L’ANALYSE LOGIQUE DES PROBABILITÉS SELON WAISMANN

KARIM BELABAS ET LAYLA RAÏD

TABLE DES MATIÈRES

1. Introduction . . . 1
2. Critique de la conception fréquentiste . . . 3
2.1. Waismann *versus* von Mises . . . 3
2.2. Le dé pipé . . . 4
2.3. Qu’est-ce qu’un *Kollectiv* ? . . . 5
2.4. Critique de l’idéalisation . . . 8
3. La conception logique de Waismann . . . 9
3.1. Préambule . . . 9
3.2. Le propos de l’axiomatique de Waismann . . . 10
3.3. Spielraum . . . 10
3.4. Mesure . . . 11
3.5. Probabilité relative, ou conditionnelle . . . 11
3.6. Rapport à l’axiomatique de Kolmogorov . . . 12
4. Une défense programmatique . . . 13
4.1. La probabilité logique face aux applications . . . 13
4.2. La loi des grands nombres . . . 14
5. Conclusion . . . 16
Références . . . 16

1. INTRODUCTION

Le calcul des probabilités se développe depuis le 17^e siècle — de Pascal, puis Jacob Bernoulli à Laplace, on les pense à partir de la combinatoire, comme quotient du nombre de cas favorables sur celui des cas possibles, souvent supposés équiprobables par symétrie, à l’exemple canonique du dé à six faces. Mais ce n’est qu’en 1933 que les probabilités obtiennent des fondations axiomatiques satisfaisantes, avec la parution des *Grundbegriffe der Wahrscheinlichkeitsrechnung* [7] du jeune mathématicien russe Kolmogorov (il a 30 ans) ; elles mettront une vingtaine d’années à s’imposer, et ne seront traduites en anglais qu’en 1950.

Jusqu’au début du 20^e siècle, l’usage universel du calcul des probabilités (des sciences de l’homme et de la nature à l’ingénierie) est privé de fondations solides.

Dans son fameux discours de 1900 au Congrès International de Paris, parmi ses 23 problèmes destinés à façonner le développement des mathématiques pour le siècle à venir, Hilbert inclut la construction de fondations pour la physique (6^e problème). Il précise aussitôt le caractère prioritaire des branches les plus proches des mathématiques : probabilités et mécanique. Toujours au tournant du siècle, Poincaré [8, Chap. XI] insiste lui aussi, par delà les récréations combinatoires du joueur de casino, sur le « rôle considérable » des raisonnements probabilistes et des tests statistiques dans les sciences de la nature. Aussi déplore-t-il, dans sa préface à la *Science et l'hypothèse*, leur état d'inachèvement, demandant qu'on en approfondisse les principes :

> Sous ce rapport, je n'ai pu donner que des résultats bien incomplets, tant ce vague instinct, qui nous fait discerner la vraisemblance, est rebelle à l'analyse. [8, p. 27]

Comme on le voit, il n'est pas clair à l'époque que les probabilités relèvent des mathématiques, ni qu'il soit possible de leur donner une forme rigoureuse.

Durant trois décennies, ce « vague instinct » est l'objet de multiples tentatives d'affûtage, dont la plus importante mathématiquement avant Kolmogorov est due à von Mises [10]. Mais celle-ci restait confrontée à des difficultés majeures[1]. L'article de Waismann, « Logische Analyse des Wahrscheinlichkeitsbegriffs » [12], propose l'ébauche d'une axiomatique concurrente[2], sous l'influence revendiquée du *Tractatus logico-philosophicus*.

Notre objet sera double. Dans une perspective d'histoire des sciences, nous présenterons cette axiomatique tant dans ce qu'elle anticipe des développements futurs — comment elle se fonde à juste titre sur la donation d'une mesure *arbitraire* —, que dans ce qu'elle laisse inachevé. Dans une perspective philosophique, nous montrerons comment Waismann défend, contre toute interprétation empirique, l'aprioricité du calcul des probabilités dans une perspective tractatuséenne. Dans ce cadre, il s'élève en particulier, de manière problématique comme nous le verrons, contre les définitions de la probabilité comme limite d'une fréquence relative.

Un des points remarquables de cet article tient dans l'interaction entre analyses philosophiques et mathématiques de la probabilité. À l'époque où Waismann intervient, l'accord scientifique sur ce sujet n'est pas acquis. Aussi le philosophe interroge-t-il les fondements formels des probabilités, en sus de leur statut épistémologique, alors que seul le second est depuis Kolmogorov affaire philosophique véritable, si on entend par probabilité l'objet de la théorie mathématique actuelle[3]. Cette configuration historique particulière conduit Waismann à défendre un article

[1] Sur ces difficultés et les corrections apportées à cette axiomatique, en conservant les intuitions de von Mises, *cf.* van Lambalgen, « Von Mises' Definition of Random Sequences Reconsidered » [9].

[2] L'article fut présenté dans la section sur la philosophie des probabilités de la *Erste Tagung für Erkenntnislehre* (Prague, 1929). Les actes furent publiés comme premier numéro de *Erkenntnis*. On y trouve également, sur les probabilités, des contributions de Feigl, von Mises et Reichenbach.

[3] Des concepts différents et non concurrents existent, *cf.* par exemple la probabilité logique de Carnap qui, contrairement à celle de Waismann, n'a pas prétention universelle.

hybride, contenant à la fois une ébauche d'axiomatique et une défense philosophique de sa thèse principale : la théorie des probabilités est une « branche de la logique ». L'histoire des sciences connaît de tels hybrides en particulier dans les périodes précédant la définition de paradigmes acceptés par la communauté des chercheurs. Aussi suivrons-nous chez Waismann cet entrecroisement, en montrant comment l'inspiration aprioriste du *Tractatus* modèle sa réflexion.

L'article de Waismann suit le plan suivant. La première partie critique deux conceptions de la probabilité : la première, tripartite et soutenue dans les seuls « cercles philosophiques », affirme le caractère primitif du probable, en sus du vrai et du faux. L'estimant moins importante dans l'économie de l'article, nous la laisserons de côté, pour considérer directement la seconde conception, définissant les probabilités comme fréquences relatives : Waismann lui oppose plusieurs critiques, en rejetant en particulier les axiomes de von Mises. La deuxième partie de l'article expose les fondations logiques proposées par Waismann lui-même. La troisième, enfin, est une défense programmatique : il annonce pouvoir retrouver les « axiomes » de Keynes, puis propose une explication du lien entre sa probabilité et l'idée de fréquence relative.

2. Critique de la conception fréquentiste

2.1. Waismann *versus* von Mises. Avant de présenter sa conception propre des probabilités, Waismann commence par les difficultés d'une certaine conception fréquentiste, qu'il appelle aussi « statistique », identifiant la probabilité d'un événement avec sa fréquence relative dans une longue série d'expériences. Cette idée a été axiomatisée en particulier par von Mises, mais ce n'est pas à celui-ci que s'attaque d'abord Waismann, qui commence par en proposer une version naïve préscientifique. Selon cette théorie, explique-t-il, « la probabilité d'obtenir 2 avec un dé est de 1/6 » signifie « sur un nombre de lancers suffisamment grand, le nombre 2 apparaîtra en moyenne dans un résultat sur six » — la naïveté résidant dans l'usage de l'expression vague « suffisamment grand », et dans une prédiction empirique qui a toutes les chances d'être fausse (une proportion valant *exactement* 1/6 est très peu probable). Tout le propos de von Mises sera de donner au fréquentisme une formulation rigoureuse à travers sa notion de *Kollectiv*. Waismann entend en donner l'idée maîtresse à travers cette première formulation naïve : il niera ensuite que le raffinement que von Mises en propose soit satisfaisant.

Cette idée naïve évite au moins d'emblée, remarque ironiquement Waismann, le problème de l'application : la question « Comment la probabilité peut-elle prédire le cours des événements ? » ne se pose pas, car l'énoncé de probabilité a directement pour source et pour objet des faits d'expérience (on observe ce qui se produit quand on jette un dé « assez longtemps »), ce qui évite le gouffre conceptuel entre calcul *a priori* et faits empiriques, dont la modélisation pourrait paraître mystérieuse. Mais, loin de résoudre ce problème, l'empirisme du fréquentisme naïf ne permet pas seulement de penser l'idéalité mathématique tout court. Devant les difficultés de von Mises, Waismann n'admettra finalement pas qu'il ait dépassé ce stade naïf :

il présente la critique du *Kollectiv* dans la foulée énumérative de celle que nous venons de voir.

Il y a certes chez von Mises un souci de l'empiricité, mais le rapport à l'expérience n'a pas chez lui ce caractère direct — si bien que le problème de l'application se pose en fait pleinement, chez lui comme d'ailleurs chez Waismann. Dans *Wahrscheinlichkeit, Statistik und Wahrheit* [11, p. 102], il affirme, ce qui a pu faire sursauter Waismann, que les probabilités sont une théorie de faits observables déterminés, d'événements répétés en masse comme les jeux de hasard, la statistique des populations, etc. Von Mises précise utiliser ici le terme de théorie au sens où on parle de « théorie des phénomènes caloriques » pour la thermodynamique, ou de « théorie des phénomènes spatiaux » pour la géométrie. Suite de comparaisons inacceptables pour un lecteur du *Tractatus* ! Chaque théorie, poursuit von Mises, commence par un certain nombre d'axiomes, dans la formulation desquels il est fait « usage de l'expérience générale », même s'il est clair qu'ils n'énoncent pas de « faits directement observables ». La différence entre un philosophe aprioriste comme Waismann et un mathématicien appliqué comme von Mises, pour qui l'axiomatique est soumise à la pression générale de l'expérience[4] (aussi paradoxal cela soit-il pour un logicien de stricte observance), tient donc à ce qu'on entend par cet *usage de l'expérience générale*. Waismann y reviendra en fin d'article en des termes pragmatistes. Notons que, contrairement à ce que suggère Waismann, toute conception fréquentiste n'est pas vouée à l'idée qu'on décrit le comportement *observable* de dés *réels* quand on définit la probabilité d'un 2 comme fréquence relative ; l'axiomatique fréquentiste de von Mises a d'ailleurs été rendue rigoureuse par Wald [13], puis Church [2].

2.2. **Le dé pipé.** L'article rappelle ensuite la critique classique du fréquentisme naïf : elle porte sur le dé pipé, qui *n'a pas* une chance sur six de produire un 2. La conception fréquentiste n'explique pas la différence de la série obtenue : se contentant d'enregistrer la modification, elle n'explique pas le droit que nous avons d'attendre un changement dans la fréquence.

Waismann poursuit ainsi la critique : selon la réponse insatisfaisante du « statisticien » (ici un interlocuteur fictif), les séries statistiques sont ses faits ultimes et cela n'a pas de sens de demander pour quelle raison telle série particulière existe ; explications physiques et statistiques seraient indépendantes. Division du travail que Waismann juge inacceptable, puisqu'on ne peut renoncer à l'idée qu'il y a bien une connexion entre le pipage du dé et la série nouvelle obtenue. Ce développement de Waismann, reposant sur l'exagération rappelée plus haut (le fréquentisme fondé sur l'observation), est peu convaincant : en quoi le fréquentiste serait-il condamné à nier une connexion, qu'il avoue simplement ne pas pouvoir établir (ce n'est pas son

[4]Indiquons ici la fort intéressante manière dont Kolmogorov présente son axiomatique dans les *Grundbegriffe der Wahrscheinlichkeitsrechnung* [7] : la section 1 énonce les axiomes ; la section 2 s'intitule « Relation aux données expérimentales », où est opérée une « déduction empirique des axiomes » ! Il déclare s'inspirer largement de von Mises dans cette section. Une comparaison épistémologique de leurs interprétations reste à faire.

objet) ? L'objet des statistiques est justement de tester des hypothèses en l'absence de toute autre considération.

Par contre, il est juste que le statisticien n'*explique* en rien le phénomène en le mettant en évidence. Il est tout aussi juste que le fréquentiste naïf doit dire de quel droit il attend un changement, *alors que* son point de départ est celui de l'application d'un principe d'équiprobabilité (le seul cas qu'il sait vraiment « expliquer », par symétrie), ce que Waismann oublie de rappeler. Le problème posé par le dé pipé n'est pas celui de la relation entre explications physiques et tests statistiques : c'est qu'il brise ce principe, sur lequel se fonde l'approche naïve des probabilités. Le fréquentiste se trouve dès lors sommé d'expliquer pourquoi il s'attend à quoi que ce soit, sachant que les séries empiriques effectivement observées peuvent être fort différentes de ce à quoi il « s'attend ». Cette attente, comme il ressortira de la conception de Waismann et comme tout un chacun l'admet maintenant, est déterminée dans un modèle *a priori*.

2.3. **Qu'est-ce qu'un *Kollectiv* ?** Suit alors l'objection essentielle de Waismann à von Mises, une objection selon son terme, purement logique, concernant la façon dont la fréquence relative est exprimée mathématiquement. Cette objection se déroule en deux temps et porte sur les difficultés qu'a eues von Mises pour construire les suites « aléatoires » au fondement de sa théorie, à savoir l'idée de *Kollectiv*. Selon l'article de 1919, un collectif est une suite infinie de nombres pris dans un ensemble E, caractérisée par deux axiomes :

(1) chaque partie de E admet une fréquence relative limite,
(2) extraire une sous-suite par une « sélection de place » admissible ne change pas cette limite.

On interprète E comme l'ensemble des résultats possibles (« le dé donne 1 »), ses parties[5] comme l'ensemble des événements possibles (« le dé donne un nombre pair »), et le collectif comme un ensemble infini de réalisations indépendantes d'une expérience idéale. Prenons l'exemple du jeu de pile ou face *équilibré*. Un collectif associé sera une certaine suite infinie de piles et de faces, respectivement codés par 0 et 1 par exemple, dans laquelle les résultats 0 et 1 ont pour fréquence relative limite $1/2$ (c'est, par définition, ce qu'on entend par un jeu équilibré). On peut choisir au sein de ce collectif différentes sous-suites en y sélectionnant des « places », par exemple les termes d'indices pairs, ou ceux suivant une série de 100 piles consécutifs. Von Mises explique l'axiome (2) en termes de casino : il n'y a pas de stratégie[6] permettant d'augmenter mes gains, pas de suite particulière d'instants qui augmenterait mes chances de tomber sur pile par exemple. Le point de cet axiome (2), dit

[5] La quantification « chaque partie de E » dans (1) n'est pas essentielle, et en fait nocive, *cf.* §3.1. Dans les versions ultérieures de sa théorie, von Mises se restreindra aux parties « opérationnellement descriptibles » si E est infini, celles qui apparaissent effectivement dans les applications.

[6] On suppose ici que je joue une infinité de fois, et que je m'autorise uniquement à choisir le moment auquel je joue (pas le montant des mises), en prenant au besoin en compte le passé au moment de miser.

d'irrégularité (*Regellosigkeit*), est que de telles sous-suites existent évidemment, mais qu'aucune stratégie *pré-établie* ne permet de les *construire* ; il s'agit bien d'un jeu « de hasard » et le propriétaire du casino peut dormir tranquille.

Von Mises donne quelques principes généraux et divers exemples pour préciser cet axiome d'irrégularité, mais sa définition reste incomplète : que sont exactement les sélections de place *admissibles* ? Quelle type de stratégie autorise-t-on ? Considérons de nouveau le collectif ci-dessus, modélisant le jeu de pile ou face équilibré. Sans autre spécification, il est clair qu'une sous-suite[7] $\{1, 1, 1, \text{ etc.}\}$ n'a pas pour fréquence relative limite $1/2$. Il reste donc à démontrer que le concept de collectif n'est pas contradictoire : s'il n'existe pas de collectif, la théorie est vide[8]. Mais Waismann adopte directement une posture critique : la formulation présente n'a pas de sens clair[9] et l'ensemble de l'idée fréquentiste s'effondre avec elle.

Waismann fait une seconde objection à von Mises (la première dans l'ordre du texte), qui explique, à ses yeux, ses difficultés définitionnelles. Elle porte sur l'idée même d'une suite mathématique irrégulière, que Waismann considère absurde. La suite mathématique, explique-t-il, dépend d'une loi. Au contraire, poursuit Waismann, en statistiques, on part d'une suite finie, que le fréquentiste poursuit à l'infini pour pouvoir lui appliquer le concept mathématique de suite (une suite infinie de jets de dés...). Mais on n'est autorisé, selon Waismann, à parler d'une suite infinie que si on dispose d'une loi capable de générer la suite. Le simple fait que la suite du fréquentiste est irrégulière prouve, selon lui, qu'elle n'est pas un concept mathématique. L'« accidentel », l'« imprévisible », selon ses termes, ne peuvent être accueillis dans les mathématiques sans détruire leur nécessité *a priori* caractéristique. Il conclut que la définition de von Mises est victime d'une confusion logique.

Mais, autant les problèmes définitionnels de von Mises sont réels, autant ce diagnostic de Waismann est-il hautement problématique, qui condamne par principe toute mathématisation du hasard, quand celle-ci est précisément l'enjeu des tentatives d'axiomatisation des probabilités des années 20 et 30. En premier lieu, remarquons qu'on démontre facilement sous les axiomes de Kolmogorov qu'une suite infinie de variables de Bernoulli indépendantes (le modèle d'une série de pile

[7]À proprement parler, von Mises ne permet que de sélectionner des *places*, c'est-à-dire des indices, et non pas les *valeurs* de ces places. Mais l'objection demande pourquoi une sélection de places admissible ne correspondrait pas à la sous-suite $\{1, 1, 1, \text{ etc.}\}$.

[8]Doob [3], puis Wald [13] montrent depuis la théorie de Kolmogorov qu'on peut préciser l'axiome d'irrégularité, sans altérer l'intuition de von Mises : le collectif n'est effectivement pas contradictoire. Church [2] a initié un autre angle d'attaque en formalisant la notion de sélection de place du point de vue de l'informatique, différents types de « sélections admissibles » conduisant à des théories différentes.

[9]Waismann ne cite pas le texte précis de von Mises auquel il se réfère (lequel restera ferme sur l'intuition sous-jacente mais variera dans sa formalisation). Il rejette l'axiome d'irrégularité en des termes sujets à caution : l'axiome d'irrégularité ne dit pas que toute partie « arbitrairement choisie » du collectif est admissible, ce que lui attribue pourtant Waismann. Cependant, à partir du moment où von Mises donne seulement des exemples, le lecteur peut choisir l'interprétation la moins charitable : il n'y a pas de charité en mathématiques...

ou face, certainement l'exemple paradigmatique de la théorie) vérifie les axiomes de von Mises (voir Feller [4, Chap. VIII]), légèrement modifiés. Les collectifs sont donc consistants dans le cadre standard.

Par ailleurs l'irrégularité s'est vue donner un sens formel dans les années 1960, y compris pour les suites finies, par Kolmogorov, Solomonoff, puis Chaitin[10] (d'une tout autre manière que von Mises). L'idée est la suivante : est aléatoire une suite qu'on ne peut essentiellement pas décrire de façon plus économique qu'en la transcrivant littéralement. L'intuition est claire : qui obéit à une loi simple se décrit facilement. Une suite répétant 1000 fois le même motif se décrit économiquement en recopiant le motif et en disant qu'il se répète 1000 fois ! Le seul vrai problème pour formaliser ceci est le concept de « description », qui n'est pas sans rappeler celui des stratégies ou sélection de places de von Mises, ce qui nous met d'ailleurs sur la voie : décrire, c'est donner un *algorithme*. L'idée de l'axiome d'irrégularité est en effet qu'un joueur n'augmente pas ses chances en suivant une stratégie explicite, l'existence non constructive d'une stratégie gagnante ne lui étant d'aucun secours. Pour quantifier l'irrégularité d'une suite binaire finie x, on définit ainsi rigoureusement sa *complexité* $K(x)$, en utilisant le formalisme de l'informatique théorique[11]. C'est un bien intéressant objet, qui fournit par exemple une preuve immédiate et « constructive » du théorème d'incomplétude de Gödel[12] : pour tout x et tout entier m assez grand, la proposition « $K(x) > m$ » n'est pas démontrable. L'entier m étant fixé, elle est par ailleurs vraie pour tout x à un nombre fini d'exceptions près.

[10]*Cf. L'héritage de Kolmogorov en mathématiques* [1, Chap. XIII].

[11]On choisit un modèle de calcul, par exemple celui de la machine de Turing, qui explique ce qu'est un « algorithme », comment et sur quels objets il opère, etc. Comme toute information finie, données et résultats se codent par des mots (finis) composés de 0 et de 1, dont la *longueur* est le nombre de chiffre. Par définition, une fonction f est calculable dans le modèle si un algorithme calcule le mot $x = f(t)$ pour tout mot donné t ; la « thèse de Church » (de Church-Turing-Post, en fait) énonce en gros que les fonctions calculables sont les mêmes dans tous les modèles de calcul raisonnables : c'est un théorème pour les modèles proposés jusqu'ici. Si f est une fonction calculable, la *complexité* (de Kolmogorov) $K(x, f)$ du mot x est la *longueur* du plus petit mot t tel que $f(t) = x$; et une valeur infiniment grande si t n'existe pas.

Intuitivement, on va choisir un f très expressif, en fait le voir comme une machine universelle, qui exécute un programme codé par t ; $K(x, f)$ est la longueur de la description la plus efficace de x, si on choisit d'utiliser f. De plus, il existe une fonction calculable optimale F (pas unique) telle que $K(x,F)$ soit inférieur à $K(x, f)+C(f)$ pour tout mot x et toute fonction calculable f, où $C(f)$ est une constante ne dépendant que de f. La *complexité de Kolmogorov* $K(x)$ de x est la valeur de $K(x,F)$ pour une fonction optimale F. Ce qui précède dit en particulier que, si G est une autre fonction optimale, alors la différence entre $K(x,F)$ et $K(x,G)$ est bornée par une constante ne dépendant que de F et de G, mais pas de x !

Toujours intuitivement, $K(x)$ est la longueur du plus petit programme calculant x, indépendamment de la machine universelle (de la fonction optimale F) choisie. En ce sens, K ne dépend pas de notre concept d'« ordinateur ». C'est une fonction bien définie (à une constante près liée au choix de la fonction optimale F) mais non calculable !

[12]La démonstration généralise le paradoxe classique du plus petit entier dont la description nécessite plus de cent mots. Le lecteur est invité à réfléchir au sens du mot *description* dans cette « définition ».

Elle fournit aussi, et c'est dans ce but que nous l'introduisons, une définition rigoureuse et intuitive de ce qu'est le hasard. Par définition, un mot est aléatoire s'il est incompressible : sa complexité de Kolmogorov est de l'ordre de sa longueur. Pour tout entier n, il existe un mot incompressible de longueur n, puisque les 2^n mots de longueur n ne peuvent être décrits par les $1+2^2+\ldots+2^{n-1}=2^n-1$ descriptions potentielles de longueur strictement inférieure. La notion s'adapte aux suites infinies, avec quelques difficultés et des idées nouvelles (Martin-Löf, Schnorr).

Si Waismann a raison de dire que la conception de l'irrégularité défendue par von Mises est incertaine, son rejet de la description de l'irrégularité du champ des mathématiques est ainsi une pétition de principe, et ne suit en rien de leur « nécessité *a priori* ». Notons tout de même que, dans un cadre intuitionniste fort, l'affirmation excessive de Waismann « opérer sur une suite infinie veut dire opérer sur les lois qui engendrent la suite », devient trivialement vraie, et exclut effectivement tous les développements ci-dessus.

2.4. **Critique de l'idéalisation.** En sus de l'affirmation de cette impossibilité intrinsèque pour le formel de penser le hasard, Waismann propose une interprétation philosophique du problème que rencontre von Mises, issu, soutient-il, d'une confusion entre *a priori* et *a posteriori*, d'une notion vague de l'idéalisation de l'expérience aux mathématiques. Nous avons vu plus haut comment l'idée de von Mises d'une axiomatique faisant « usage de l'expérience générale » signalait pour Waismann qu'une telle confusion était à l'œuvre. Von Mises dit que les concepts des probabilités seraient des abstractions ou des idéalisations de l'expérience, sans s'attarder sur les dimensions philosophiques d'une telle question. Waismann s'oppose à l'idée d'un tel processus, et voit (à tort) la source des difficultés (réelles) pour définir le collectif dans ce qui est à ses yeux un empirisme intempestif.

Ainsi se défendrait donc, selon Waismann, un fréquentiste : en géométrie quand nous raisonnons sur des formes exactes, nous idéalisons les formes sensibles de la perception ; pareillement dans le calcul des probabilités nous idéalisons les fréquences établies empiriquement, en définissant un concept de fréquence limite. Waismann nie dans les deux cas la pertinence du concept usuel d'idéalisation ; il illustre l'erreur commise en prenant l'exemple de la *définition* de π comme idéalisation de mesures de plus en plus fines sur un cercle. Tout au contraire, π est l'étalon grâce auquel nous jugeons de la précision des mesures (qu'entend-on par « de plus en plus fines » ?). Pour Waismann, les propositions de la géométrie sont un système de règles, qui forment la syntaxe des concepts employés pour décrire les relations spatiales. Et on ne peut en tout cas (cela n'a pas de sens) approcher des règles de syntaxe en raffinant des observations. Selon les termes de Waismann, on ne parvient pas à l'idéal, on en part.

Von Mises serait d'accord avec l'idée que la donation d'un ensemble d'axiomes n'est pas une *telle* idéalisation, par « passage à la limite », mais pose en réalité la question du rapport des axiomes aux pratiques empiriques, pour lesquelles on recherche précisément des fondations. Il insistera toujours[13] sur l'idée qu'une application rationnelle des probabilités au monde réel *doit* être fondée sur une théorie fréquentiste. Plus généralement, les pratiques précèdent, en un sens qui reste à élucider, l'axiomatique. La conception de Waismann (et du premier Wittgenstein) ne permet pas d'éclairer philosophiquement cette relation. *De la certitude* [15], par contre, soutiendra l'idée d'une perméabilité des frontières entre grammaire et expérience.

Donc Waismann oppose une objection épistémologique à tout fréquentisme : l'idéalisation ne peut transformer des fréquences en une limite. (Mais von Mises ne commet pas cette erreur.) Les deux points de Waismann sont justes, mais indépendants : le *collectif* de von Mises est mal défini, et l'idéalisation naïve se trompe sur la nature de l'idéal. Nous ajoutons : il faut un concept pré-défini de la probabilité pour éventuellement l'interpréter ensuite comme limite.

Puisqu'on ne parvient pas à idéaliser le concept de série statistique, dit Waismann, excluons-le de nos fondations. Il passe alors à l'exposé de sa propre conception.

3. La conception logique de Waismann

3.1. **Préambule.** Le probabiliste moderne, après Kolmogorov [7], commence par définir un espace probabilisé, composé d'un *univers*, d'une *tribu*, et d'une *mesure*, puis étudie des *variables aléatoires*. L'univers est l'ensemble des résultats possibles d'une expérience, ou *atomes* : en lançant ce dé, j'obtiendrai 1, 2, ..., ou 6 ; dans cet exemple, l'univers a six éléments. Une tribu est un ensemble de parties distinguées de l'univers, qui sont les *événements* qui nous intéressent (« le résultat du dé est pair » est réunion de trois atomes) ; chaque événement est donc composé d'atomes[14]. Une mesure associe à chaque élément de la tribu un réel positif ou nul. On peut décider que toutes les parties sont dans la tribu, mais ce n'est en rien nécessaire, et souvent contre-indiqué[15]. Une variable aléatoire est une fonction qui

[13]face au scepticisme général (voir en particulier les recensions des éditions successives de [11] dans les *Mathematical Reviews*) : après bien des efforts, son formalisme est, pour les applications, essentiellement le même que celui dérivé suivant Kolmogorov (il lui apporte quelques restrictions ennuyeuses). Par contre, bien des auteurs, Kolmogorov le premier, citent avec respect son point de vue et, si son axiomatique est aujourd'hui oubliée par les probabilistes, elle a motivé des travaux importants en informatique théorique.

[14]lesquels sont d'habitude appelés événements élémentaires pour cette raison, mais la terminologie est maladroite : un événement élémentaire n'est pas un événement, le singleton le contenant en est un.

[15]Sous l'axiome du choix, la mesure naturelle sur l'intervalle $[0,1]$ (pour laquelle la mesure d'un intervalle est sa longueur) ne *peut pas* être définie sur l'ensemble de *toutes* ses parties. Plus fondamentalement, une tribu décrit une information ; restreindre la tribu diminue les possibilités d'aléa, donc augmente l'information supposée connue. Le concept moderne de probabilité conditionnelle repose sur ce principe : le remplacement d'une tribu par une sous-tribu, et le remplacement d'une variable aléatoire par sa projection orthogonale sur l'espace des fonctions ayant accès à cette nouvelle information.

associe à chaque atome un nombre réel. Ainsi, le lancer d'un dé à 6 faces se modélise par la variable aléatoire prenant la valeur 1 quand l'événement « le résultat est 1 » se produit, etc.

3.2. **Le propos de l'axiomatique de Waismann.** Waismann estime que les probabilités sont une branche de la logique ; en particulier, il se restreint dès l'abord aux problèmes de décision[16]. Dans le vocabulaire moderne rappelé plus haut, il laisse l'univers et la tribu dans l'ombre, se concentre sur la notion de mesure, et considère des variables aléatoires ne prenant que deux valeurs, 0 et 1 (vrai et faux si l'on veut).

Il désire donner un sens utile à un énoncé du type « la probabilité que p soit vraie est de 1/6 », pas estimer le résultat numérique probable de mesures expérimentales[17], et encore moins déduire la loi physique la plus probable expliquant un ensemble de mesures ; toutes choses dans lesquelles Poincaré voit l'apport essentiel des probabilités à l'activité scientifique. Waismann estime que le rôle des énoncés de probabilité, comme de tout énoncé logique, n'est pas de prédire le résultat de l'expérience (ils ne sont ni vérifiables ni falsifiables empiriquement), mais de dérouler les conséquences de nos assomptions initiales. Elles sont un langage pour décrire toute situation dont nous n'avons pas une connaissance parfaite (ni une connaissance nulle, d'ailleurs). Si on leur attribue souvent une valeur prédictive, Waismann entend ainsi que son axiomatique évite les problèmes philosophiques associés. Il insiste donc sur le fait que le contenu empirique des probabilités (et en général des mathématiques ou de la logique) est nul : on n'obtient rien qui ne soit déjà contenu dans les données explicites du problème. Là où la mécanique attend le verdict de l'expérience, le calcul des probabilités nous dit quelque chose que nous savons déjà : il porte sur les relations entre des ensembles de propositions. Dans cette mesure, il est logique.

3.3. **Spielraum.** On suppose donc un univers, c'est-à-dire l'ensemble des faits et événements élémentaires possibles, et on étudie les propositions qui en disent quelque chose. Waismann ne précise ni l'univers considéré, ni les propositions étudiées, mais admet la propriété suivante : une proposition donnée est vraie si et seulement si un ensemble bien déterminé de faits élémentaires se produisent ; et fausse sinon. Il appelle *champ* (*Spielraum*) de la proposition cet ensemble de faits[18]. Les relations logiques (négation, implication, équivalence, conjonction, disjonction) entre propositions se traduisent immédiatement en opérations ensemblistes sur leurs champs (complémentaire, inclusion, égalité, intersection, réunion).

[16]Outre les inévitables jeux de hasard, la problématique du jugement éclairé est fréquente chez les probabilistes classiques. Ainsi Poisson, qui inventa le terme de "loi des grands nombres", intitule son livre *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités* (1837).

[17]On peut toutefois estimer le résultat d'une mesure par une décision, en considérant une proposition p du type « le résultat est dans l'intervalle $[a,b]$ ».

[18]Dans la terminologie du *Tractatus*, 5.101, le *Spielraum* est l'ensemble des *Warheitsgründe*.

Waismann fait mine de considérer que le monde ordinaire est un univers valide et qu'un énoncé ordinaire « Mon ami est à Paris » est justiciable de ce type d'analyse, mais un langage et un monde largement formalisés sont bien entendu nécessaires. Nous n'en saurons pas plus, mais la formalisation précise choisie est sans incidence sur la suite.

3.4. **Mesure.** Ayant défini un champ, et des relations ensemblistes qu'il qualifie de topologiques, Waismann franchit un pas supplémentaire et fixe une *mesure* de certains énoncés :

(1) La mesure d'un énoncé est un réel positif ou nul.
(2) Un énoncé vide (une contradiction) est de mesure nulle.
(3) La mesure de la conjonction de deux propositions *incompatibles* est la somme de leurs mesures.

Waismann laisse explicitement ouverte la possibilité d'énoncés non mesurables, sans s'attarder sur ce fait ou sa signification. L'idée intuitive est que l'on mesure le champ des propositions qui nous intéressent ; à condition de respecter les axiomes, on peut affecter aux autres propositions une mesure nulle. Dans le cas d'un univers fini, on peut *choisir* la mesure de comptage, convention correspondant à décréter les événements équiprobables. Waismann insiste sur l'arbitraire du choix d'une mesure, et cite le cas du paradoxe de Bertrand, où il s'agit de choisir des cordes sur un cercle (de façon que leur longueur soit supérieure au rayon). Le « paradoxe » de la multiplicité des probabilités envisagées comme réponse possible s'explique par plusieurs choix, naturels mais incompatibles, pour la mesure définie sur des ensembles de cordes. Ce dernier exemple est non-trivial ; en particulier, il ne ressort pas de la simple combinatoire comme les traditionnelles urnes ou jeux de pile ou face.

3.5. **Probabilité relative, ou conditionnelle.** Pour a et b deux propositions mesurables, la « probabilité de a sachant b » est le quotient de la mesure de « a et b » par la mesure de b (qu'il serait utile ici de supposer non nulle). Nous le noterons $p(a,b)$. Waismann interprète ce quotient, qui est compris entre 0 et 1, comme une évaluation de la connexion déductive[19] entre a et b, du degré de plausibilité que la connaissance de b imprime à a : il est nul si les énoncés sont incompatibles, égal à 1 si b entraîne a. Remarquons que de $p(a,b) = 1$, on ne peut déduire réciproquement que b implique a, mais uniquement que la partie de l'univers où l'implication est falsifiée est de mesure nulle[20].

Dans cette définition, il n'est nul besoin de considérer l'univers tout entier, ni d'y définir une mesure : seul le champ de b importe et une mesure qui s'y restreint. Si b est de champ fini, on peut choisir la mesure de comptage, qui associe à un

[19] L'image de proximité logique qu'il utilise aussi est plus trompeuse : une distance est symétrique, alors que $p(a,b) = p(b,a)$ si et seulement si a et b ont même mesure.

[20] En supposant que l'ensemble des propositions mesurables est stable par les opérations logiques. Waismann ne le fait pas, mais c'est probablement un oubli.

ensemble le nombre de ses éléments. Dans ce cas on retrouve la formulation classique[21] du nombre de cas favorables divisé par le nombre de cas possibles, *supposés équiprobables*. Avec, bien sûr, l'avantage qu'on n'a plus à définir la problématique dernière condition.

Il serait naturel de définir ensuite la probabilité absolue de a comme $p(a,b)$, où b est une tautologie quelconque, par exemple « a ou non-a » dans un cadre non-intuitionniste : le champ d'une tautologie est l'univers, et la supposer ne nous apprend rien sur a. Le procédé de Waismann est différent, et difficilement formalisable : il choisit b comme la conjonction de toutes les phrases que nous savons vraies, en y incluant les lois de la nature (qui sont seulement supposées). La raison épistémologique est claire : il obtiendrait ainsi le degré de plausibilité que l'ensemble de nos connaissances empiriques actuelles donne à a (vu tout de même à travers le filtre d'une mesure qui reste arbitraire). Le point philosophique est que la probabilité n'est pas une propriété de a, mais de notre connaissance (incomplète) de a, laquelle doit donc intervenir dans la définition. En contrepartie, le concept devient vide, puisqu'introduire le monde réel dans notre axiomatique en signe la ruine complète : la construction est foncièrement atomique (l'univers est un ensemble d'éléments), et on ne voit pas très bien ce que seront les événements élémentaires ; ni quelles seront les conditions de vérité des propositions du langage ordinaire. On retrouve par un autre biais la critique classique de l'atomisme logique du *Tractatus*.

3.6. **Rapport à l'axiomatique de Kolmogorov.** Waismann ne restreint pas le cadre en raisonnant ainsi sur des propositions, plutôt que sur des parties arbitraires de l'univers. Les deux langages sont équivalents : tout événement peut être considéré comme le champ d'une proposition, précisément celle qui dit que l'événement se produit. On pourrait aussi bien s'exprimer ainsi :

(1) La mesure d'une partie est un réel positif ou nul.
(2) L'ensemble vide est de mesure nulle.
(3) La mesure de la réunion de deux parties *disjointes* est la somme de leurs mesures.

Il serait préférable de rajouter explicitement aux axiomes les propriétés de l'ensemble des parties mesurables. Implicitement, (2) suppose que l'ensemble vide est mesurable, et (3) qu'une réunion de deux mesurables est mesurable. Waismann ne donne pas d'application concrète détaillée de son axiomatique, renvoyant à un travail ultérieur jamais publié la preuve qu'elle entraîne « toute la structure formelle du calcul des probabilités ». Ce faisant, il aurait probablement rajouté les axiomes suivants, nécessaires pour retrouver les théorèmes habituels des probabilités élémentaires et non déductibles des siens :

(4) L'ensemble des parties mesurables est stable par passage au complémentaire et par réunion finie. Il contient l'ensemble vide.

[21] C'est en particulier celle du *Tractatus*, 5.15, qui par ailleurs n'introduit pas l'idée d'une mesure *arbitraire*.

(5) L'univers n'est pas de mesure nulle.

(4) assure que l'ensemble des propositions mesurables est stable par les opérations logiques élémentaires. Et (5) que les quotients définissant $p(a,b)$ sont légitimes : dans le cas trivial de la mesure uniformément nulle, le procédé de Waismann ne définit rien.

On obtient alors essentiellement les axiomes de Kolmogorov pour les probabilités sur des univers *finis*[22]. Kolmogorov impose que l'univers soit de mesure 1 ; de nos axiomes complétés, on déduit seulement que l'univers est mesurable (comme complémentaire de l'ensemble vide, qui l'est), de mesure M fixée. C'est une légère complication technique : on peut normaliser la mesure, en la divisant par M, sans rien changer aux quotients définissant les probabilités conditionnelles. On peut aussi remplacer (5) et imposer que l'univers soit bien de mesure 1. Dans ce cas, la mesure donne directement la probabilité absolue, qui devient donc un concept primitif, contrairement à l'intention de Waismann qui tient à la priorité de la probabilité conditionnelle. Pour lui, et plus généralement pour les partisans de la probabilité logique, la probabilité doit d'abord exprimer le rapport logique entre deux propositions.

À condition de remplacer la définition exotique de la probabilité absolue de a par la définition naturelle (conditionnellement à une tautologie), on retrouve exactement l'axiomatique de Kolmogorov, donc la théorie usuelle des probabilités. D'un point de vue moderne, la grande intuition de Waismann et de faire porter l'accent sur une théorie de la mesure (certes incomplètement formalisée). Mesure librement choisie, mais pas frivole : changer de convention à chaque application ôterait toute valeur au calcul des probabilités ; une fois la mesure choisie pour un champ de phénomènes, elle le gouverne désormais.

4. UNE DÉFENSE PROGRAMMATIQUE

4.1. La probabilité logique face aux applications. En dernière partie, Waismann revient sur le point délicat des relations entre sa théorie et les usages courants des probabilités, en particulier du concept de fréquence relative. Il distingue entre les statistiques strictement empiriques (détermination du pourcentage de naissance de garçons par rapport aux filles sur une année donnée), et les probabilités *a priori*. Il reconnaît que dans le premier cas la probabilité se réduit à la fréquence relative toutefois pensée sans limite à l'infini. Dans le cas *a priori* cependant, cette équation est refusée : le calcul des probabilités n'est pas une théorie empirique. L'application est seconde par rapport à l'axiomatique, et ce n'est qu'alors qu'une épistémologie de l'expérience peut débuter.

[22] Pour des univers infinis, Kolmogorov rajoute la stabilité de la tribu par réunion *dénombrable* et un axiome d'additivité pour les unions dénombrables d'ensembles disjoints. C'est une complication technique, nécessaire à l'application de techniques d'analyse pour étudier rigoureusement les problèmes de convergence.

Que les axiomes soient défaits de tout sens empirique implique, selon Waismann, le caractère immuable du calcul des probabilités. Seule changera la mesure adoptée, qui déterminera les valeurs numériques de telle ou telle application des probabilités à l'expérience. Cette détermination permet précisément le genre d'anticipations de l'expérience statistique qui induit en erreur, selon lui, les partisans d'une théorie empirique de la probabilité. Ainsi, pour appliquer le calcul à la roulette, on adoptera une mesure où les différents lieux de taille identique sont équiprobables (la mesure uniforme), en se soumettant aux faits d'expérience, qui ne nous suggèrent pas d'intervention d'autres facteurs (couleur, etc.) que les relations spatiales.

À quoi obéit plus généralement le choix de la mesure ? Waismann est programmatique et avance plusieurs possibilités. Premièrement la stipulation doit être telle que nos expériences statistiques s'y accordent. Deuxièmement, les déviations par rapport à la norme doivent être le moins nombreuses possibles, requête qu'il propose en général de ranger au sein de considérations pragmatiques d'utilité et de simplicité. Avec ces remarques, Waismann espère pouvoir retrouver en particulier l'application des probabilités à la physique, où, dit-il, il est connu qu'elles sont indispensables.

4.2. **La loi des grands nombres.** En général, si, un peu nettoyées, ses fondations des probabilités sur les ensembles finis sont remarquablement lucides, Waismann passe à côté des théorèmes limites qui, pour avoir une longue histoire remontant au moins à Bernoulli et Gauss, n'en sont pas moins centraux dans la théorie moderne : ils *sont* le véritable sujet mathématique des probabilités, et conditionnent leurs applications réelles, aux statistiques en particulier.

Considérons en effet la fin de l'article : un long passage sur la loi des grands nombres, dont malheureusement Waismann ne cite pas de forme précise (il en existe de nombreuses variantes, y compris à l'intérieur des mathématiques), mais à laquelle il assigne un statut épistémologique. Le lecteur se pose dès lors la question légitime des formes et du statut mathématique de cette loi[23] dans l'article. Deux interprétations sont possibles de ce final : soit Waismann sait qu'au moins une forme faible de la loi est considérée comme théorème mathématique, auquel cas il est étrange qu'il ne dise pas comment il peut l'obtenir à partir de ses propres axiomes, ni même qu'il se proposerait de la dériver[24]. Soit Waismann ne sait pas que la loi

[23]Qui énonce qu'avec probabilité 1, la moyenne des résultats successifs obtenus dans une suite indépendante d'expériences analogues tend vers une constante fixée, sous des hypothèses techniques précisant « indépendante », « analogue », et la constante. Un énoncé plus faible (Chebyshev, 1859 ; des cas particuliers remontent au moins à Bernoulli) dit que la probabilité d'un grand écart entre la moyenne des n premiers résultats et la constante tend vers 0. Cette version faible dit que la fréquence des grands écarts est petite ; la version forte dit qu'avec probabilité 1, les écarts restent petits. La forme rigoureuse de la loi forte des grands nombres est due à Kolmogorov (1927), précisée l'année suivante (avec Khintchin) par la loi du logarithme itéré, qui estime la taille maximale des écarts. De l'un quelconque de ces résultats, on déduit par exemple que les fréquences relatives des résultats d'un grand nombre de lancers de dés équilibrés tendent à s'égaliser.

[24]La forme faible de la loi, due à Chebyshev, est démontrable à partir des axiomes de Waismann, amendés fidèlement à son rejet du concept de limite de la définition des probabilités. Elle énonce

des grands nombres est un théorème, et, dans ce cas, c'est à une véritable critique de cette loi même qu'il procéderait : il n'y a pas de telle loi[25], écrit Waismann, ce qu'on serait en droit d'interpréter radicalement. Mais un problème s'ouvre aussitôt : s'il ne s'agit pas pour lui d'une loi, alors *à quoi* exactement attribue-t-il un statut épistémologique ? Le texte ne le dit pas. Si, en rebours, on adopte une autre interprétation de l'affirmation d'inexistence, selon laquelle cette loi n'est pas une loi *de la nature* (sous-entendu), alors on attend à nouveau une présentation mathématisée (ou son annonce) qui précéderait l'interprétation épistémologique, étant donné l'importance de cette loi dans l'application des probabilités.

Considérons cependant, malgré ces difficultés du texte de Waismann, l'interprétation épistémologique qu'il donne de la « loi des grands nombres », quoi qu'il entende par là. Elle signifie, selon Waismann, la pure et simple limite que l'on met à la recherche de l'explication. Posant une forme d'égalisation par le hasard, elle ne dit en réalité rien d'autre que : là commence le hasard, et là les lois de la nature. Si on considère les lois de la nature comme le fait Waismann, comme des structures formelles, la loi des grands nombres aurait un rôle méta-structurel indiquant où leur usage est légitime, et où illégitime. Ainsi, la loi n'est que la conséquence, selon le terme de Waismann, de l'existence des lois de la nature. Mais, autant l'idée de rendre épistémique la distinction entre hasard et loi de la nature est intéressante, autant cet usage du terme de conséquence est étrange dans un texte de portée axiomatique où le lecteur est légitimé à prendre le terme de conséquence en son sens logique. Cette idée que l'existence de lois de la nature a pour conséquence la loi des grands nombres milite en faveur non pas de l'interprétation suivant laquelle la loi est pour Waismann un théorème, mais plutôt pour la lecture selon laquelle elle exprime simplement (comme il le dit) que nous ne recherchons pas d'autre loi : une limite à l'enquête empirique. Waismann, selon cette lecture, reste fidèle à l'idée qu'il n'y a pas de loi, là où il y a irrégularité. Mais un des enjeux principaux des fondations des probabilités est bien dans la recherche d'un discours rigoureux sur le hasard.

Tout à son combat contre le fréquentisme, Waismann relègue ainsi la loi des grands nombres à une forme d'expression : quand on constate empiriquement qu'un grand nombre de résultats s'égalisent, on *parle* justement de loi des grands nombres. On aurait tort d'invoquer une prétendue loi de la nature qui aurait *causé* cette égalisation. Le lancer du dé, son rebond sur la table, ses mouvements de rotation sont sujets aux lois de la nature ; ayant appliqué ces lois, il ne reste plus rien à expliquer. Il n'y a pas de raison de s'alarmer de ce que les six résultats possibles apparaissent avec des fréquences proches : vu les symétries du problème, on ne pouvait s'attendre à autre chose. Waismann conclut que la loi des grands nombres n'est qu'une

---

simplement qu'une certaine fonction tend vers 0, ce qui s'exprime avec les outils standards de l'analyse ; le fait que les valeurs de cette fonction soient des probabilités ne joue aucun rôle.

[25]Notons que la loi des grands nombres a été victime du succès même de l'expression : Waismann rappelle à juste titre qu'elle a été l'objet d'usages fantaisistes depuis Bernoulli. Elle a exercé, selon ses termes, l'imagination des philosophes… Un exemple en est l'usage peu rigoureux qu'en fit Poisson. *Cf.* Hacking, *The Taming of Chance* [5, Chap. 12].

expression pour indiquer que nous avons atteint le point où nous ne cherchons plus de nouvelle loi de la nature, que nous sommes satisfaits du niveau d'explication déjà atteint. Il n'envisage pas que l'on puisse démontrer mathématiquement un théorème (la loi faible des grands nombres) à partir de ses propres axiomes. Certes, il reste la question épineuse du rapport d'un théorème sur la convergence d'une suite de variables aléatoires avec une prédiction empirique. Mais le problème se pose pour tout modèle mathématique, et n'a rien de spécifique aux probabilités.

5. Conclusion

L'interprétation la plus plausible de l'article est bien que Waismann veut exclure tout concept de limite de la définition des probabilités. Dans ce cas, le concept logique de la probabilité défendu par Waismann est tout compte fait fort différent de ce qu'on appelle probabilité tant chez Poincaré en amont, qui l'utilise avec le soutien de l'intuition physique, que chez Kolmogorov en aval, dont le propos est de retrouver, mieux que Von Mises mais en sympathie avec lui, un bon usage du concept de limite dans les probabilités, et par là la loi des grands nombres, dans sa version la plus forte.

Ajoutons un dernier point sur la question cruciale dans l'économie de l'article des liens entre expérience et axiomatique. Si pour Waismann l'arbitraire ne s'étend qu'au choix de la mesure (le calcul des probabilités étant ce qu'il est de manière immuable), la stratégie globale de von Mises laisse ouverte la possibilité de nouvelles manières, sous la pression de l'expérience, d'axiomatiser les probabilités. La persuasion waismanienne qu'il n'existe qu'une arithmétique, qu'une logique et qu'un calcul des probabilités appartient bel et bien à une conception anhistorique de l'*a priori*, dont le second Wittgenstein montrera les difficultés. La paradoxale position de von Mises d'une axiomatique sous pression empirique (quand bien même il n'a pas eu le succès de Kolmogorov) est plus proche de l'épistémologie de *De la certitude* que la position philosophique inflexible de Waismann sur l'existence d'un unique calcul des probabilités.

Références

[1] É. Charpentier, A. Lesne & N. Nikolski (eds.), *L'héritage de Kolmogorov en mathématiques*, Belin, Paris, 2004.
[2] A. Church, « On the Concept of a Random Sequence », *Bull. Amer. Math. Soc.* **46** (1940), pp. 130–135.
[3] J. L. Doob, « Note on Probability », *Ann. of Math. (2)* **37** (1936), pp. 363–367.
[4] W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley, New York, 1968, 3^{e} édition.
[5] I. Hacking, *The Taming of Chance*, Cambridge University Press, Cambridge, 1990.
[6] M. Heidelberger, « Origins of the Logical Theory of Probability : von Kries, Wittgenstein, Waismann », *Int. Stud. Philos. Sci.* **15** (2001), no. 2, pp. 177–188.
[7] A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer-Verlag, Berlin, 1933. Tr. ang. par N. Morrison, Chelsea Publishing Company, New York, 1950.
[8] H. Poincaré, *La science et l'hypothèse*, Champs Flammarion, Paris, 1968, 1$^{\text{ère}}$ édition 1902.

[9] M. VAN LAMBALGEN, « Von Mises' Definition of Random Sequences Reconsidered », *Journal of Symbolic Logic* **52/3** (1987), pp. 725–755.
[10] R. VON MISES, « Grundlagen der Wahrscheinlichkeitsrechnung », *Mathematische Zeitschrift* **5** (1919), pp. 52–99.
[11] R. VON MISES, *Wahrscheinlichkeit, Statistik und Wahrheit*, Springer-Verlag, Vienne, 1928.
[12] F. WAISMANN, « Logische Analyse des Wahrscheinlichkeitsbegriffs », *Erkenntnis* **1** (1930), pp. 228–248. Tr. ang. « A Logical Analysis of the Concept of Probability » in *Friedrich Waismann, Philosophical Papers* (pp. 4–21), B. F. McGuinness (ed.), Reidel, Dordrecht, 1977.
[13] A. WALD, « Die Widerspruchsfreiheit des Kollectivbegriffes der Wahrsheinlichkeitsrechnung », in *Les fondements du calcul des probabilités* (P. Cantelli & al., eds.), Actualités scient. et industr., vol. 735, Hermann, Paris, 1938, pp. 79–99.
[14] L. WITTGENSTEIN, *Tractatus logico-philosophicus*, Routledge & Kegan Paul, Londres, 1922. Édition bilingue, tr. ang. par C. K. Ogden, et F. P. Ramsey. Seconde traduction par D. F. Pears et B. F. McGuinness, Routledge & Kegan Paul, Londres, 1961. Tr. fr. par G.-G. Granger, Gallimard, Paris, 1993. Le texte est paru pour la première fois sous le titre « Logisch-Philosophische Abhandlung » dans *Annalen zur Naturphilosophie* **14** (1921), pp. 185–262.
[15] L. WITTGENSTEIN, *Über Gewißheit / On Certainty*, G. E. M. Anscombe & G. H. von Wright (eds.), Blackwell, Oxford, 1969, tr. ang. par D. Paul et G. E. M. Anscombe. Tr. fr. *De la certitude* par J. Fauve, Gallimard, Paris, 1976.