de leurs possibilités.

En effet, si le cardinal de S est m , il y a 2^m sous-ensembles, donc parties définies, ce qui revient à devoir définir 2^m coefficients de possibilité pour faire le tour de la question. Une distribution de possibilité simplifie cela en associant à chaque élément θ_j une mesure $\Pi(\theta_j)$. Elle répond alors à la définition suivante :

Une distribution de possibilité π est une fonction définie sur S , appliquant ses valeurs dans $[0, 1]$, et satisfaisant la condition de normalisation, à savoir :

$$(8) \quad \sup_{\theta \in S} \pi(\theta) = 1$$

Si on veut associer une possibilité à tout ensemble de S , il suffit de réunir des θ_j pour former une partie S_j et appliquer la propriété de la possibilité de l'union, à savoir :

$$(9) \quad \forall (S_j) \in S, \quad \Pi(S_j) = \sup_{\theta \in S} \pi(\theta)$$

C'est donc une mesure monotone vis-à-vis de l'inclusion, et dès lors elle répond correctement à la condition des mesures floues. Évidemment, la connaissance de la distribution permet aussi d'associer une mesure de possibilité à chaque élément. Une exploitation logique et facile de (9), mais qui n'apparaît pas tout de suite via sa formulation, est :

- Soit S un sous-ensemble de Ω ,
- et soit qu'il soit tout-à-fait possible que l'élément θ appartienne à S , donc :

$$\Pi(\theta)=1 \quad \forall (\theta) \in S$$
- et soit qu'il soit impossible qu'il n'y appartienne pas, donc :

$$\Pi(\theta)=0 \quad \forall (\theta) \in S^c \text{ dans } \Omega.$$
- Soit d'autres sous-ensembles de Ω , disons A, B, \dots ;

Sur la base de la connaissance du sous-ensemble S auquel θ appartient, il serait intéressant de faire une assertion de possibilité que θ appartienne à A, B, \dots dans Ω . C'est vite fait, car la simple logique indique que dans ces conditions :

$$\Pi(A) = 1 \quad \text{si } A \cap S \neq \emptyset.$$

Si l'intersection est vide (pas d'élément commun avec le "tout-à-fait possible"), alors $\Pi(A) = 0$. Lorsqu'on applique cela aux performances cyclistes de BINAMÉ Joseph, il boude pendant des semaines. En effet, soient les possibilités de retard suivantes :

- Impossible de gagner plus de 10 minutes : $P\{]-..., -10]\}=0$;
- Tout-à-fait possible de gagner jusque 5', jusqu'à en perdre 15 : $P\{[-5, +10]\}=1$;
- Impossible d'avoir un retard $\geq 20'$, car c'est le camion-balai et la face perdue. Donc :

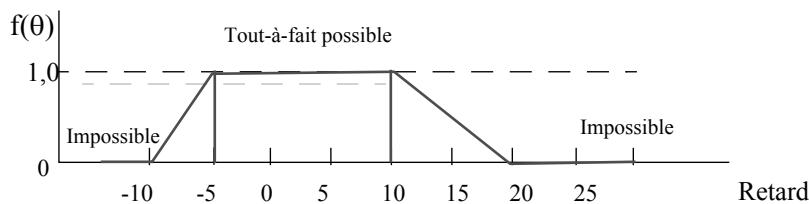
$$P\{[20, \dots]\}=0.$$

La Figure 24 montre la distribution de possibilité du retard, sans donner de détails sur les intervalles non-cités ; les segments en pente illustrent la monotonie de la fonction sur les intervalles inclus. Cette Figure 24 montre aussi la tête que peut tirer un nombre flou.

L'attention des spectateurs est attirée sur le fait que les événements "performance" et "crash" sont tous deux "tout-à-fait possibles", alors qu'ils ne se recouvrent pas. Ceci confirme que des événements complémentaires peuvent être absolument possibles, ce que n'expriment pas les axiomes de probabilité.

On construit aussi sur les mesures de possibilité des mesures composées et dérivées, telles que des distributions de possibilités conjointes, marginales et conditionnelles. Avec regrets, ces aspects resteront confiés aux maisons d'édition sérieuses, car ils ne s'imposent pas dans le parcours présent.

Figure 24. Distribution de possibilité du retard



6.4 Une mesure de nécessité

De ce qui précède, on comprend que le degré de possibilité, et même le tout-à-fait possible, ne renseigne pas sur le degré de certitude d'un état, événement – défini par le degré d'appartenance à un sous-ensemble. On peut donc avancer qu'un événement (S) et son contraire (S^c , son complément) sont tous deux possibles, ce qui laisse perplexe, ou encore constater que S^c est impossible, sans pour autant dire grand-chose sur le degré de certitude – la conviction ? – de réalisation de S .

Les mesures de nécessité complètent le tableau. Elles sont fondées sur la possibilité – on les qualifie parfois de duals de la possibilité – et fournissent un coefficient lui aussi défini sur l'intervalle fermé $[0, 1]$.

Plus précisément, une mesure de nécessité est une fonction définie sur l'ensemble des parties de Ω , s'appliquant sur le fermé $[0, 1]$, et spécifiée par les propriétés suivantes :

- (10) $N(\emptyset) = 0$; $N(\Omega) = 1$;
- (11) $\forall (S_j, \dots, S_k) \in S, N(\bigcap_{j,k} \dots S_j) = \inf_{j,k, \dots} N(S_j)$

On lit donc en (11) que la nécessité de l'intersection est le minimum des nécessités respectives des ensembles concernés.

Les propriétés formelles sont du même acabit que celles des mesures de possibilités :

- Si l'ensemble A inclut $B \in S$, la nécessité de $A = 1$, sinon elle est nulle ;
- La nécessité de l'union de A et B est : $N(A \cup B) \geq \max \{N(A), N(B)\}$.

6.5 Incertitude et dispersion

Les mesures d'incertitude et d'information peuvent être construites à partir d'une expérience aléatoire binaire, c'est-à-dire décrivant le fait qu'un événement A peut avoir lieu ou ne pas avoir lieu : il n'y a donc que deux issues possibles. Soit $P\{A\}$ la probabilité a priori de réalisation de A dans le chef du "client" de l'information (disons un récepteur, un décideur). Le montant d'information contenue dans un message signalant quel événement A s'est réalisé (la probabilité associée à cet événement devient égale à 1) est une fonction monotone croissante de l'écart entre $P\{A\}$ et 1.

Cette fonction, développée par SHANNON, est $I_g(1/P)$, où I_g sera écrit pour le logarithme en base 2. Ce montant d'information peut représenter l'effet de surprise du client qui apprend la réalisation de l'événement. Cette surprise est d'autant plus grande que sa probabilité a priori P était faible, donc que $1/P$ était plus élevée.

Plus généralement, soit une expérience aléatoire α dont les k issues sont exclusives et exhaustives, et forment donc la liste A_1, \dots, A_k , et soit $P\{A_i\}$ la probabilité associée à l'issue A_i . Il est clair dès lors que :

$$\sum_i P\{A_i\} = 1$$

Chaque issue possible A_i contribue à l'incertitude à concurrence de $1/P_i$, et a une probabilité d'occurrence de P_i . Dès lors, l'espérance mathématique de l'incertitude contenue dans cette situation aléatoire α peut s'exprimer par (en logarithme de base 2, écrit I_g) :

$$H(\alpha) = \sum_i P\{A_i\} I_g[1/P\{A_i\}]$$

Cette grandeur est l'entropie (discrète) de cette expérience. Elle mesure l'espérance mathématique de l'information contenue dans un message qui signalerait au client qu'une de ces issues s'est réalisée. Ses propriétés le plus utiles sont les suivantes :

- L'entropie (discrète) est maximale si toutes les probabilités sont égales, donc que la diffusion des événements dans l'espace probabiliste est maximale, disons comme celle des molécules d'un gaz parfait dans un milieu clos ;
- L'entropie tend vers 0 si la probabilité d'une des issues tend vers 1 (devient "certaine"), les autres devenant nulles ;
- Elle est dès lors croissante avec le nombre n d'issues (il y a plus de possibilités) et avec la diffusion (les issues sont plus équiprobabiles). C'est en effet n l'nombre d'issues qui correspondra au nombre d'états potentiels d'un système dans la version thermodynamique, source de cette théorie.

On met déjà cela dans un coin de la mémoire du Lecteur, car cela explicitera bien des choses à partir de la section 3.

L'intervalle est donc : $0 \leq H(\alpha) \leq I_g(n)$

L'entropie est mise en œuvre dans l'exposé sur « L'Information et sa thermodynamique », et des probabilités conditionnelles sont servies par l'exposé « Vers le Décideur artificiel ».

Probabilité

Les termes de "probable" et "probabilité" sont génériques dans le langage courant. Le plus répandu est de dire "c'est probable" lorsque, dans le chef du répondant, il y a plus de "chances" que cela se produise plutôt que cela ne se produise pas. BINAMÉ va-t-il gagner la course cycliste ? La réponse "c'est probable" le place quasi comme favori. Le "c'est peu probable", en revanche, reflète plutôt une impossibilité tout empreinte de courtoisie. Ainsi en va-t-il du crédible, du possible, du vraisemblable, etc.

Ces mots ont une signification distincte et bien établie dans la langue française, et leur emploi judicieux est preuve de raffinement. Mais, comme toujours, le professionnel se dénote par son langââge précis et châââtié. Cependant, le problème est double :

- D'une part, les mêmes mots utilisés par des sujets différents n'ont pas nécessairement la même signification, et certes pas la même intensité ;
- D'autre part, les mots courants, on peut les dire mais pas en faire des formulations. On ne peut pas sans ambiguïté "calculer ses chances", ou les combiner, les additionner, bref, faire des descriptions stables et des constructions – des systèmes ? – non ambigus avec les machins du hasard.

C'est une des bonnes raisons pour lesquelles les notions avec lesquelles on souhaite faire de telles constructions cohérentes reposent sur une axiomatique. Il en est typiquement ainsi de la probabilité. Cette notion traduit la propension à se réaliser. On ressent bien que ce champ de tension a des intensités variées, en ce sens que certaines éventualités ont une propension plus forte que d'autres à se réaliser.

Cette intensité de la propension peut :

- Soit relever du chef de celui qui simule l'univers aléatoire, en parcourant les possibilités par la pensée, – on la dit alors subjective ;
- Soit être une propriété inhérente au milieu ou au phénomène auquel cette propension est associée – on la dit alors objective.

Bien sûr, certaines propriétés d'éventualités favorisent leur propension à se réaliser :

- Dans un univers aléatoire fini, si un événement est défini par un effectif relativement plus élevé, cet événement a plus de "chances" de se réaliser, puisqu'il a relativement plus de candidats ;
- La répétition est aussi un facteur augmentant les chances de réalisation, dans un contexte de processus temporels ;
- Certains événements ont une plus forte propension à se réaliser parce qu'un ensemble de facteurs sous-jacents y contribuent, mais ces facteurs ne sont pas élucidés ;
- Il y a aussi le phénomène de non-réalisation ; ainsi, dans le domaine génétique, toutes les prédispositions peuvent être présentes pour retrouver une certaine caractéristique héréditaire, mais celle-ci n'apparaît pas, parce qu'une faible perturbation, due au polymorphisme, a fait une dérivation vers une variété. L'attendu, "le plus probable" n'est alors pas réalisé ;

- Enfin, il y a de fortes intensités de la propension à se réaliser qui sont simplement... comme cela. Il y a des gens qui semblent gagner plus souvent que d'autres aux tombolas, aux distributions de pur hasard, donc qui sont comblés ou frappés par le sort. Il y a aussi cette infinité de processus naturels qui font apparaître certaines de leurs issues, ranimant par leur inattendu les jours sans surprise, les pluvieuses soirées de novembre, les siestes sans amour et les repas sans vin.

Mais l'inexorable axiomatic est là pour tempérer les espoirs de ceux qui attendent des cadeaux du hasard. Dès qu'elle lui apprendra qu'il a moins de chances de gagner un million au Loto que de s'éclater la tronche en avion, le péquenot commencera à se méfier des caprices de l'aléatoire comme on se méfiait des coups du sort jetés par les sorcières.

Bon, on rentre dedans. Une mesure de probabilité est associée à un espace $\{\Omega, \mathbf{T}\}$. Elle est définie comme une application répondant aux conditions suivantes :

- Ayant pour ensemble de définition la tribu \mathbf{T} ;
- Ayant pour ensemble d'arrivée l'ensemble \mathbf{R} des réels;
- Satisfaisant aux axiomes suivants :

(12) La probabilité associée à un événement quelconque est positive ou nulle :

$$\forall A \in \mathbf{T}, \quad P\{A\} \geq 0$$

(13) La probabilité de l'événement certain est égale à 1 :

$$P\{\Omega\} = 1 \text{ donc } \sum_{i=1}^n P\{A_i\} = 1$$

(14) Pour toute suite d'éléments de \mathbf{T} disjoints deux à deux, on a :

$$P\left\{\bigcup A_i\right\} = \lim_{n \rightarrow \infty} \sum_i^n p[A_i]$$

Lorsque l'univers des éventualités Ω est fini – la tribu est un clan – l'axiome (14) est plus facile à lire par son implication, pour deux événements A et B , que :

$$(15) \quad P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}$$

Les développements faits à partir des axiomes de définition construisent évidemment la théorie des probabilités, dont beaucoup d'exemplaires ont été tirés chez des éditeurs sérieux.

Quelques implications initiales en sont écrites puis dessinées ci-dessous ; comme on atteint alors les limites fixées pour la promenade présente, elles ne servent que de repères pour suivre les Voies des maîtres errant sur les sentiers probabilistes.

La principale est celle de l'inclusion :

$$(16) \quad A \subseteq B \Rightarrow P\{B - A\} = P\{B\} - P\{BA\} \text{ et } P\{A\} \leq P\{B\}$$

Du fait que $\emptyset \subseteq A \subseteq \Omega$, ceci implique, que $0 \leq P\{A\} \leq 1$

Deux des mesures qui ont le plus de succès sont les probabilités conditionnelles et celles du produit. La probabilité de B conditionnelle à A ($\neq \emptyset$) et une application définie par:

$$(17) \quad P\{B|A\} = \frac{P\{A \cap B\}}{P\{A\}} \quad \forall (B \in \mathbf{T})$$

Une autre lecture de (17) définit la probabilité de l'intersection de deux ensembles A et B :

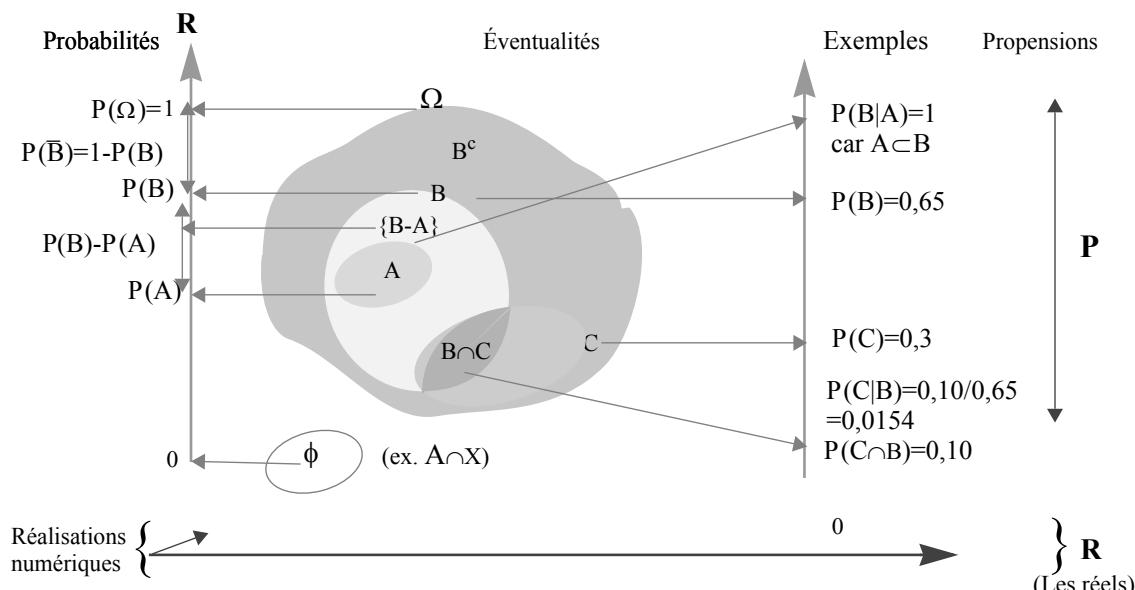
$$(18) \quad P\{A \cap B\} = P\{A\}P\{B|A\} = P\{B\}P\{A|B\}$$

L'expression (18) est une des clés du développement de la théorie de l'information (B "informe" sur A en modifiant la probabilité de réalisation de A); elle est située dans l'exposé sur «L'Information et sa thermodynamique».

D'autre part, (18) est aussi la clé du théorème de Bayes, qui fait un tabac dans l'exposé sur la décision dans l'incertitude, du Tome d'Ouessant. On est donc prié de ne pas effacer cette expression; il vaut même mieux la tagger sur le mur du voisin.

La Figure 25 tente quelques représentations des expressions (12) à (18); il y est bien spécifié que les surfaces des ensembles et la probabilité associée n'est en principe qu'une illusion d'optique, la clé étant la densité. Mais une image c'est cela: elle n'est pas barométrique, c'est encore de la topologie flasque des sciences molles. Le vrai graphique devrait être calibré, et l'épure de l'ingénieur doit respecter la géométrie.

Figure 25. Descripteurs de probabilité



! Il n'y a aucune relation formelle entre les proportions "visuelles" des ensembles et leur probabilité de réalisation; cela dépend surtout de leurs densités en événements et de leurs propensions à se réaliser. Ce piège graphique est néanmoins difficile à contourner.

6.6 L'apport mutuel des mesures d'incertitude

Comme certains ménages, les mesures de probabilité, de possibilité, de nécessité, de vraisemblance et de dispersion ont des relations complémentaires. Celles-ci peuvent être synthétisées comme suit:

- Le point important qu'elles ont en commun, et qui les spécifie comme mesures floues, est la monotonie par rapport à l'inclusion ;
- La certitude $C(\theta)$ dit s'il est vrai ou faux que θ s'est réalisé. Une version subjective de la certitude est qu'on en est absolument certain si, aux yeux de l'agent, rien ne peut empêcher qu'il se produise (et que l'information qui l'affirme soit vraie) ;
- En théorie classique des ensembles, l'appartenance de θ à S , $\mu_S(\theta)$, est bien nette, donnée par une application sur l'ensemble $\{0; 1\}$ – donc soit appartient à S , soit pas ;
- En théorie des sous-ensembles flous, l'appartenance de θ à S est graduée, et exprimée par la fonction d'appartenance $\mu_S(\theta)$ qui s'applique sur tout l'intervalle fermé $[0; 1]$;
- La probabilité $P\{\theta\}$ est une mesure associée à la propension à se réaliser d'un événement, et s'applique sur l'intervalle $[0; 1]$, avec 0 pour impossible, et 1 pour certain :
 - $P\{S\} = 1$
 - $\forall A$ et $\forall B \in S$ tels que $A \cap B = \emptyset$, $P\{A \cup B\} = P\{A\} + P\{B\}$
- En théorie des possibilités, la mesure $\Pi(\theta)$ évalue le degré avec lequel l'occurrence de l'événement θ est possible. Cependant, elle ne donne pas une direction de la détermination, puisque si un événement S_j (partie de S) est possible, son contraire (S_j^c , son complément dans S) peut aussi être possible ;
- La nécessité donne un complément d'indication sur le degré de détermination dans un contexte de possibilité. L'idée est qu'un événement est d'autant plus certain que la possibilité que son contraire se réalise est faible. Une expression qui convient à cela est alors, en gardant comme "S" l'ensemble des parties de Ω ;

$$(19) \quad \forall S_k \in S, \quad N(S_k) = 1 - \Pi(S_k)$$

Si la relation est à établir à partir de la distribution de possibilité $\pi(\theta)$, (19) devient :

$$(20) \quad \forall S_k \in S, N(S_k) = NF_{q \in S} [1 - \pi(\theta)].$$

On sent que la possibilité enseigné n'est sur la vraisemblance, laquelle demande qu'on puisse observer la réalisation (constater le "vrai"), ni sur la certitude.

6.7 Variable aléatoire

6.7.1 L'aléa numérique

L'interprétation vulgaire de l'expression "variable aléatoire" est celle d'une variable définie sur un ensemble de valeurs, mais telle qu'on ne puisse pas contrôler sa réalisation ; elle se comporte comme si les valeurs qu'elle présente au monde réel, à l'observateur, étaient engendrées par un processus imprévisible, désordonné ou non élucidé.

C'est une acceptation suffisante pour parler d'éventualités définies sur un clan, et traiter de problèmes impliquant des listes finies, des "cas favorables et cas possibles", des proportions observables, les comptages et leur combinatoire. Elle admet la définition d'une application de l'espace des éventualités sur la droite des nombres réels, cette dernière portant la "valeur" de la variable aléatoire, qui par hasard se révèle être celle-là.

On peut écrire cela $X(\omega_i) = x_i$, où $X(\cdot)$ désigne l'application, et x_i désigne une valeur réalisée lorsque ω_i se produit.

Cette acceptation a dû être rapidement généralisée et spécifiée pour pouvoir construire le calcul des probabilités, dont l'apport principal est celui de pouvoir traiter de compositions de phénomènes et de variables, de sommes, de produits, etc.

Le terme de variable aléatoire pourrait alors être correctement remplacé par celui d'application mesurable, ou, plus explicite encore, d'aléa numérique, disons $v(\omega)$:

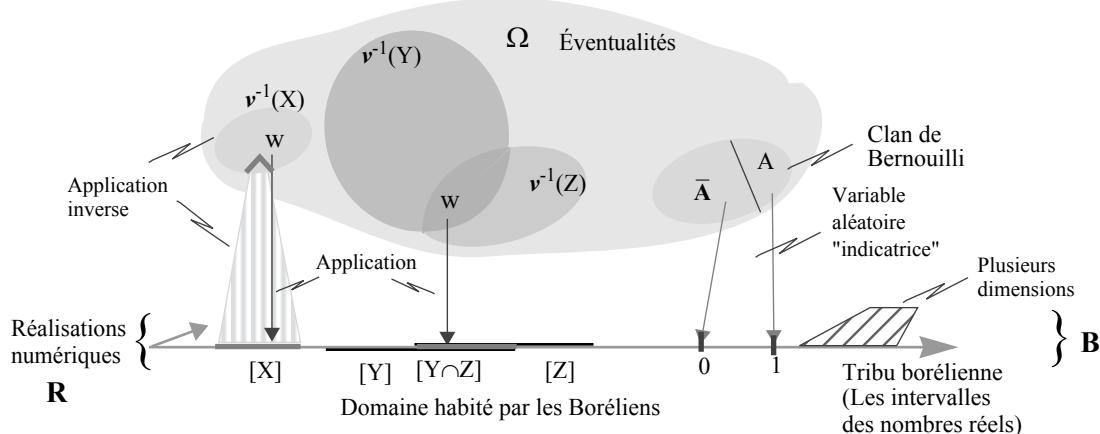
- Dans l'espace aléatoire Ω , soit donc la tribu \mathbf{T} des événements potentiels ;
 - Dans l'espace numérique (des nombres réels), soit la tribu borélienne \mathcal{B} , c'est-à-dire la tribu engendrée par tous les intervalles ouverts de \mathbb{R} . Elle comprend les réunions et intersections d'intervalles, et ceux réduits à un seul point. Un élément de cette tribu est une réalisation borélienne.

Un aléa numérique (ou variable aléatoire) est défini comme :

Toute application de \mathcal{B} vers Ω telle que l'application inverse de toute réalisation borélienne appartienne à la tribu \mathbf{T} des événements potentiels.

La Figure 26 présente cette interface entre les éventualités et des valeurs.

Figure 26. Aléas numériques



Un petit exemple: soit que Ω désigne comme éventualités le fait que le prochain bonhomme qui apparaît de derrière le coin de la rue est glabre (ω_a), barbu (ω_b), ou moustachu (ω_c). La tribu va comprendre les éventualités d'indices a, b, c, ainsi que l'événement formé par $\{\text{barbe} \cup \text{moustache}\}$, ce qui donne quatre événements, disons A, B, C, D.

Dans les applications aléatoires, la correspondance numérique entre les aléas et les valeurs est affaire de convention. Ainsi par exemple, on peut choisir l'application $\{A \Rightarrow 1, B \Rightarrow 0,5, C \Rightarrow 1, D \Rightarrow 2\}$, l'idée étant de "scorer" la touffe de poils. Dans bien des cas, les conventions sont préétablies; ainsi, les faces des dés habituels sont déjà marquées à cette fin, mais on pourrait fort bien le faire autrement. Ainsi DE MOIVRE (vers 1720), fondateur des fonctions génératrices, a utilisé un dé à dix faces et s'en très bien sorti.

6.7.2 Structure d'algèbre aléatoire

Le fait de définir les variables aléatoires comme des applications numériques – des fonctions, dans les cas courants – a la vertu de pouvoir en faire des compositions et des transformations en préservant leur qualité d'aléa. Quelques clefs en sont citées ci-dessous.

Soit donc $v(\omega)$ l'image de ω sur \mathbf{R} par l'application aléatoire $v(\cdot)$; $v(\Omega)$ est l'application issue de l'ensemble aléatoire considéré.

La v_a (variable aléatoire) est dite:

- Discrète finie, si l'ensemble $v(\Omega)$ est fini;
- Discrète infinie, si l'ensemble $v(\Omega)$ est dénombrable et admet un minorant;
- Continue, si l'ensemble $v(\Omega)$ est un intervalle de \mathbf{R} non réduit à un point.

Soient u et v deux v_a . On montre que les compositions suivantes sont des v_a :

- Le changement de variable $u(\omega) = a v(\omega) + b$;
- La bijection numérique $u(\omega) = \varphi v(\omega)$;
- La somme $u(\omega) + v(\omega)$;
- Le produit $u(\omega) \cdot v(\omega)$;

Ceci s'étend aux constructions de même style faites avec plus de deux variables. En bref, l'ensemble des aléas numériques défini de cette façon a, comme annoncé par le titre de cette section, une structure d'algèbre. Dans ce texte, on ne fera flotter que deux bouées sur cet océan des probabilités, mais elles sont choisies pour leur importance dans les analyses des systèmes aléatoires: l'espérance mathématique et les convergences, ainsi que les mesures de probabilité associées à des variables jointes.

6.8 Les fonctions génératrices

Les fonctions génératrices et les processus stochastiques sont les fleurons des aspects aléatoires de la systémique. En effet, ce sont des engins qui engendrent quelque chose; ce sont des processeurs, bref des foncteurs, alors que les autres notions sont surtout des descripteurs, des "combineurs" et des transformations d'écritures.

Les premières transformées sont nées dans un tripot, lors d'une partie de dés historique entre Jules CÉSAR et DE MOIVRE. Franchissant la rue BICON, Jules joua le premier et s'écria «Alea jacta est! (les dés en sont jetés)». Puis il recommença et demanda : ça fait combien ?, le mathématicien DE MOIVRE s'en occupa et trouva le moyen d'exprimer les probabilités associées aux sommes des valeurs obtenues en lançant n dés réguliers "généralisés" à k faces numérotées de 1 à k , ce qui revient à engendrer les masses de probabilité de sommes de variables aléatoires à valeurs entières : ce seront les premières fonctions génératrices de probabilité.

L'expression (21) ci-dessous définit la fonction génératrice directe, laquelle s'adresse aux réalisations de variables aléatoires finies entières, telles le nombre de "succès", etc.

$$(21) \quad \Phi(t) = P\{v(\omega) = 0\} \cdot t^0 + P\{v(\omega) = 1\} \cdot t^1 + P\{v(\omega) = 2\} \cdot t^2 + \dots$$

Cette expression est plus synthétique sous la forme de l'espérance mathématique (22) :

$$(22) \quad \Phi(t) = \sum_{k=0}^{\infty} p_k \cdot t^k = E(t^x)$$

La série numérique des probabilités p_k somme sur 1. Donc, si l'on admet la convention $0 \leq t \leq 1$ pour la variable opérationnelle t , la série définie par (22) est convergente.

Ceci correspond donc aux transformées-z présentées dans l'exposé «Analyse de processus». Poursuivant le même scénario que dans cet exposé (qui préparait celui-ci), les transformées de LAPLACE (1812), puis celles de FOURIER permettent d'engendrer de façon univoque toutes les fonctions de densité de probabilité et leurs moments.

Remplacer la variable opérationnelle t ci-dessus par e^{-s} , où s est non-négative, forme la transformée de LAPLACE :

$$(23) \quad L(s) = E(e^{-sx})$$

Si les réalisations x de la variable aléatoire $v(\omega)$ sont positives entières, on retrouve facilement la génératrice de DE MOIVRE :

$$(24) \quad L(s) = E(e^{-sx}) = P\{v(\omega) = 0\} + P\{v(\omega) = 1\}e^{-s} + P\{v(\omega) = 2\}e^{-2s} + \dots$$

Ce qui implique la relation suivante entre les deux fonctions génératrices, si $s \geq 0$:

$$\Phi(e^{-s}) = L(s)$$

Mais la génératrice $L(s)$ développe les probabilités pour les variables aléatoires positives quelconques, disons $V(\omega) = x_1, x_2, \dots$:

$$(25) \quad L(s) = E(e^{-sx}) = P\{v(\omega) = x_1\}e^{-sx_1} + P\{v(\omega) = x_2\} \cdot e^{-sx_2} + \dots$$

Lorsque s est positive, l'expression (25) est manifestement convergente pour $x \geq 0$. Si la variable n'est pas astreinte à être positive, l'espérance mathématique (23) peut ne pas exister – il y a divergence ; on fait alors appel à la transformée de FOURIER, également rendue familière par l'exposé «Analyse de processus».

A cette fin, l'expression opérationnelle utilisée est cette fois $e^{-i\theta x}$, où i est l'imaginaire $\sqrt{-1}$.

En fine, l'expression générale (26) donne leur vraie esthétique aux génératrices, montrant simplement leur expression pour une variable possédant une fonction de densité continue $f(x)$:

$$(26) \quad L(s) = \int_0^{\infty} e^{-st} \cdot f(t) dt$$

La transformée inverse définit alors de façon univoque la fonction de densité de probabilité. Celle-ci, déjà évoquée dans le cas de l'ensemble d'événements, donne la propension à se réaliser sur tous les intervalles définis sur les valeurs réelles. L'expression (26) peut même s'appliquer aux situations mixtes discrètes-continues, à condition de l'exprimer via l'intégrale de STIELTJES. Mais pour se faire servir celle-là, il faut croire à la métémpsychose, et se réincarner en autre chose qu'en gestionnaire.

Elle peut aussi engendrer tous les moments factoriels par dérivation successives, mais il faut laisser un peu de place aux dizaines de bons livres qui s'en occupent, et se focaliser à présent sur deux de ses retombées sur la systémique.

6.9 Les processus aléatoires

Selon E. PARZEN (Stochastic Processes, Holden Day, 1962, pp. 8, 11 et 74-75), un processus stochastique est une famille de variables aléatoires $v(t)$, $t \in T$, indexées par un paramètre t variant dans un ensemble indiqué T . La description en appelle dès lors l'expression de la probabilité jointe des $v(t)$, ce qui n'est en général praticable que via les fonctions génératrices (dites aussi "caractéristiques"), dont on vient de tirer un exemplaire.

Il s'agit en chaque point t de décrire $v(t)$ en termes d'une famille de variables aléatoires dont la loi de probabilité est connue. Il convient donc de se situer dans l'espace de réalisation associé au phénomène aléatoire concerné; cet espace de réalisation est l'ensemble de toutes les descriptions de toutes les issues du phénomène géré par le hasard – comme il a été bien accroché à la section 3.

Le processus aléatoire n'est pas toujours une séquence. Un processus stochastique (avec PARZEN, p. 24) est une fonction de deux arguments, dont l'un est l'indexation et l'autre est la réalisation: $\{v(t,s), t \in T, s \in S\}$.

Dans les notations de cet exposé, la rédaction de PARZEN correspond à la suivante:

- « Si t est fixé, $v(t,.)$ est une fonction de l'espace de probabilité S ou, de façon équivalente, $v(t, .)$ est une variable aléatoire;
- D'autre part, pour s fixé (dans S), $v(., s)$ est une fonction de t qui représente une observation possible du processus stochastique $\{v(t), t \in T\}$. La fonction $v(., s)$ est appelée une réalisation, ou une fonction d'échantillon, du processus. C'est celle qui est représentée comme une séquence (ou ensemble) de valeurs ou de symboles dans un diagramme conventionnel ».

On la choisit, on l'embrasse, on entre dedans, puis on se retire...

C'est la carrière, évidemment.

7 Les contrefaçons

7.1 L'aléa synthétique

L'expression « les contrefaçons du hasard » est de B. MANDELBROT (op. cit.). Bien que cet auteur n'explicite pas ce qu'il entend de la sorte, elle sera récupérée subrepticement ici pour désigner les façons de traiter formellement les outputs des agents du hasard qui ont été cités à la Figure 23 et dans ses environs immédiats. Parmi ces agents on trouvait les expériences, les échantillonnages, les phénomènes non maîtrisés, les interactions trop complexes pour être innocentes.

L'idée est de gérer ces outputs (qui sont des éventualités, des conjonctions d'événements, des surprises) avec des instruments adéquats. Après tout, la statistique mathématique est la modélisation de la gestion du hasard. De l'océan de contributions en ce domaine, on fera surnager ici deux épaves auxquelles peuvent s'accrocher quelques gestionnaires privilégiés. Les autres resteront dans les ténèbres de l'ignorance, ou alors ils sont déjà éclairés par de bons auteurs, ou bien ils s'en foutent complètement et on les comprend. Et d'ailleurs, disent ces dames à juste titre, un "beau ténébreux", ça n'éclaire pas loin.

Pourquoi "contrefaçon" ? Parce qu'une version de l'expression du hasard est "ce qui ne peut pas être résumé". Dès lors, lorsqu'on construit des expressions qui le décrivent sans lui laisser tout son potentiel d'éventualités, on en réduit la variété, on ne laisse pas aux agents du hasard tous leurs degrés de liberté. Ce faisant, on augmente les chances de décrire succinctement ces manifestations et de pouvoir les maîtriser en partie.

C'est cette capacité de descriptions succinctes qui permet d'élaborer, de gérer, puis de commander, des systèmes soumis à des agents hasardeux, à des bourrasques ou à des perturbations. Les domaines offrant de tels exemples de systémique aléatoire sont aux premières loges la décision et la commande aléatoire – qui consomment des descripteurs synthétiques – les processus de commande floue – qui conservent une forme d'incertitude. Elle est complétée par une problématique plus générale, dite du contrôle adaptatif.

Plus associée à la gestion des EAH, se dessine la gestion intégrée des risques, dans laquelle la version "statistique", ou dite "actuarielle", fait aussi une grosse consommation de descripteurs de synthèse. D'autres approches, plus récentes et qui concernent plutôt les événements (graves) occasionnels que répétitifs, relèvent de la théorie du potentiel et de la crédibilité. L'exposé « Le Goût des risques » y est consacré dans le Tome d'Ouessant.

Les modes de génération des contrefaçons d'aléas synthétiques sont :

- Les concentrations (les tendances centrales – comme l'espérance mathématique – et les paramètres de synthèse – comme les moments) ;
- Les convergences (comme la loi des grands nombres, les distributions-limites) ;
- Les compositions aléatoires (décrivant dans une seule expression des aléas multiples, des conjonctions de variables aléatoires).

Comme les ahuris qui sont incapables de choisir sur un chariot de fromages, on ne prendra qu'un petit morceau de chaque – cela fera une économie de 15 ans et de 10000 pages.

7.2 Les concentrations

7.2.1 Le culte de la moyenne

C'est dans la version du hasard qui semble correspondre au hasard "bénin" de MANDELBROT que se situent diverses voies de prise en charge par des formulations "concentrées" dans des lois, dont les fonctions génératrices de DE MOIVRE ont donné le ton juste.

Comme dans un bon régime politique, ces lois sont appliquées par le pouvoir exécutif qui est la statistique, laquelle est appliquée à ses citoyens, les statistiques. Enfin, de même que le pouvoir judiciaire "dit le droit", ces "lois" de l'aléatoire sont sanctionnées par la rigueur de la mathématique.

De là, on passe aux régularités et aux convergences, puis à la sacro-sainte moyenne. La source étant les mesures physiques, le centre d'intérêt était d'abord les "lois de l'erreur", et la domination des effets du "grand nombre" conduisant à sanctifier la moyenne. Ces régularités sont notamment très bien typifiées par les travaux de GALTON et la loi dite "Normale" qui en est issue :

« [...] Chaque fois qu'un grand échantillon d'éléments chaotiques sont pris en mains et rangés selon leur ordre de grandeur, on s'aperçoit qu'une forme de régularité insoupçonnée et très belle existait, latente, depuis le début [...] ».

(Francis GALTON, *Natural Inheritance* (traduction de M. ROBERT).

La loi des grands nombres et la loi Normale sont cependant remises en cause lorsque :

- Des comportements sont connectés de telle sorte que l'hypothèse d'indépendance ne prévaut pas (par exemple les acheteurs en bourse) ;
- Les convergences et homogénéités de distributions initiales ne sont plus valables ;
- Des transformations comme la superposition (somme de variables identiquement distribuées) ne prévalent pas, ou que des perturbations rendent les variables autonomes.

L'espérance mathématique, ainsi que les concentrations sur la moyenne, trouvent leur épaulement chez QUÉTELET (dans *Physique sociale ou essai sur le développement des facultés de l'homme*). Les travaux de QUÉTELET sont typiques de cette attitude consistant à exprimer un comportement ou un phénomène "général" à partir du comportement "moyen" issu d'un grand nombre, dont les variations sont lissées par l'opération dite en Américain "Averaging out". QUÉTELET poussait sa vision des choses jusqu'à reconstruire l'homme moyen comme étant un personnage objet d'une juste et morale "régulation sociale".

Le thème dominant des "lois des erreurs" se manifeste en décrivant comment les valeurs observées s'assemblent autour d'une moyenne. La contribution espérée était prédictive : fournir une estimation des risques pour qu'une erreur de taille donnée survienne. Or, c'est ce qui a beaucoup changé récemment : les "lissages" à des fins prédictives escamotent ces anfractuosités qui engendrent les variétés de comportement.

Comme le montrait MANDELBROT, cela n'apprend rien de décrire la côte bretonne par une trajectoire lisse, car aucun segment de droite ne donne d'idée sur son extrapolation. Il en va de même pour la surface d'un brocoli ou des nuages, mais le modèle des fractals peut y contribuer.

Cette vision du rôle de la "valeur centrale" comme objet type de comportement – n'expliquant et ne décrivant rien des écarts – est donc remise en cause par d'autres descriptions. Ceci restreint donc les possibilités des régulateurs, donc de la commande, par le fait de ne pas accepter les convergences écrasant les écarts.

Cette vision peut être formalisée par l'opposition entre la distribution Normale et la distribution de CAUCHY. Alors que la première, bien sentie par GALTON, formulée par EDGEWORTH, et fleuron de GAUSS, est utilisée avec une grande fécondité, celle de CAUCHY sert plutôt de support formel à la remise en question des abus de la domination de la Normale. Sources de grands discours, elles ne diffèrent cependant que par un petit paramètre a , qui vaut 2 pour la Normale, et 1 pour la CAUCHY. Une petite perturbation, et voilà un autre monde du hasard?

7.2.2 L'espérance mathématique

L'espérance mathématique s'adresse aux variables aléatoires, lesquelles sont aux événements. Un événement élémentaire est une des issues d'un processus que l'observateur ne maîtrise pas, selon les deux versions suivantes :

- Soit la source est un processus qu'il ne connaît pas, et qu'il ne peut donc pas commander: l'événement est alors une des manifestations du hasard, et on n'y peut rien;
- Soit il s'agit d'un processus que l'on peut décrire – et même provoquer, comme lorsqu'on joue à un jeu – mais dont l'issue ne peut être ni prédite avec certitude, ni obtenue volontairement. Celle-ci est dite aléatoire.

L'espérance mathématique est concernée par la deuxième version; elle suppose une forme de "régularité" sous-jacente, laquelle peut être masquée, ou même indistinguable, étant donné les interférences ou la complexité qui ne permet pas de l'élucider. Lorsque l'ensemble des issues de tels processus peut être exprimé, même s'il ne peut être dénombré, alors on entre dans la théorie des probabilités.

L'espérance $E[\nu(\omega)]$ est l'output d'un opérateur E [...] appliqué à une telle variable aléatoire simple ou composée, disons $\nu(\omega)$. Soit donc \mathbf{A} un sigma-champ d'événements sur Ω ; si un événement ω_i se passe, l'application $\nu(\omega_i)$ en donnera une réalisation scalaire x_i dans l'intervalle V de l'espace réel \mathbf{R} .

La définition se fait à partir de la fonction de répartition (donc cumulative), que celle-ci soit discrète ou continue :

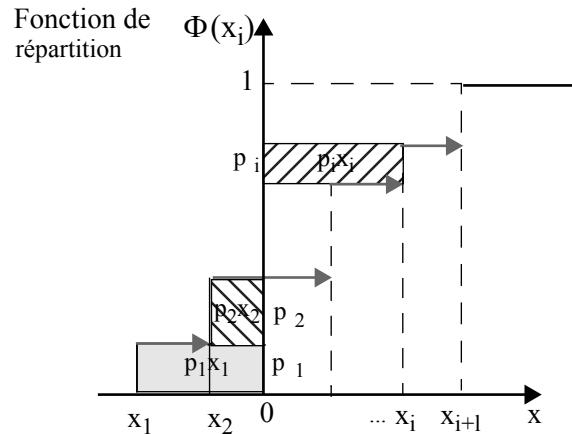
- Lorsque la variable aléatoire est discrète finie, les fréquences relatives jouent le rôle de poids des valeurs associées. L'espérance mathématique correspond alors au barycentre des valeurs x_1, \dots, x_N pondérées par leurs probabilités respectives p_1, \dots, p_N .

Cette version est celle de la classique expression (27) :

$$(27) \quad E[x] = \sum_{i=1}^N p_i \cdot x_i$$

- Dans le cas où la variable est discrète et infinie (N est infini), la notion de barycentre devient inadéquate, et on demandera que la série soit absolument convergente. La Figure 27 représente les composantes entrant dans cette définition.

Figure 27. Les composantes de l'espérance

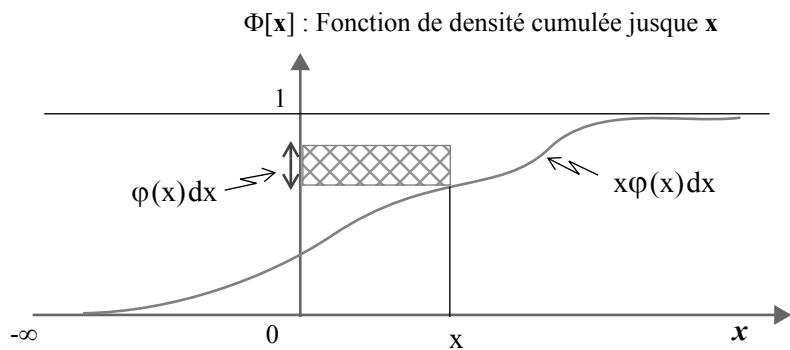


Dans le cas de variable aléatoire absolument continue, l'expression de l'espérance est donnée par l'intégrale (28), si elle veut bien se donner la peine de converger:

$$(28) \quad E[x] = \int_{-\infty}^{\infty} x\varphi(x) \cdot dx$$

La Figure 28 montre l'intégrande de (28).

Figure 28. L'intégrande de l'espérance mathématique



7.2.3 Définition générale de l'espérance

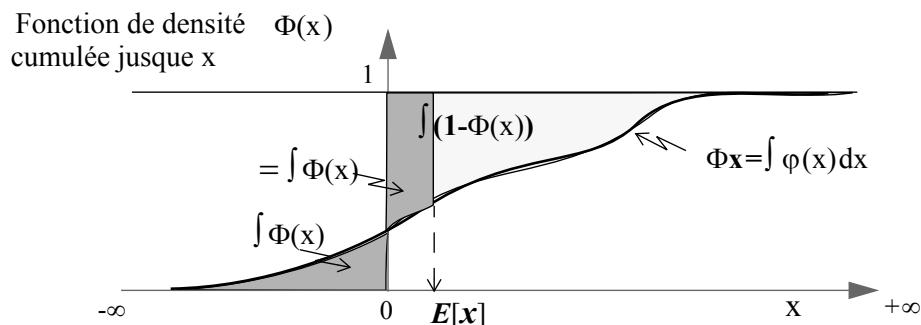
Soit $\Phi(x)$ la fonction décrivant la répartition de la variable aléatoire $v(\omega)$ sur l'intervalle numérique réel $[-\infty, x]$. L'expression formelle utilisera l'opérateur E , et s'écrira $E[.]$.

$$(29) \quad E[x] = \int_0^\infty [1 - \Phi(x)] \cdot dx - \int_{-\infty}^0 \Phi(x) \cdot dx$$

Selon cette définition, l'espérance mathématique demande que les deux intégrales existent ; ce qui n'est pas nécessairement le cas de toutes les fonctions de répartition – on verra le contre-exemple que présente celle dite "de CAUCHY".

La représentation géométrique est la différence de deux aires définies par les fonctions de densité cumulées, situées de part et d'autre de l'axe de placé par commodité sur la valeur 0. L'espérance mathématique est le lieu sur l'intervalle $[-\infty, +\infty]$ indiquant le centre de gravité de la dispersion de la masse de probabilité décrite par la fonction cumulative Φ .

Figure 29. L'espérance mathématique – densité continue



Une formulation très légère de l'espérance mathématique, valable seulement dans le cas où les réalisations sont exclusivement des valeurs positives entières, est la suivante :

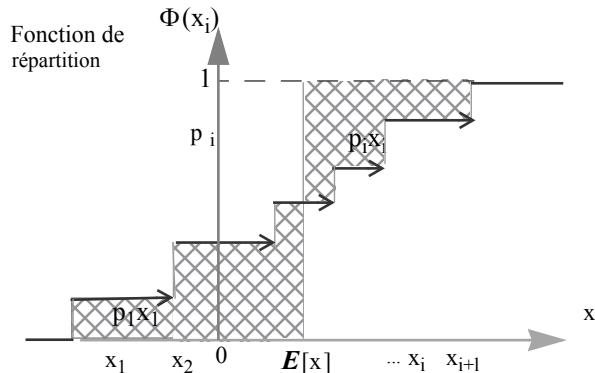
$$(30) \quad E[x] = P\{x>0\} + P\{x>1\} + P\{x>2\} + \dots$$

Cette formulation (30) est peu familière, et inattendue car elle ne fait pas apparaître explicitement les valeurs spécifiques (1, 2,...) de la variable. Pourtant, il suffit de développer chacune de ces probabilités (qui est un cumul) pour obtenir (31) qui convient bien :

$$(31) \quad E[x] = \sum_{i=0}^{\infty} i \cdot P\{x = i\}$$

Dans le cas discret, sa représentation est la Figure 30, qui répond mieux à l'intuition que sa collègue continue.

Figure 30. L'espérance discrète



7.2.4 Une Application de 40 milliards

Soit l'exemple quasi fictif de LOTO. Celui-ci a lieu par "tranches", et disons que chacune de ces éditions fasse 10^8 écus de bénéfice sur 10^5 joueurs. La "perte moyenne" par joueur par tranche est donc de 10^3 . L'espérance mathématique étant nettement inférieure à l'utilité, la propension au risque et la perspective d'un gain potentiel élevé mais exceptionnel peut seule motiver la participation, ce qu'expose la théorie de la décision dans l'incertitude. En effet, les "chances" de ne pas perdre sont les plus élevées en ne jouant qu'une seule fois, et jouer à répétition, ou "jusqu'à ce qu'on gagne" n'est pas plus malin. Il en va ainsi, d'ailleurs, pour tous les jeux qui ne sont pas à somme nulle entre les participants.

7.3 Les convergences

7.3.1 Les "grands nombres"

Le succès de "loi des grands nombres" est peut-être dû à la simplicité de son expression populaire, quelque chose comme "tend vers la moyenne". Or, cette loi dit seulement que, sous certaines restrictions, la moyenne arithmétique non pondérée est un estimateur convergent en probabilité de l'espérance mathématique. Formellement,

- Soit ε une grandeur positive arbitrairement petite;
- Soient p variables aléatoires x_1, \dots, x_p qui suivent la même fonction de densité, de même espérance mathématique μ .

La propriété s'écrit:

$$\lim_{p \rightarrow \infty} P\{|\bar{x}_p - \mu| < \varepsilon\} = 1$$

On a bien lu que les variables doivent être de même "loi", d'espérance finie et surtout, indépendantes. Ces hypothèses peuvent être maintenues dans certains cas évidemment, mais la gestion des systèmes aléatoires n'y croit pas beaucoup: au lieu d'un "paquet" d'observations simultanées, elle suit des yeux des processus, et ceux-ci, surtout ceux qui apportent des événements "innovants" ou parfois dangereux, présentent des associations de valeurs successives qui contredisent les hypothèses de la loi des grands nombres.

C'est ainsi que la vision des processus boursiers, leurs séquences et interactions – à des fins prédictives, on imagine – est très différente de celle de calculs de rendements "moyens", surtout sur des périodes élastiques pour attirer des gogos. De même, les "statistiques" d'accidents, traduites dans des moyennes, n'engagent pas le gestionnaire du risque, mais c'est l'intégration des générateurs d'événements qui en forme la gestion moderne, c'est-à-dire depuis qu'elle écoute quelques systémiciens.

7.3.2 L'inégalité de Chebychev et la convergence

La formulation de l'espérance mathématique fournit immédiatement – sous les mêmes hypothèses – l'inégalité dite de MARKOV avec une constante c :

$$(32) \quad P\{\tilde{x} \geq c\} \leq \frac{E_x}{c}$$

Dans (32), remplaçons c par ε^2 , et appliquons l'inégalité de MARKOV à une variable définie par $z = |x - E_x|^2$. Dès lors, $E[z] = s^2$, la variance de x . Ceci permet d'écrire l'inégalité de CHEBYCHEV:

$$(33) \quad P\{|x - E_x| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

Soit à présent:

$$k = \varepsilon/\sigma$$

L'expression (33) est alors généralisée, et devient valable pour les variables aléatoires tant continues que discrètes:

$$(34) \quad P\{|x - E_x| \geq k\sigma\} \leq \frac{1}{k^2}$$

Les gens ne comprennent pas bien l'expression (32), mais l'utilisent implicitement en pensant toujours qu'il est "peu probable" que des valeurs fort écartées de la moyenne (en nombre de fois l'écart-type) se présentent. Il est vrai que (32) est très générale, n'étant réservée à aucune restriction sur le processus aléatoire ou la fonction de distribution. Évidemment, si l'écart-type est grand, les bornes proposées sont rapidement très écartées.

Ainsi, supposons un tire-fesses de 20 places sur la piste de ski. L'espérance de poids des skieurs est de 75 kg, d'écart-type 10. Donc l'espérance mathématique du poids porté est de 1500 kg. Le truc est calibré sur 1800 kg, ce qui est confortable, car $1800 - 1500 = 300 = 30\sigma$. L'application de (34) dirait:

$$P\{|x - E_x| \geq 300\} \leq \frac{1}{30^2}$$

Le tire-fesse présente donc une probabilité majorée par 1/900 d'être en surcharge. C'est pourquoi dans les stations de sports d'hiver, et dans un but évident de conscience professionnelle, les préposés regardent toujours les fesses des skieuses – et il y en a une sur 900 qui doit rester en bas ?

Mais la relation (32), bien que limitée aux entiers, est plus inattendue car elle ne fait pas intervenir l'écart-type. Ainsi, en travaillant à la pelle, par kg de touriste à la fois, la probabilité qu'un touriste pèse plus de 90 kg est :

$$P\{x \geq 90\} \leq \frac{75}{90}$$

La réponse est donc de l'ordre de grandeur de 0,83. Cette relation est peu spectaculaire, mais elle est utile en systémique car elle donne une certaine aisance naturelle à évaluer les grandeurs et la vraisemblance des assertions qui les concernent.

En ce qui concerne la gestion des processus, la question qui se pose est le fait qu'il s'agit bien de variable aléatoire libre. Dans la plupart des systèmes gérés, justement, les objets sont calibrés pour amoindrir les probabilités de mauvaises surprises. Ainsi, les bagages des voyageurs sont pesés – selon des limites –, les véhicules ont une tare connue etc. Ces dispositions se retrouvent typiquement dans la prévention des risques, ainsi que le signale l'exposé «Les Goûts du risque» du Tome d'Ouessant.

7.3.3 Convergence et Normalité

Bien des gens du peuple statistique se demandent comment dans des fonctions de densité apparaissent des êtres mathématiques tels que $\sqrt{2\pi}$, qui n'ont à première vue rien à voir avec des fréquences et autres êtres des comptages, et concernent des cercles et de la trigonométrie dont on n'a rien à cirer ici.

La réponse est dans les bons livres, et pas ici, mais comme on s'est plaint plus haut des "contrefaçons du hasard", on se doit de donner une petite indication sur la source du phénomène : c'est encore une affaire de convergence.

Tout part de l'analyse combinatoire construite sur l'espace probabiliste de BERNOULLI. Celui-ci considère l'événement A (appelé "succès", on a dit pourquoi) de mesure de probabilité p et son complément A^c , de probabilité $(1-p)$. L'interprétation courante est qu'elle décrit le nombre aléatoire $r=1$ ou 0 de succès en un nombre $n=1$ donné d'essais, et pour une probabilité p donnée de réalisation de A par essai.

Donc, le modèle de base est :

$$P\{r|n=1, p\}$$

De là, on peut construire les extensions. Par exemple, considérer le nombre d'essais jusqu'à obtention du premier succès ; c'est alors n l'aléa, et le modèle, dit géométrique, est :

$$P\{n|r=1, p\}$$

Le nombre r de succès en n essais correspond au modèle binomial (et pas la "distribution biennale", comme une copie d'examen l'a suggéré) :

$$P\{r|n, p\}$$

Le nombre d'essais jusqu'à obtention du r^e succès est le modèle de PASCAL :

$$P\{n|r, p\}$$

Ces modèles ont leur correspondant continu. Par exemple, on peut considérer cette fois comme aléa la probabilité p de BERNOULLI, exprimée via le nombre de succès en n essais; ceci est évidemment une fonction de densité (appelée la fonction " β ") sur $[0; 1]$ puisque p est une mesure de probabilité, soit :

$$P\{p|n, r\}$$

Dans la même veine, la densité de probabilité du temps d'attente jusqu'à l'obtention du premier (ou "prochain") succès est décrite par l'exponentielle négative. Celle du temps d'attente jusqu'au r^e succès est la fonction d'ERLANG, correspondant donc à celle de PASCAL.

L'argument présent n'est pas de développer ces constructions, mais de signaler que, dans leur version discrète en nombres entiers positifs, elle s'exprime par les "formules" combinatoires, ce qui implique des expressions factorielles, dont l'exemplaire le plus reproduit est le coefficient de la binomiale :

$$(35) \quad C_n^r = \frac{n!}{r!(n-r)!}$$

La requête en calcul pour de telles fonctions arithmétiques peut être considérable, surtout lorsque le problème demande des compositions de telles formules, et aussi lorsque les paramètres (n et r) deviennent grands. En pratique, n correspond à un nombre d'éléments d'un échantillon, de répétitions dans un processus etc., de sorte que la gageure de calcul est le plus pertinente.

Aussi, un appel a été fait à une approximation, dite "de STIRLING", qui s'exprime par (36) :

$$(36) \quad \lim_{n \rightarrow \infty} n! \approx \left(\frac{n}{e}\right) \cdot \sqrt{2\pi n} \cdot \left(1 + \frac{1}{12n}\right)$$

Et voilà le $\sqrt{2\pi}$ qui apparaît (le $1/12n$ devient rapidement négligeable). Avec cela, des gens intelligents ont construit les fonctions de densité telles que la Normale et associées.

Justement, c'est là qu'intervient la deuxième convergence célèbre, via une des versions du théorème dit de la limite centrale. Cette expression est issue de la propriété de convergence suivante : soit un ensemble de variables aléatoires indépendantes x_i , d'espérance μ et de variance σ^2 identiques ; dans ce cas :

$$\text{var}\left[\frac{x_1 + x_2 + \dots + x_n}{n}\right] = \frac{\sigma^2}{n} \rightarrow 0$$

La distribution de cette somme devient donc concentrée sur un seul point, s'interprétant comme une "limite centrale".

Lorsqu'une transformation linéaire est effectuée sur une telle somme de sorte que la variance reste constante, on est sur la voie du théorème de la limite centrale. Une première version en a été présentée et démontrée par DE MOIVRE, et par LAPLACE ensuite; la version qui tient encore aujourd'hui est dite de LINDEBERG-LÉVY, bien qu'on déplore le décès de ces deux gentlemen.

Le théorème dit simplement que toute somme standardisée de p variables aléatoires identiquement distribuées (ce qui implique même espérance et variance, non-nulles) a pour limite la fonction de répartition Normale réduite " $N(0;1)$ ".

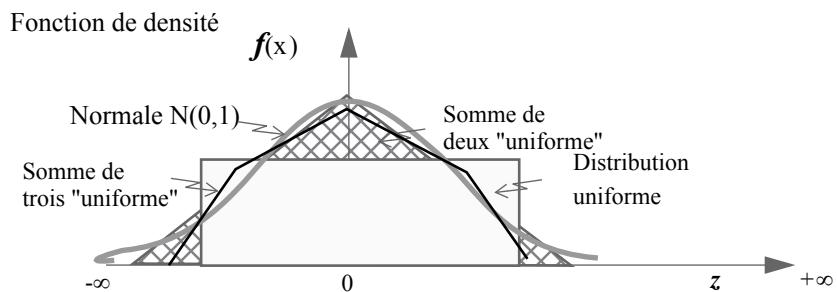
Si z_p est une telle variable-somme, alors:

$$(37) \quad \lim_{p \rightarrow \infty} (z_p \leq k) = \frac{1}{(2\pi)^{1/2}} \cdot \int_{-\infty}^k \exp\left(-\frac{v^2}{2}\right) dv$$

L'implication est considérable, car les convergences sont rapides et la tentation est grande d'utiliser la distribution Normale en trop de circonstances.

Les convergences sont d'autant plus rapides que les variables entrant dans la composition de la somme ont des distributions simples et peu bizarres. Le record en ce sens est alors évidemment l'"uniforme", que l'on schématise ici sur la Figure 31 sans la paramétriser.

Figure 31. Convergence de la densité uniforme



7.3.4 Leçon de morale

L'expression (37) est celle d'un pouvoir totalitaire qui a sévi sur toute la statistique, en concentrant tout sur la tendance centrale. Beaucoup de phénomènes, en effet, présentent une complexité de causes et de manifestations qui ne peut être élucidée, et dès lors l'hypothèse maintenue est que leurs réalisations sont la résultante d'un grand nombre de processus aléatoires identiquement distribués, dont la somme réduite est soumise alors à la convergence vers la Normale. Les cas en sont très diversifiés, telles les erreurs de mesures, les comportements des agents économiques, des grandeurs échantillonnées et mille autres.

Seulement, les hypothèses sont beaucoup plus restrictives que les clients du théorème ne le croient; combien y a-t'il de variables aléatoires et indépendantes dans un même contexte de problématique? Et ces moyennes, existent-elles, et veulent-elles vraiment dire quelque chose? Même des "événements" statistiques paraissant spectaculaires résistent parfois mal à une deuxième analyse. Ainsi le phénomène suivant a déferlé sur tous les médias, tant et si bien que le Journal de Gériatrie de l'ASLJSIGASRH en a parlé:

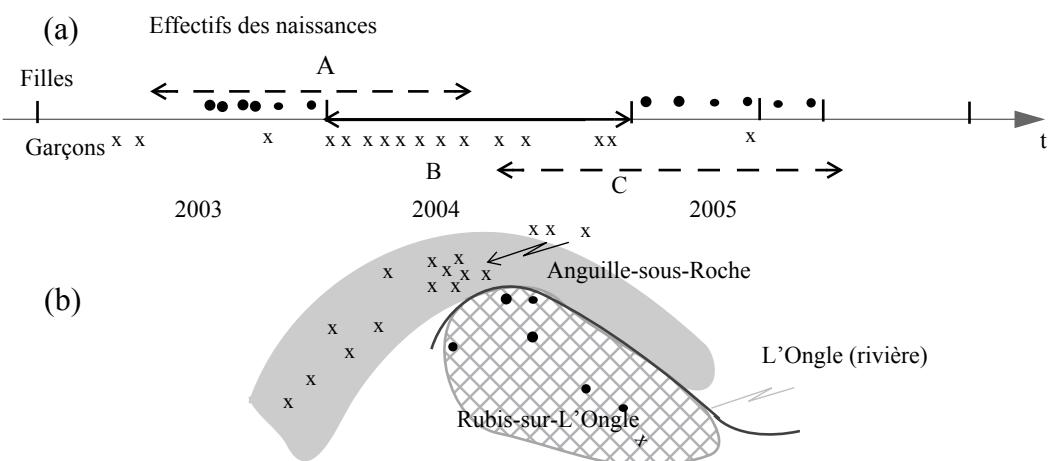
«Cette année (2004), il y a eu 12 naissances dans la paroisse d'ANGUILLE-SOUS-ROCHE (ici: "AsR"), et les 12 sont des garçons!»

À cette annonce, les croquantes et les croquants d'AsR se sont jetés sur leurs tables de la distribution Binomiale (l'expression (31) ci-dessus), et ont lu avec stupéfaction que la probabilité d'avoir $r=0$ fille ou encore moins en $n=12$ essais indépendants "au hasard" est bornée par 0,000244141 ou, si on préfère, une chance sur 2^{12} , soit 1/4096. Or, l'instituteur de statistique a dit qu'on rejette les hypothèses des curés quand les probabilités d'erreur de première espèce (il a dit "de la pire espèce") sont supérieures à 0,05.

La conclusion est alors de rejeter une des hypothèses de l'expérience, par exemple le fait que les essais sont indépendants (il y aurait des "corrélations de ménages"?), ou alors que la probabilité de genre, féminin ou masculin, n'est pas de $p=0,5$. Bref, c'est extraordinaire, il s'est passé quelque chose de troublant dans la paroisse en 2004.

Un systémicien qui y prenait son verre a cependant dessiné la Figure 32 sur la nappe en papier du Café des Sports.

Figure 32. La Nativité d'Anguille-sous-Roche



Ce qui a amené les remarques suivantes:

- Les effectifs observés de 12 garçons se situent dans l'année civile 2004, c'est-à-dire dans l'intervalle (B) de l'axe temporel. Si on avait pris les effectifs du 30 juin au 30 juin (pourquoi pas), ce serait par exemple l'intervalle C, présentant disons 6 filles contre 8, et la probabilité associée ne serait inférieure qu'à 0,143, ce qui n'attire pas l'attention.

Il en irait de même pour un intervalle de temps antérieur, tel que A sur la Figure 32 ;

- D'autre part, la forme de la paroisse est montrée dans la partie (b) de la Figure 32. Ceci est dû au parcours de la rivière (l'Ongle) qui la sépare de sa voisine du Sud. Si on met ensemble les deux paroisses, ou si on en réunit les parties de droite, on a des proportions de naissances très ordinaires – comme le montre bien la Figure 32 ;
- Enfin, si l'on considère que trois des garçons d'AsR sont issus d'une autre commune, mais venus au monde là parce qu'il y a une maternité, que deux sont issus de familles où il y avait déjà quatre filles, et qu'un garçon est né dans une roulotte de Bohémiens qui se rendaient aux Saintes Marie de la Mer en Camargue, on se demande pourquoi on vient nous bassiner les oreilles avec des statistiques à la noix.

En conclusion sur le rôle de ces deux points importants de description de hasards dans les systèmes, on ne peut que recommander de poursuivre la lecture du présent exposé, car c'est à ce genre de "résumés" que s'opposent beaucoup de phénomènes aléatoires relevant, selon l'expression de MANDELBROT, du "hasard sauvage". La leçon de morale est alors que la systémique ne signifie pas du tout de se contenter de régularités apparentes, de concentrations et résumés des manifestations aléatoires. Au contraire, les processus qui sont l'objet de ses analyses sont censés conserver leur personnalité, et même les caprices occasionnels de leurs composantes doivent être adulés comme ceux des stars-pétasses.

8 Les compositions aléatoires

8.1 Variables conjointes

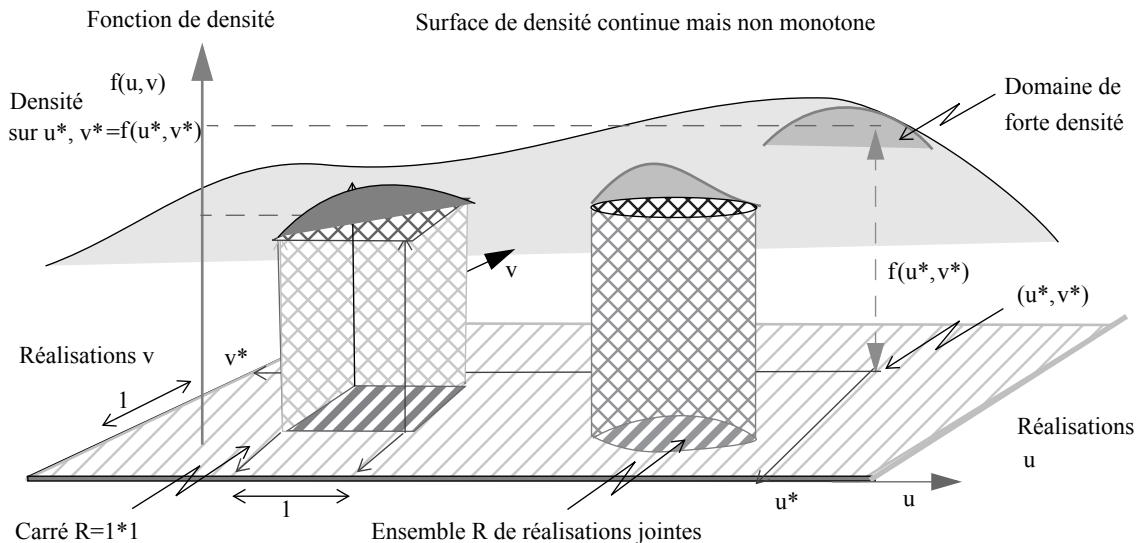
Les agents du hasard ne présentent pas qu'un seul aléa à la fois, tant s'en faut. C'est en projetant seulement des événements selon des fonctions aléatoires conjointes sur l'espace des réalisations (numériques) que l'on devient très humble en voyant combien la complexité apparaît rapidement... Aussi, on ne parlera ici que des altercations dans un ménage de deux conjoints ; les ménages à trois et plus donnent de telles partouzes aléatoires qu'il vaut mieux les confier au... hasard des rencontres.

Soit donc la densité jointe de deux variables aléatoires u et v . La Figure 33 montre des niveaux différents de la densité de probabilité, certaines zones ombrées plus sombre correspondant à des événements pour lesquels la propension à se réaliser est plus élevée que le reste de la surface. On voit ainsi une densité assez forte sur (u^*, v^*) . Elle signifie que, si on distingue un espace de réalisation (donc "par terre", ici, sur le plan $u \perp v$), le produit de la densité par ce volume donnera la probabilité associée à cet espace.

Dans tous ces cas, la probabilité associée à un espace de réalisation est la densité fois le volume. Si la constitution du volume peut se faire par composition différentielle du du et dv , il peut alors s'exprimer par l'intégration sous la fonction de densité, qu'on écrira :

$$P\{(x, y) \text{ dans } R\} = \iint_R f(u, v) du dv$$

Figure 33. Imagerie des projections de densités bivariées sur le plan



Pour un domaine de réalisation à une seule dimension, et numérique, ce serait un intervalle, par exemple de $\{u^* - du \text{ à } u^* + du\}$.

Pour un espace à deux dimensions, cela pourrait être R sur la Figure 33, disons un cercle de rayon 1 centré sur l'origine, formant donc une région définie par $u^2 + v^2 < 1$. Ce serait comme un green au golf, où le trou est quelque part dans ce cercle; chercher à atteindre une telle cible correspondrait à $P\{x^2 + y^2 < 1\}$.

La densité est donc telle que :

$$\begin{aligned} f(u, v) &= 1/\pi & \text{si } 0 \leq u^2 + v^2 \leq 1 \\ &= 0 & \text{ailleurs} \end{aligned}$$

La densité uniforme $1/\pi$ sur le cercle est donc bien conforme, puisque dans ce cas l'intégration sur le cercle est $\iint f(u, v) du dv = 1$. Ceci donne donc la probabilité de taper quelque part, mais n'importe où, dans le cercle, sachant qu'il est impossible de taper ailleurs.

À présent il faut atteindre le trou, ce qui est d'autant plus difficile qu'il est petit. Si, par exemple, le trou fait une partie S valant $1/100$ de la surface de ce cercle, la joueuse qui tape "au hasard" en téléphonant, a une probabilité de tomber dedans qui est donnée par :

$$\iint_S \frac{1}{\pi} \cdot du \cdot dv = \frac{\text{surface } S}{\pi} = \iint_S \frac{\pi \cdot (1/(100))^2}{\pi} du dv = \frac{1}{10000} \text{ où } s := (u^2 + v^2)^{1/2} \leq 1$$

Voilà pourquoi certaines blondes préfèrent taper au hasard sur le green : si elles visent le trou, elles ne l'auront sûrement pas; tandis qu'en s'en remettant au hasard...

8.2 Une différence de deux variables aléatoires

La section 8.1 s'est terminée élégamment par une considération sur les convergences en donnant le petit exemple de la somme de variables uniformes ; obéissant à une inexorable logique, cette section 8.2 en donne leur différence. Le propos, toutefois, est évidemment de portée plus générale : montrer doucement comment la composition de variables conduit à la complexité (quand celle-ci n'est pas "résumée"), et à la nécessité de grader l'individualité des processus quand on doit gérer quelque chose de hasardeux.

Formellement, soient deux points choisis au hasard dans l'intervalle fermé unitaire $[0, 1]$. Soit que les variables soient indépendantes, de même fonction de densité uniforme :

$$(38) \quad g(u) = h(u) = 1 \text{ si } 0 \leq u \leq 1, \text{ et } = 0 \text{ ailleurs}$$

La fonction de densité jointe a donc la forme :

$$(39) \quad f(u, v) = 1 \text{ si } 0 \leq u \leq 1 \text{ et } 0 \leq v \leq 1$$

$$= 0 \text{ ailleurs}$$

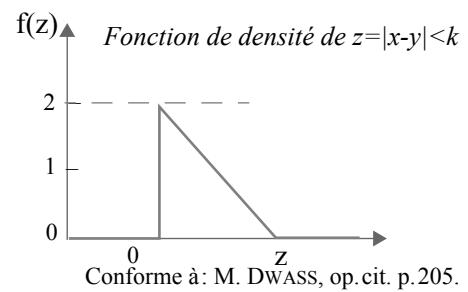
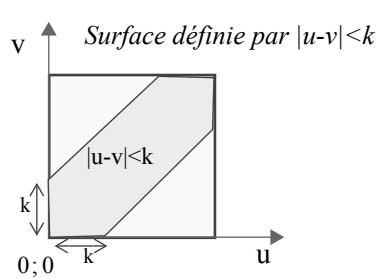
La signification de (39) au golf lapon est la suivante. La région de réalisation, le green du golf, soit un carré R de côté égal à 1 – donc de surface 1. Soit deux joueurs qui bazzotent au hasard leur petite balle sur le green, de façon indépendante. La densité uniforme sur chacune des régions est donnée par (38) pour u et v .

Quant à la fonction de densité jointe sur la situation des deux balles à la fois, elle est directement :

$$\begin{aligned} f(u, v) &= 1 && \text{pour } (0 \leq u \leq 1) \cap (0 \leq v \leq 1) \\ &= 0 && \text{ailleurs} \end{aligned}$$

Quelle est la probabilité que les deux balles soient à une distance l'une de l'autre inférieure à z ? La région de réalisation est donc telle que $|u-v| < k$. Conformément à M. DWASS, Probability and statistics, New York Benjamin, 1970, p. 205, exploité ci-dessous, cette région est représentée à la Figure 34. Amusant, n'est-il pas?

Figure 34. Région de réalisation et densité



La probabilité qui nous intéresse est l'intégrale de la densité sur cette région, soit:

$$\int_{(|u-v| \leq z)} \int f(u, v) du dv$$

La surface de cette région est $1 - (1-z)^2$. Comme la densité est uniformément 1 sur la surface du carré, cette surface définit aussi la probabilité qui y est associée, puisqu'il suffit de la multiplier par 1 pour former le volume; dès lors,

$$P(|x-y| \leq k) = 1 - (1-k)^2 \quad \text{pour } 0 \leq k \leq 1$$

Pour obtenir la fonction de densité de la différence, il convient de dériver cette fonction:

$$\frac{dF}{dk} = \frac{d}{dk} [1 - (1-z)^2] = 2(1-k)$$

Le deuxième membre est donc la densité qu'il faut intégrer pour obtenir la fonction de répartition, c'est-à-dire:

$$1 - (1-k)^2 = \int_0^z 2 \cdot (1-x) dx$$

De là on obtient la densité recherchée, à savoir la fonction de densité de la distance d'ordre 1 entre les réalisations de x et y, qui est bien la Figure 34:

$$\begin{aligned} f(z) &= 2(1-z) \text{ si } 0 \leq u^2+v^2 \leq 1 \\ &= 0 \text{ ailleurs} \end{aligned}$$

Cet exemple naïf montre que les compositions d'aléas conduisent très rapidement à des résultats qui ne sont pas intuitifs, sinon "surprenants".

La leçon à en tirer ici est que dès que l'on est en présence de plusieurs aléas, leurs comportements conjoints dans une combinaison même très simple peut devenir chaotique, et dès lors leur description déjà, et ensuite leur prédiction, puis leur maîtrise, se trouvent rapidement hors de portée.

8.3 Distribution d'un produit et d'un rapport

8.3.1 Le produit et sa transformation

Bien des phénomènes de la vie courante sont définis par un rapport de deux aléas – par exemple le "rapport de chances". Ainsi, soient deux parieurs au "turf", dont les "rapports de chances" ("odd ratios") sont respectivement de 3 sur 17 et de 5 sur 29 – cela dépend du nombre de haquenées en état de marche. Quel est le plus favorisé par la "chance"?

En fait, le rapport des deux est de 1,023 en faveur du premier, mais qui se soucierait de ces décimales dérisoires?

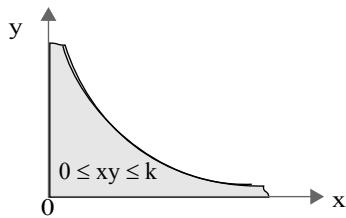
Plus généralement, soit que les probabilités intéressantes soient engendrées par le rapport de deux aléas. Soient x et y deux variables aléatoires indépendantes, de fonctions de densité f et g dans la région non-négative. On a rappelé que les probabilités associées à des intervalles sont données par les fonctions de densité intégrales, mais pour élucider ce problème à présent de façon générale, il faut utiliser le jacobien de la transformation.

On peut partir de la fonction de densité d'un produit qui s'obtient, pour $1/y$ fini, par :

$$(40) \quad h(x) = \int_0^\infty \left(\frac{1}{y}\right) \cdot f\left(\frac{x}{y}\right) \cdot g(y) dy \quad \text{et} \quad h(x)=0 \text{ dans la région inadmissible.}$$

La région d'un produit de deux variables positives borné est dessinée sur la Figure 35.

Figure 35. Région de réalisation d'un produit borné



Avec M. DWASS (op.cit. p.241), on peut donc écrire, pour $z = xy$:

$$(41) \quad P(xy \leq k) = P(0 < z \leq k) = \int_0^k du \cdot \int_0^\infty \frac{1}{v} \cdot f\left(\frac{u}{v}\right) \cdot g(v) \cdot dv = \int_0^k h(u) \cdot du$$

On applique ceci à la région R , le carré tel que :

$$\begin{aligned} f(x) \text{ et } g(y) &= 1 \text{ si } 0 \leq x \text{ et } y \leq 1 \\ &= 0 \text{ ailleurs} \end{aligned}$$

Dans ces conditions,

$$\begin{aligned} f(x/y) &= 1 & \text{si } y \geq x \text{ et } 0 \text{ ailleurs. Alors:} \\ 1/y f(x/y) g(y) &= 1/y & \text{si } 0 \leq x \leq y \leq 1 \text{ et } 0 \text{ ailleurs.} \end{aligned}$$

L'application de l'expression (40) est alors immédiate et donne :

$$h(x) = \int_0^1 \frac{1}{y} \cdot dy = -I_2 x$$

Cette fonction de densité (inattendue?) est montrée ici en bits d'information, c'est-à-dire en logarithme de base 2; on lit d'ailleurs 1 bit d'information sur la valeur $x=1/2$. On se rappelle, toutefois, qu'il s'agit ici d'une densité continue et non discrète.

8.3.2 Densité directe d'un rapport

La fonction de densité d'un rapport de deux variables positive s'écrit:

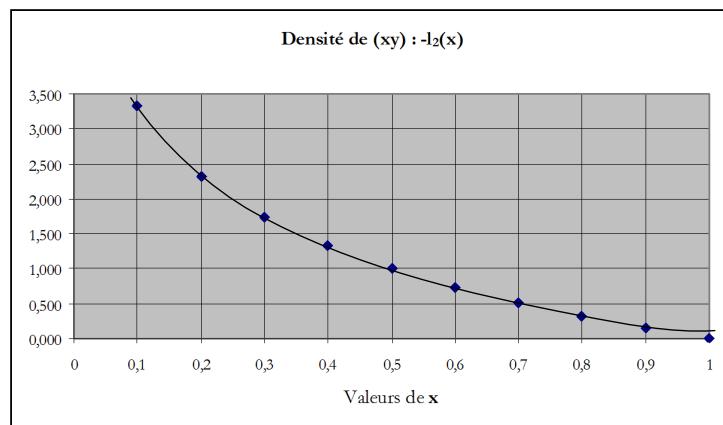
$$\mathbb{P}\{x/y \leq k\} = \int_{(x/y) \leq k} \int f(x)g(y)dx dy$$

Par un changement de variable (et son jacobien) analogue à celui du produit, on obtient la densité:

$$h(x) = \int_0^{\infty} y \cdot f(xy) \cdot g(y) \cdot dy, \text{ pour } x > 0, \text{ et } 0 \text{ ailleurs.}$$

Ceci conduit à la Figure 36.

Figure 36. Densités d'un produit dans le carré unitaire



Soit que les variables x et y sont réalisées dans un carré unitaire. Lorsque x et y sont de même densité 1 sur le carré unitaire, et 0 ailleurs, cela veut dire que les chances sont égales sur tous les sous-ensemble de même surface que l'on peut prélever sur le carré; de plus, la probabilité sera proportionnelle à la surface distinguée.

Comme, dans ce carré: $y \cdot f(xy)g(y) = y$, la densité du rapport est:

$$h(x) = \int_0^{\infty} y \cdot f(xy) \cdot g(y) dy = \int_0^1 y dy = \frac{1}{2} \text{ si } 0 \leq x \leq 1, \text{ sinon c'est autre chose.}$$

Amusant, n'est-il pas? La densité du rapport y est une constante!

De plus, ce rapport n'a pas d'espérance mathématique. En effet,

$$E\left(\frac{x}{y}\right) = \int_0^{\infty} x(dy)h(x) = \infty$$

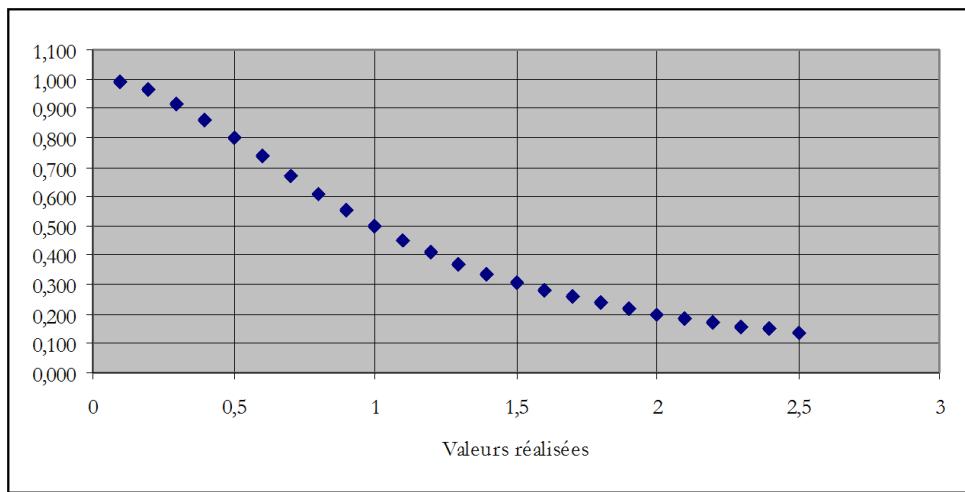
Mais il y a mieux. La densité du temps d'attente d'un (heureux ?) événement est exprimée par une exponentielle négative, $f(x) = e^{-x}$, et il en est de même pour y , soit $f(y) = e^{-y}$, pour des arguments > 0 évidemment.

Mais quelle est la densité de leur rapport? C'est plus facile ici, avec cette fois un changement de variable $y(1+x)=u$:

$$h(x) = \int_0^\infty y e^{xy} \cdot e^{-y} dy = \int_0^\infty y \cdot e^{-y(1-x)} dy = \frac{1}{(1-x)^2} \cdot \int_0^\infty e^{-u} du = \frac{1}{(1-x)^2}$$

Pour les mêmes raisons, cette densité (Figure 37) n'a pas non plus d'espérance mathématique finie, et à présent on va exploiter cette "anomalie" via la distribution de CAUCHY.

Figure 37. Densité d'un rapport d'exponentielles ($x \geq 0$)



8.3.3 Application: la distribution de Cauchy.

On peut entrer dans CAUCHY de différentes façons. Celle qui est préconisée ici est le rapport de deux variables aléatoires Normales réduites. La densité de CAUCHY s'écrit selon (42), et est représentée par la Figure 37.

$$(42) \quad f_Z(z) = \frac{1}{\pi(1+z^2)} \quad \text{pour tout l'intervalle } -\infty < z < \infty$$

On peut d'abord s'assurer qu'il s'agit bien d'une fonction de densité, dont l'intégrale sur le domaine de définition vaut bien 1. En effet, comme:

$$\int_0^t \frac{1}{1+x^2} \cdot dx = \arctan t$$

Dès lors, l'intégrale sur le domaine de définition est bien:

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} \cdot dx = \pi$$

On va montrer facilement la propriété la plus importante: si x_1, x_2, \dots, x_n sont des variables indépendantes dont la fonction de densité est de CAUCHY, $(x_1 + x_2 + \dots + x_n)/n$ a la même distribution de CAUCHY que chacune des variables composant la somme.

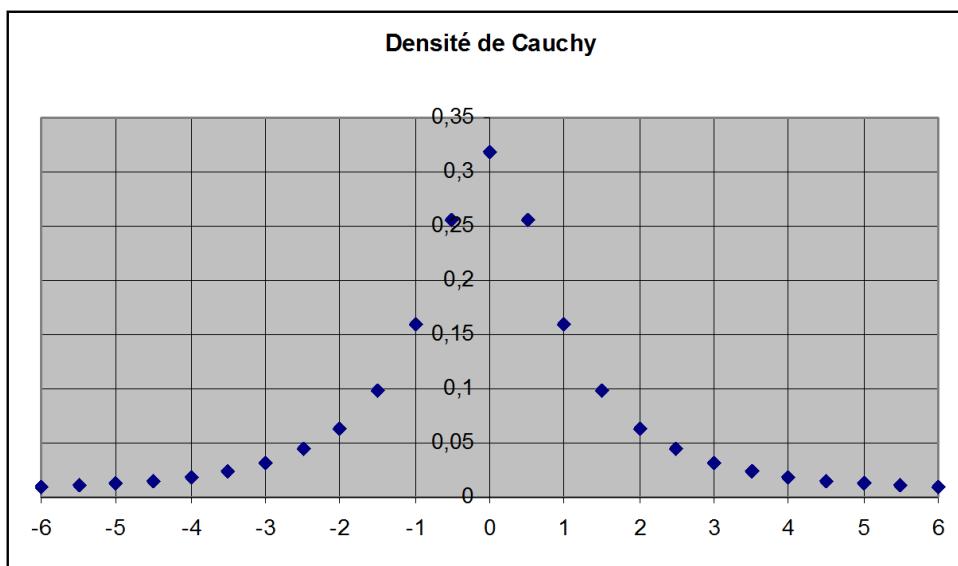
À cette fin, on utilisera la fonction caractéristique de cette fonction de densité, à savoir sa transformée de FOURIER, laquelle est unique pour toute fonction de densité, et la définit uniquement par la transformée inverse. Il suffit de glisser subrepticement un petit "i" imaginaire qui généralise la fonction génératrice de LAPLACE, $E[e^{i\theta}]$, ce qui donne

$$F[\theta] = E[e^{i\theta x}]$$

Appliqué à la densité de CAUCHY, montrée sur la Figure 38, cela donne:

$$\int_{-\infty}^{\infty} e^{i\theta x} \cdot \frac{1}{1+x^2} \cdot dx = e^{-|\theta|}$$

Figure 38. La fonction de densité de Cauchy



La transformée inverse est bien:

$$\frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} e^{-ixy} \cdot e^{-|y|} \cdot dy = \frac{1}{\pi(1+x^2)}$$

Donc, cela tient la route. Avec cet engin-là, on va avancer très vite. En effet, la transformée d'une somme de variables indépendantes est le produit des transformées :

$$F[\theta] = E[e^{i\theta(x+y)}] = E[e^{i\theta x}] \cdot E[e^{i\theta y}]$$

Dès lors (si on se souvient de ce que $i^2=-1$), la transformée de FOURIER de la somme est :

$$E[e^{i\theta(x+y)}] = E[e^{-2|\theta|}]$$

En remplaçant θ par $\theta/2$ ci-dessus, on lit que :

$$F_{\theta/2}[x+y] = e^{-|\theta|}$$

Dès lors, comme les transformations de FOURIER sont univoques, on lit que la moyenne arithmétique de deux variables aléatoires indépendantes de même densité de CAUCHY a la même distribution que chacune d'entre elles.

En généralisant l'argument sur n telles variables, on obtient :

$$E[\exp(i\theta/n(x_1+\dots+x_n))] = (e^{-|\theta|/n})^n = e^{-|\theta|}$$

Ces variables ne peuvent donc avoir la propriété d'ergodicité dont on parle ci-après.

On montre d'ailleurs que la fonction de densité de CAUCHY elle-même n'a pas d'espérance mathématique – bien qu'elle puisse donc être définie par le rapport de Normales, dont les espérances sont sans doute l'être statistique le plus exploité.

En termes de cette distribution, on peut donc écrire :

$$P\left\{\left|\frac{x_1 + \dots + x_n}{n}\right| \geq \varepsilon\right\} = P(|x_i| \geq \varepsilon)$$

Comme cette probabilité ne tend pas vers 0 quand $n \rightarrow \infty$, la "loi des grands nombres" ne peut être satisfaite. Il en résulte, par exemple, qu'une estimation fondée sur la moyenne n'est pas meilleure que l'une quelconque des mesures ou observations.

L'espérance $E(x_i)$ n'existe pas, mais la tendance centrale doit être zéro puisque la fonction de densité est symétrique par rapport à cette valeur – qu'elle n'atteint pas, y étant discontinue.

Dès lors, la "loi des grands nombres" qui a fait la statistique populaire, ne s'y applique pas non plus. La loi des grands nombres dit simplement que la moyenne est convergente en probabilité. Plus précisément, les variables indépendantes x_1, x_2, \dots, x_n de même espérance mathématique μ de même variance, et soit ε une grandeur arbitrairement petite :

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{x_1 + \dots + x_n}{n} - \mu\right| \geq \varepsilon\right\} = 0$$

Une implication que tout le monde croit connaître est l'application à la binomiale. Soit, dans une expérimentation répétée sous les mêmes conditions, que $N_i(A)$ désigne le nombre d'essais dont l'issue est dans A .

La loi des grands nombres montre que :

$$(43) \quad \lim_{n \rightarrow \infty} P \left\{ \left| \frac{N_n(A)}{n} - P(A) \right| \geq \varepsilon \right\} = 0$$

Ceci serait une justification initiale de la théorie fréquentielle de la probabilité $P\{A\}$ en tant que limite de proportions pour de suffisamment grands échantillons.

8.3.4 Processus stochastique et ergodicité

La propriété qui gère ce problème exprimé par (43) ci-dessus est celle de l'ergodicité—celle qui est primordiale en gestion des systèmes séquentiels stochastiques, où elle se comprend d'ailleurs le plus aisément (dis-je).

Soit donc $g[x(t)]$, défini sur un ensemble de valeurs $x(t)$, où le codage est fait par l'argument t pour indiquer que le truc se passe dans le contexte de séquences d'événements. Il est d'ailleurs issu du problème de la succession d'observations de processus physiques.

Pour cette brève évocation, on n'aura besoin que des trois descripteurs statistiques classiques, à savoir :

La fonction de densité intégrale: $F_{x(t)}(x) = P\{x(t) \leq x\}$

L'espérance mathématique: $\mu(t) = E[g(x(t))]$ (qui ne dépend pas de t)

La moyenne des réalisations: $M_T = 1/T \sum_t x(t)$

La fonction de covariance: $C(s,t) = E[x(s)x(t)]$

La question relève de l'inférence : sur base de l'observation d'un sous-ensemble, que peut-on avancer sur les propriétés "moyennes" de son ensemble-source, lequel est non-exhaustivement observable ?

Présentée dans le contexte de processus stochastiques, l'observation se fait soit sur un sous-ensemble discret: $\{x(t), t=1, 2, \dots, T\}$, soit sur un enregistrement d'un processus continu dans un intervalle borné: $\{x(t), 0 \leq t \leq T\}$.

Les systèmes physiques qui possèdent la propriété (44):

$$(44) \quad \frac{1}{T} \cdot \sum_{t=1}^T g[x(t)] \rightarrow E[g(x(t))] \quad \text{pour } T \rightarrow \infty$$

sont dits ergodiques.

Soit alors la séquence de moyennes M_T pour T croissant ($T=1, 2, \dots$ en discret). Une telle séquence est qualifiée d'ergodiques si la variance des échantillons tend vers zéro lorsque leur taille tend vers l'infini, ce qui s'écrit :

$$\lim_{t \rightarrow \infty} \text{Var}[M_T] = 0$$

Ce sont donc les processus dont la variance des moyennes d'observation tend vers 0 quand se poursuit l'échantillonnage qui sont ergodiques. Mais quelle est la condition de base pour obtenir cette propriété ? Selon PARZEN (op. cit; pp. 73-75), les conditions nécessaires et suffisantes s'expriment en termes de la fonction de covariance entre les moyennes atteintes :

$$C(t) = \text{Cov}[x((t), M_t)] = \frac{1}{t} \cdot \sum_{s=1}^t C(s, t)$$

La condition est que pour que la variance de M_T tende vers zéro, il suffit que la limite de cette covariance soit nulle lorsque $t \rightarrow \infty$:

$$(45) \quad \lim_{t \rightarrow \infty} C(t) = 0$$

Ceci est l'argument fondamental auquel veut arriver cette section.

En effet, (45) signifie que lorsque l'échantillon s'accroît, il y a de moins en moins de corrélation (donc de covariance) entre les moyennes d'échantillonnage M_t du processus et la dernière observation $x(t)$. On pourrait donc lire que la moyenne devient raisonnable, et ne s'accroche pas à la volatilité de toute nouvelle observation.

C'est justement ce qui ne se passe pas avec la distribution de CAUCHY, et cela paraît le gros souci de MANDELBROT, conduisant à sa proposition de "hasard sauvage". Ce dernier ne contribuerait-il pas à éviter un monde sans innovation, où tout se fondrait dans des lois, mais à développer une complexité croissante associée à de constantes nouvelles avenues de volatilité ?

8.3.5 Que fait le rapport ?

En fine, l'intuition et la formalisation peuvent se rejoindre via le rapport. Soient donc deux variables aléatoires x et y indépendantes, de distribution χ^2 (chi²) respectivement de m et n degrés de liberté. Leur rapport a une distribution F avec m, n d.l. :

$$\frac{x/m}{y/n} = F_{m, n}$$

Une χ^2 de m degrés de liberté étant la somme de carrés de m variables Normales réduites indépendantes, il est logique que la Cauchy soit le rapport de deux carrés de Normale.

De plus, si x est distribuée Normalement, χ^2 suivra la distribution Gamma avec $\alpha=1/2$ et $\lambda=1/2$.

À présent, comme elle est familière, chacun peut "voir" des valeurs distribuées selon la Normale, ainsi que leur dispersion – exprimée en nombre de fois l'écart-type par rapport à l'espérance. Quand on fait le rapport de carrés de tels écarts, il peut fort bien arriver que l'on prélève des valeurs qui soient "loin" à gauche ou à droite de la moyenne, et les divise par des grandeurs qui sont proches de zéro : il est dès lors intuitivement raisonnable que la distribution résultante soit extrêmement sensible, et que son espérance mathématique n'apparaisse pas par convergence.

8.3.6 Déchets

Une référence intéressante à propos de ce dialogue, si l'on peut dire, entre les distributions Normale et de Cauchy et celle de H. JEFFREYS, (*Theory of probability*, Oxford, the University Press) 1948, p. 244, repris (donc traduit?) par MANDELBROT (op. cit. p. 96) est : « Par ailleurs, l'étude détaillée des erreurs d'observation montre en général qu'elles sont loin de la loi de CAUCHY [...] Le point de vue du physicien vieux-jeu n'est donc pas absurde. Il comporte deux parties... dont la première est fausse et la seconde exagérée... ».

Transposant cette brillante assertion dans le cas des présents exposés, on aurait :

« L'ouvrage du systémicien vieux-jeu (CdB?) n'est donc pas absurde. Il comporte quatre Tomes (Nord, etc.) dont le premier est faux, le deuxième exagéré, et... » et les deux autres ?

9 La maîtrise du hasard

Le chanceux est celui qui arrive à tout
Le malchanceux est celui à qui tout arrive.
Georges Courteline

9.1 Se fier au hasard ?

La maîtrise du hasard est le privilège suprême d'un vrai grand dieu. Si j'en étais un, c'est le hasard que je créerais, car c'est le hasard qui s'oppose aux projets, au contrôle et à la prise de pouvoir sur des choses qui peuvent ébranler Ma Suprématie. C'est pourquoi, à défaut d'être dieu, des esprits trop téméraires tentent – en vain j'espère – d'en découdre avec le hasard.

Le rapport entre la systémique et cette maîtrise du hasard pose alors la question suivante : « Où se trouve le monde réel que l'Homme peut commander ? »

Le monde que l'on peut commander est celui que l'homme a fait pour cela. C'est un tout petit monde technologique qui a été construit pour pouvoir être exploité et s'en servir avec une bonne fiabilité de prédiction des issues des processus dont il a fait le design à cette fin : obtenir des outputs contrôlés dans des contextes et paramètres donnés, et auxquels seuls on les soumet – donc dont on a la "variété nécessaire" (disait R. ASHBY) pour en effectuer le contrôle. Ce sont quelques "minuscules univers" qui sont spécifiques à un contexte, une époque et une tâche donnée, univers auquel cette tâche est confiée : ce sont des systèmes.

Mais ce ne sont que des systèmes artificiels, "fabriqués" de façon à se comporter de façon déterministe et, s'ils s'en écartent, c'est par des aléas que l'on peut contrer soit par des adaptations, soit en les confiant dans des limites d'homéostasie qui leur permet néanmoins d'accomplir leur mission.

Lorsqu'on sort de ces petits univers dérisoires que l'on a bricolés pour en être les patrons, on se trouve confronté, parfois reçu en audience, par les "maîtres du hasard". Ce sont parfois de bons maîtres (ils se font alors appeler Providence), mais ce sont plus souvent des dieux de la nature, dont les caprices de patriciens ne se soumettent pas à nos lois des aléas dérisoires.

De plus, on parle de l'univers en expansion, mais son expansion, avance-t-on ici, se ferait "par l'intérieur", c'est-à-dire par la prolifération de la complexité et des relations. Cette expansion est peut-être plus redoutable que celle que lui attribue la physique cosmique.

D'ailleurs, les "maîtres après Dieu", les terriens qui ont voulu contrôler des peuples et leurs civilisations, en ont tout d'abord diminué la complexité. Pour mieux établir leur règne, ils ont commencé par diminuer la variété (par exemple par l'uniforme, ou les vêtements des chinois), les mœurs (telles les conventions japonaises), les relations (la hiérarchie au lieu de l'interaction), et ils ont privilégié la science cybernétique sur les sciences parlant de relations humaines.

À défaut de pouvoir le dominer, ces orgueilleux croient aussi pouvoir influencer ce hasard, ce qui n'est déjà pas si mal. Il y a bien des habitants de la mythologie (et certaines personnalités de la Bible) qui disent eux-mêmes ou ont acquis la réputation d'y avoir réussi, en des occasions exceptionnelles, appelées miracles, dont la chronique est friande. Mais il est toutefois symptomatique que, lorsqu'ils réussissent, ils appellent cela un "miracle" : cela veut bien dire que cet exploit est exceptionnel et que normalement ils ne sont pas capables de le répéter.

Ne pouvant influencer soi-même le hasard, certains tentent d'intercéder auprès de ceux qui pourraient le faire ; ils rétrogradent alors au niveau du Grand Prêtre. Celui-ci doit pour cela se ruiner (ou ruiner les autres) en offrandes, faire (ou plutôt imposer aux autres) des sacrifices cruels, se prosterner en des positions déférantes et humiliantes. Mais, plus proche du Ciel, et pouvant y lancer des signaux (souvent de la main, comme le séma-
phore), c'est de haut, tout de même, que le Grand Prêtre toise les vulgaires citoyens.

Ces vulgaires citoyens, quant à eux, ne peuvent que solliciter à grands frais les Grands Prêtres de prier ceux qui intercèdent auprès de ceux qui peuvent influencer le Hasard, et de leur faire part des résultats.

Plus bas encore dans l'échelle, il reste ceux qui n'ont même pas le droit de les en prier, car ils n'ont que des axiomes et pas de dogme – ce sont les probabilistes.

Et sous eux, sous leur fumier, sans axiomes, sans dogmes et sans foi, et privés même de l'oreille d'une providence qu'ils nient, il y a ceux que l'on méprise, traite d'escrocs de vérités, de faussaires d'opinion, de fabricants de chiffres illisibles et non-crédibles – ce sont les statisticiens.

Il est faux que «les voies de Dieu sont impénétrables» ; Jésus-Christ, ses poursuivants (de son œuvre), et des cohortes de gens de diverses Églises se sont escrimés à expliquer (à leur façon) comment il faut faire (et surtout comment il ne faut pas faire) pour espérer les emprunter. Mais les voies du hasard, elles, sont d'un accès difficile, car si "le hasard fait bien les choses", c'est tout un art de bien faire les choses au hasard.

Et il ne faut pas croire au destin tout écrit dans des étoiles; celles-ci sont des stars très pudiques et qui ont tellement honte de ce qu'elles offrent aux hommes qu'elles le cachent derrière des couches de nuages. Quelques voyageurs du Destin – qui se font appeler voyants – les regardent avec d'immenses jumelles, des télescopes qu'ils incrustent dans des créneaux du ciel, crevant la couche d'eau jaune qui baigne la terre.

Il n'y a pas lieu non plus de croire à la prédestination des prénoms. Ainsi la grande courtisane, la call-girl des rois et l'espionne batave de luxe, la troublante MATA-HARI, dite Margarita ZELLE, s'appelait GERTRUDE.

9.2 Les voies du hasard et leurs commandes

Comme ces exposés sont très pratiques et très orientés vers les recettes au lieu des idées, voici quelques moyens par lesquels les citoyens peuvent solliciter ceux qui intercèdent pour influencer le Hasard, selon les besoins et les infortunes :

- Pour retrouver un objet, et éviter la contagion : saint ANTOINE;
- Se protéger des accidents de voyage et les intempéries : saint CHRISTOPHE;
- Dans les causes désespérées : sainte RITA;
- Contre une voisine enquiquinante : la photo de saint EXPÉDY au-dessus de la porte;
- Pour obtenir un enfant à bref délai (mais après une bonne sieste au Club MED) : faire la procession de saint MAXIMIN de Le Dorat (56 km);
- En vue de la fortune (par la santé et le travail) : santo PANCRACIO;
- Un accouchement moins douloureux : sainte MARGUERITE;
- Pour trouver un mari dans l'année : caresser le pied gauche du gisant de la rue des Pierres à Bruxelles. A ce sujet, on sait comment ont disparu les bras de la VÉNUS de MILO : elle flanquait des baffes à ceux qui lui caressaient la fesse gauche pour obtenir quelque chose que la tradition n'a pas transmis. Lassés des beignes sur la gueule, les mecs lui en ont coupé les moyens;
- Contre les convulsions : saint GHISLAIN;
- Pour que l'argent rentre, s'adresser à saint GÉRARD, mais cela ne concerne que les pérégrinées (une messe dédicacée à cette fin a lieu à Liège le mercredi);
- En Bretagne, invoquer sainte ANNE contre la pauvreté, saint BENOÎT contre les méléfices et la fièvre, saint FIACRE contre les hémorroïdes, saint FRANÇOIS pour préserver de la peste, saint GEORGES contre les maladies de la peau; les spasmes sont écartés par saint Jean, les voleurs par saint NICOLAS, les morsures de serpents par saint PAUL, et la réussite en affaires est affaire de saint PIERRE.
- Obtenir des vents favorables : sacrifier sa fille (IPHIGÉNIE, fille d'AGAMEMNON);
- Obtenir un destin exceptionnel : sacrifier son fils (ABRAHAM va trucider ISAAC);
- Se refaire une santé : faire trempette avec Notre Dame de Lourdes;
- Enfin, on peut conjurer le mauvais sort en touchant du bois, et mettre l'aléa de son côté en croisant les doigts, mais dans ces cas il faut surtout éviter les saints STATISTICIEN et SYSTÉMICIEN...

