Les probabilités boréliennes et la certitude exprimée par des nombres (formellement) différents de l'unité

M. Matschinski

LES PROBABILITES BORELIENNES ET LA CERTITUDE EXPRIMEE PAR DES NOMBRES (FORMELLEMENT) DIFFERENTS DE L'UNITE.

M. MATSCHINSKI.

Summary.

The absolute certitude in the classical Probabilities and in the Borelian one. Four methods to express mathematically the fundamental Borelian concept : 1. Characterization of the absolute certitude by the « very great number » as it is in the especial kind of the Mathematics in which the infinity is replaced by the « very great ». 2. Consideration of the numbers, the difference of which is smaller than the exactitude of our measurements, as « equals ». (These two possibilities are rather only mentioned ; the two following are considered more in detail). 3. Introduction into the Theory of Probabilities of the scheme of a drawn (the patterns of a urn), in which some proprieties of the content of urns will be known after a finite number of the drawns. 4. The last possibility to base the Borelian concept is connected with the substitution of the fundamental axiom of Probabilities : « it is always possible to repeat the independent measurements of the same kind and even the infinite quantity of independent one », through the new real axiom (given here) : « it is not possible to repeat the independent measurements of the same kind in the infinite quantity, but only in the finite one ».

INTRODUCTION.

En probabilités classiques on a insisté beaucoup sur la différence primordiale entre la certitude dite « absolue » (= 1 exactement) et la probabilité « très grande » (= $1-\epsilon$, où ϵ est très petit, par exemple 10^{-20}, 10^{-40}, etc., mais jamais égal à zéro). On a dit dans ce sens que l'impossibilité d'un «miracle dactylographique » (singe frappant fortuitement sur une machine à écrire et produisant de cette manière l' «Iliade » d'Homère) où « miracle physique », (brique montant en l'air sous l'influence de forces moléculaires) est « très probable » , mais non pas « certaine ». Contrairement à ceci le regretté Emile Borel a insisté, dans sa dernière œuvre /1/ , sur l'identité de ces deux notions : il pense par exemple que $1-10^{-40}$ *est* la certitude. Il a exposé plusieurs raisons pour justifier cette affirmation, très profondes et très curieuses, mais il n'a pas donné (peut-être parce qu'il n'en a pas eu le temps : le livre cité a été publié peu avant sa mort) de formes mathématiques dans lesquelles cette idée pût être décrite. Sans prétendre ressusciter la pensée du grand disparu, on indiquera ici quatre possibilités de donner à l'idée en question un habillement mathématique.

Mais, avant même toute argumentation mathématique ou philosophique, il faut observer ici que, si l'impossibilité non absolue - c'est-à-dire la «possibilité» - des briques volantes est tant bien que mal acceptable, il est difficile de trouver un esprit sain et clair capable de concevoir une « possibilité » , même minime, que l' « Iliade » soit composée involontairement par un singe. Mais, comme nous l'avons déjà dit, on n'entrera pas ici dans les raisons générales - priant le lecteur de se référer par exemple au livre cité de Borel - c'est uniquement, des formes mathématiques que peut revêtir l'idée de la certitude non égale à l'unité qu'il est question ici.

§ 1. La première possibilité ne sera que mentionnée. Elle est très générale, mais sort décidemment du cadre de la théorie des probabilités. C'est la possibilité de bâtir une mathématique où les nombres « très grands » jouent le rôle de l'infini de la mathématique classique. Il existe plusieurs essais de ce genre.

Une telle mathématique appliquée aux probabilités remplacera l'infinité des cas favorables par un « très grand nombre » (« égal » , dans cette mathématique, à l'infinité) de chances. Particulièrement là où l'infinité des cas favorables donne - dans la théorie classique des probabilités - la certitude (« l'unité »), un « très grand nombre » de ces cas favorables ne donnera qu'une quantité « un peu » moindre que l'ùnité : certitude inégale, en général, de l'unité. Cependant ici ce n'est pas la seule théorie des probabilités qui est transformée, mais la mathématique entière avec toutes ces branches. Si curieuse que soit cette éventualité, elle n'est pas, et de loin, l'objet propre à la théorie des probabilités on la laissera de côté ici.

§ 2. La deuxième possibilité est basée sur le point de vue, d'ailleurs assez répandu, qu'on ne peut pas (et par conséquent qu'il ne faut pas) distinguer des nombres dont la différence est au-dessous de l'exactitude de nos mesures les plus précises. Plusieurs passages de l'œuvre de Borel /1/ incitent à croire que ceci a été l'idée (ou, au moins, une des idées) fondamentale pour lui. Cependant - de même que la première possibilité est généralement mathématique, mais non pas spécialement probabiliste - cette deuxième est plutôt physique, et, en dépit de toute son importance aux yeux de certains, n'apporte rien d'essentiel ni de nouveau pour la théorie des probabilités. *).

*) - Ce n'est pas cette *deuxième* possibilité qui est l'essence de l'œuvre de Em. Borel /1/. Dans ce domaine il a eu tellement de prédécesseurs qu'il est assez difficile de les mentionner tous. Pour ne prendre sur nous-mêmes la responsabilité de choisir les uns et d'omettre les autres, limitons nous à citer M. Frechet qui écrit : « En ce qui concerne l'assimilation *pratique* des probabilités infiniment petites d'une affirmation quelconque et de la certitude de l'affirmation opposée, Borel, probablement sans le savoir, a été précédé par Buffon, puis Cournot ». Pourtant, ce n'est pas - repétons-le - seulement ces deux noms qu'il faudrait citer si l'on veut traiter exhaustivement l'histoire de cette deuxième possibilité.

De plus, d'un point de vue rigoureusement formel, cette éventualité n'est qu'apparente. En effet, du fait que *pratiquement* on ne peut jamais atteindre la certitude absolue, - et il faut se contenter de l'approximation d'une certitude proche de l'unité, mais pas égale à cette dernière, - il ne découle point que la certitude absolue «change» sa valeur qui diminue, et devient plus petite que cette, classique « unité » .Nier (sur la base de l'expérience et de l'impossibilité expérimentale d'obtenir la valeur «exacte» de l'unité ou exactement l'« unité » pour une grandeur quelconque) l'existence théorique de l'unité arithmétique ou algébrique, c'est quitter définitivement le sol des mathématiques et passer dans le domaine des sciences purement expérimentales ; or, il y a actuellement très peu de sciences qui veulent être et restent purement expérimentales. Ainsi la deuxième possibilité, étant admissible d'un point de vue général, n'est pas à reconnaître pour une éventualité mathématique ou spécialement probabiliste. Contrairement à celles-là les deux autres possibilités sont profondément liées aux notions statistiques et probabilistes ; la troisième l'est moins, la quatrième l'est davantage.

§ 3. Les applications classiques de la théorie des probabilités sortent de l'image du tirage dans une urne, « machine » à propos de laquelle on ne sait rien, sauf qu'on peut en tirer des boules et (dans certains cas) les y remettre. Cependant on peut avoir une autre vision du monde des expériences, matérielles et spirituelles, et on peut, en conséquence, partir d'un autre modèle d'urne symbolisant ces expériences. Sans croire épuiser les possibilités logiques, nous nous bornerons ici à proposer un modèle dont la description suit. Supposons notamment que l'urne contient des boules dont nous voulons connaître («mesurer») une « propriété » : le poids, le diamètre, la longueur d'onde corespondant à la couleur, ou simplement un chiffre inscrit sur chaque boule. Supposons qu'il y a N sortes de boules et qu'on connaît la valeur numérique N. Chaque sorte peut contenir plusieurs boules (ayant identiquement la même « propriété ») ou une seule ; cela - quoique étant important en général - ne l'est pas pour le problème en question. Introduisons ici une propriété spéciale de notre modèle à savoir que, la composition de l'urne et les propriétés des boules étant absolument aléatoires, seul le nombre de boules ne peut être infini. On tire alors une boule, on « mesure » - on prend connaissance - de sa « propriété » et on la remet dans l'urne. On répète cette opération. On calcule quelque fonction qui nous est nécessaire, par exemple la moyenne, des résultats acquis. Après un petit nombre de tirages, on ne connaît cette moyenne que très imparfaitement; mais notre connaissance s'améliore au fur et à mesure qu'on tire plus de boules. Toutefois pour l'expérience fondamentale classique (où non seulement on ne connaît pas la valeur de N, mais où l'on suppose qu'elle peut atteindre l'infini) on n'arrive jamais à la certitude. Jusqu'ici tout va relativement bien. Mais si le nombre de boules est fini, et si, de plus, le nombre des tirages effectués le dépasse largement, on arrive à une connaissance de la moyenne (ou de toute autre fonction des résultats expérimentaux) dont la probabilité est « très élevée », quoique toujours inférieure à 1. C'est justement la situation qui conduit à certains non-sens logiques, situation critiquée par Borel.

§ 4. Revenons donc au modèle non-classique proposé. Distinguons deux genres de moyennes de « propriétés » des boules (ou des fonctions de ces propriétés). Soit n_i, le nombre des boules de la sorte « i » ($i=1,2, \ldots N$), ces valeurs n_i ne nous sont pas connues. La grandeur p_i étant une propriété (ou une fonction de cette propriété) des boules de la sorte « i » :

$$(1) \qquad \overline{p_i} = \frac{1}{\mathcal{N}} \sum_{1}^{N} n_i \, p_i$$

(elle caractérise la composition de l'urne), tandis que la moyenne caractérisant non plus cette composition, mais la propriété moyenne des sortes (indépendamment de leurs nombres relatifs) sera :

$$(2) \qquad \widetilde{p_i} = \frac{1}{\mathcal{N}} \sum_{1}^{N} p_i$$

Dans la théorie classique toutes ces moyennes restent « connues » seulement avec la certitude « très élevée », mais inférieure à 1.

Par contre, si on prend, pour l'expérience fondamentale, celle du modèle non-classique proposé, où l'on connaît la valeur numérique de N, on arrive à des situations tout à fait autres que la situation classique bien connue. En effet si la situation au commencement est la même - on ne sait que très peu de choses après les premiers tirages, et on améliore ses connaissances avec les tirages suivants - tôt ou tard (après un *nombre fini* de tirages) on peut atteindre la situation où, ayant tiré au moins une boule de chaque sorte, on connaît la valeur moyenne $\widetilde{p_i}$ cherchée (ou toute autre fonction pareille) avec *certitude* (avec une probabilité strictement égale à 1). Au contraire, si après un nombre de tirages même très considérable on ne connaît pas encore les propriétés de toutes les sortes de boules (parce que par exemple on n'a observé que N-1 sortes) on sait (également avec *certitude*) que la moyenne $\widetilde{p}_i$ ne décrit pas encore l'ensemble entier. Cette dernière remarque n'est ni triviale, ni vide de sens parce que justement dans le modèle classique (où N n'est pas connu) on peut arriver à cette situation bizarre - sans le savoir - qu'on connaît la moyenne exactement (au moins pour un instant, les expériences suivantes introduisant de nouveau des inexactitudes), mais qu'on estime sa probabilité inférieure à 1. Ainsi, si l'on s'impose d'avance d'accomplir un nombre N_1 d'expériences (N_1 fini, $N_1 \gg N$), on aura dans l'écrasante majorité des cas la valeur de la moyenne avec toute certitude (probabilité $\equiv 1$) ; tandis que dans une infime minorité des cas on aura la certitude qu'on ne connaît encore pas cette valeur. Mais jamais on n'arrivera aux nombres insensés (d'après Borel) du genre $1\text{-}10^{-40}$ ou $1\text{-}10^{-35}$.

Pour les moyennes $\overline{p_i}$, la situation n'est évidemment pas si simple. On ne peut estimer les résultats avec une certitude absolue que dans le cas négatif. En effet, si, en dépit de $N_1 \gg N$, à un moment quelconque les N sortes ne sont pas encore toutes apparues, on sait avec certitude (« absolue ») que la moyenne, $\overline{p_i}$ n'est pas exacte, en toute analogie avec $\overline{\overline{p_i}}$. Ce n'est que dans le cas de la détermination de $\overline{p_i}$ en présence de toutes les sortes apparues, que l'on ne connaît pas encore cette dernière moyenne, de même que dans le cas classique. Mais même ici la certitude que l'on connaît effectivement la valeur exacte de p_i est, - pour le cas du modèle non-classique et pour celui des boules des N sortes apparues, - beaucoup plus grande que dans le cas classique (voir /2/). Pourtant cette façon de présenter la troisième interprétation n'est qu'une simplification extrême, même pas la première approximation. On trouvera ailleurs de plus amples détails (voir /2/).

Telle est la troisième possibilité. N'étant ni de la mathématique générale ni essentiellement du domaine des sciences dites exactes - voir § 2 - elle est, il faut le reconnaître, liée à la théorie des expériences, qui, si importante soit-elle, et aussi proche des problèmes statistiques qu'on peut le désirer, n'est point la matière propre du calcul des probabilités. La *quatrième* et dernière possibilité, à l'exposé de laquelle nous passons maintenant, n'est liée qu'à un axiome fondamental des probabilités. Jamais cet axiome, à notre connaissance, n'a été exprimé explicitement, en dépit du fait que tout l'édifice de la théorie classique ne tient que par lui.

§ 5. Commençons par une analogie. Dans un espace tridimensionnel, la détermination d'une distance nécessite - dans le cas général - trois mesures non coplanaires. Toute autre - quatrième, cinquième, etc. - est inutile. De même, pour n dimension il faut n mesures, et elles suffisent. Ainsi dans l'espace « de phases » de mesures - répétées pour arriver à une exactitude plus grande et à une probabilité plus élevée - les mesures sont supposées (- dans la théorie classique -) indépendantes. Ceci correspond dans notre analogie à l'hypothèse que leur espace « de phases » a une infinité de dimensions. Cependant aucune expérience ne peut nous dire quoi que ce soit à propos d'une infinité. Tout jugement de cette sorte n'est qu'une conjecture, ou, en langage mathématique : un axiome. Ainsi à l'axiome classique : *les mesures répétées de même genre peuvent être indépendantes, même s'il y a une infinité de mesures* , on peut opposer un axiome réel : *les mesures répétées de même genre ne sont indépendantes qu'en nombre fini* ; à partir d'un certain nombre (différent, peut-être, pour différents cas), toute mesure suivante dépend entièrement des précédentes.

A ces deux axiomes, sans changer leur sens, on peut donner des formes légèrement différentes de celles que l'on vient de citer. Le choix entre ces deux formes ne dépend que du goût. Nous les citons toutes dans l'espoir qu'ainsi la compréhension de notre idée sera facilitée.

Axiome classique : Il est possible de réaliser indépendamment des mesures répétées de même genre, et ceci même s'il y a une infinité de mesures.

Axiome réel : Il n'est possible de réaliser indépendamment des mesures répétées de même genre que si elles sont en nombre fini.

De cette façon, à partir d'un certain nombre d'expériences, on n'améliore plus les résultats : ils sont définitifs, et il leur correspond une probabilité égale à 1, c'est-à-dire la *certitude*, en dépit du fait que la probabilité calculée formellement n'atteint pas l'unité. Ici aucune hypothèse sur le genre des tirages - comme celle du 3° - n'est désormais nécessaire. On arrive, après un nombre fini de tirages, expériences, observations, etc. à des résultats définitifs, parce qu'aucune série infinie de mesures indépendantes n'est possible (axiome réel).

Ainsi en prolongeant une série de mesures au delà des limites où ces mesures cessent d'être indépendantes, on n'a plus le droit d'utiliser le théorème de multiplication des probabilités sous sa forme élémentaire (qui correspond aux probabilités indépendantes). Ainsi, par exemple, si la probabilité individuelle est μ, nombre assez petit ce n'est que pour des séries de mesures répétées assez, courtes - une centaine peut être - que la probabilité d'une centaine de répétitions est - plus ou moins - égake à μ^{100} et la probabilité du contraire est $1 - \mu^{100}$, mais au fur et à mesure que l'on multiple les répétitions, leur probabilité ne se forme plus au moyen de multiplications par μ ,mais descend ěventuellement plus vite, de même que la probabilité $1 - \mu^n$ (n, nombre de mesures dépassant la limite des répétitions « indépendantes »). s'approche du 1 d'une façon accélérée et l'atteint pour n fini, quoique très grand. C'est de cette manière qu'à la probabilité formelle $1\text{-}10^{-35}$ correspond une probabilité réelle égale à 1, c'est-à-dire la certitude.(Pourtant cet alinéa présente un certain compromis, l'axiome « réel » sous la forme rigoureuse dit beaucoup plus).

§ 6. Les interprétations des raisons pour lesquelles « l'axiome classique » doit être remplacé par « l'axiome réel » - en général et particulièrement là où le nombre n est grand - peuvent être très variées. On peut naturellement revenir à l'une des hypothèses déjà énoncées ci-dessus. Même si ces hypotheses pouvaient tant bien que mal nous servir dans ce nouveau but, cette opération serait assez inutile, parce que les dites hypothèses, chacune à elle seule, expliquent bien la situation, et les introduire simultanément avec l'axiome du § 5 ne serait qu'entasser des hypothèses parallèles. Parmi les hypothèses propres à « l'axiome réel » on peut mentionner en premier lieu l'existence de liaisons entre les grands ensembles de phénomènes individuels étudiés. Supposons par exemple qu'entre les 10^{30} phénomènes étudiés existent des liaisons dont le nombre augmente très vite quand on passe de 10^{30} à 10^{31} etc. phénomènes. Ainsi nos phénomènes au nombre de 10^{10} sont pratiquement indépendants, tandis que, pris au nombre de 10^{33} ou 10^{35}, ils ne le sont plus. « L'axiome réel » serait une conséquence immédiate d'un tel état de choses. Cette interprétation n'est point seule dans son genre, dans d'autres publications nous considérons plusieurs autres hypothèses. L'appareil mathématique correspondant est développé dans notre Note /3/.

Il est évident que dans toutes les considérations de la Note présente et dans toutes leurs conséquences, il faut soigneusement distinguer les probabilités établies théoriquement de celles établies en pratique. Il ne s'agit ici en principe que d'établir (- de faire un « calcul » -) des probabilités théoriques; aucune possibilité n'existe de mesurer quelque chose de comparable avec ces probabilités $1\text{-}10^{-35}$ et autres pareilles que l'on traite ici (voir Borel /1/).

Pourtant rien ne nous empêche, - quoique ici notre pensée diverge de la borelienne, - de considérer théoriquement l'éventualité des « mesures » (pratiquement « irréalisables »). Même ici des interprétations de ces « mesures » qui donnent des résultats hypothétiques, restent possibles.

De même que ci-dessus, nous nous bornerons à n'en mentionner qu'une seule : « quantisation des probabilités » . En effet il est possible de supposer que non seulement $1\text{-}10^{-35}$ *est* la certitude, mais aussi toute probabilité P (non égale à l'unité) ne se distingue pas de la valeur $1\text{-}10^{-35}$ et d'autres semblables. Ainsi au lieu d'une variable P continue, il n'existera qu'une série de o, P_1, P_2, . . ., très longue, mais finie, des valeurs qui « sont » réellement ; toute autre valeur (entre deux P de la série citée) n'a qu'une existence « apparente » , formelle ayant exactement le même sens que P_i voisin. L'auteur reprend le détail de cette question dans son essai /2/.

Comme pour la plupart des phénomènes de la nature, « la rectification » des axiomes classiques est possible dans les deux directions. On peut penser que la position classique se présente toujours comme un « juste milieu » entre deux « extrêmes ». Dans ce but il suffit de rappeler la géométrie classique euclidienne (espace « sans courbure ») se trouvant entre la riemannienne (espace à courbure positive) et la lobatschevskienne (courbure négative). Il en est exactement de même pour notre présent sujet. Ainsi dans la théorie classique des probabilités on admet l'existence des phénomènes pratiquement impossibles, mais théoriquement admissibles, tandis que la théorie borelienne nie l'admissibilité, même théorique de tout ce qui est pratiquement impossible (le contraire a, dit-on, la certitude absolue). Enfin, on peut construire la théorie « lullienne » qui affirmera que tout ce qui est théoriquement possible, l'est pratiquement. Quand on utilise ici l'expression « lullienne » , on pense évidemment à Raymond Lullius, à son « Ars magna » et à sa « machine logique » dont le fonctionnement « réel » suppose l'acceptation de l'affirmation antiborelienne que l'on vient de formuler .

BIBLIOGRAPHIE

[1] BOREL (E.) Probabilité et certitude. Paris 1956.

[2] MATSCHINSKI (M.) Sur les systèmes de probabilités, borélien, classique et antiborélien (sous presse).

[3] MATSCHINSKI (M.) Nombre des séries . . etc. - *C.R. Acad. Sci. Paris* t. 252, p. 3799,. 1962.

Reçu en Décembre 1970.

41, Rue Albert Peuvrier

91240 - SAINT-MICHEL-SUR-ORGE