# Andrei Nikolaievitch KOLMOGOROV (1903-1987)
## *Un aperçu de l'homme et de l'œuvre probabiliste*

Laurent MAZLIAK
Laboratoire de Probabilités
Université Paris VI
75252 Paris Cedex 05 France
mazliak@ccr.jussieu.fr

February 7, 2003

**Abstract**

Andrei Nikolaievitch Kolmogorov fut un des plus grands mathématiciens soviétiques du vingtième siècle. Son nom reste spécialement attaché à la théorie des probabilités à laquelle il apporta des contributions majeures. Cet article tente de donner une idée des principales, et de replacer l'homme dans le contexte très particulier de la science soviétique.

# 1 Introduction

Il faut naturellement commencer le présent exposé par dire combien l'étendue du sujet ne va permettre d'en aborder que quelques grandes lignes. Andrei Nikolaievitch Kolmogorov, autant le poser d'emblée ainsi, a été l'un des plus grands scientifiques du XXème siècle, et sa créativité mathématique va très au delà de sa seule œuvre probabiliste. Il fut curieux de tout et il y eut peu de domaines des mathématiques mais aussi de la physique dans lesquels il ne s'aventura pas, ne serait-ce que pour des essais de vulgarisation à l'usage des étudiants, voire du grand public. De ce fait, la mise à part d'un domaine spécialisé dans le travail si polymorphe d'un tel homme risque de faire manquer un certain nombre de concordances intéressantes qui lui permirent, d'une façon inattendue pour ses contemporains, de mettre en relation des parties *a priori* sans lien des mathématiques. "J'ai toujours voulu comprendre comment Andrei Nikolaievitch passait d'un thème à l'autre: les travaux se succédaient chez lui sur des sujets divers d'une manière visiblement non programmée." écrira V.I.Arnol'd ([Ширяев, 1993]). Mais le talent incontestable de ce savant complet n'explique pas tout le personnage en regard des conditions assez spécifiques dans lequelles sa personnalité put s'épanouir. Et ce point concerne tout spécialement le domaine des probabilités qui lui a donné ses titres de gloire les plus reconnus.

Pour tenter d'approcher le phénomène, un problème se pose immédiatement comme c'est souvent le cas pour les personnalités qui furent en vue dans le monde soviétique.

La plupart des documents de témoignage ou de reconstitution historique que nous avons s'apparentent plus à des hagiographies qu'à un véritable travail de mise en perspective (cela concerne en particulier les différents mémoires de Shiryaev sur ce thème comme [Shiryaev,1989],[Ширяев, 1993]). Pour des raisons que nous allons voir, Kolmogorov fit partie, bon gré mal gré (je reviendrai sur ce point), de ces savants que le régime a "canonisés" pour en faire des emblèmes de la réussite de la société socialiste. On peut voir comme une indication en ce sens les articles commémoratifs de la revue *Успехи Математических наук (Succès des sciences mathématiques)* pour le 50 ème (VIII,3,1953,176-200) et le 60 ème (XVIII,5) anniversaires de Kolmogorov. Il fut à lui seul un exploit de la science soviétique. En conséquence, et en raison de l'énorme poids hiérarchique qui avait cours en URSS dans les milieux académiques, ceux qui l'ont fréquenté ne savent plus ou moins en parler qu'avec une déférence dévote. Je voudrais essayer ici de m'écarter de cette vision des choses et tâcher d'inscrire l'homme Kolmogorov et ses réalisations dans son temps. Qu'il soit clair, si besoin était de le signaler, que ce que je vais dire ne diminue en rien la valeur et le mérite du grand mathématicien qu'il fut mais veut simplement souligner ce qui fut la fortune de son destin scientifique, à savoir l'adéquation qui exista entre son génie et les conditions dans lesquelles il s'est exercé.

Cet exposé comporte deux parties. Dans la première, après avoir présenté quelques éléments biographiques concernant Kolmogorov, je vais essayer de brosser un tableau du contexte historique dans lequel a pris place son évolution scientifique. Ensuite, dans le deuxième partie, je prendrai chronologiquement les grandes étapes de l'œuvre probabiliste la plus imposante du siècle dernier.

Une table chronologique sommaire a été ajoutée dans la section 4. Le lecteur désirant avoir des renseignements rapides sur l'histoire de l'URSS de la période concernée pourra se reporter à [Werth, 1995].

# 2 L'homme et son contexte historique

Puisqu'il est né en 1903, on peut dire que la formation scientifique de Kolmogorov, au sens de sa prise de contact avec les mathématiques les plus récentes, appartient entièrement à la période qui suit la révolution bolchevique de 1917. À n'en pas douter, cette "pureté" idéologique fondamentale a joué positivement dans sa carrière. Comme tout nouveau régime, le pouvoir bolchevique a très rapidement senti la nécessité de mettre en place une caste d'élite dont il puisse revendiquer la paternité et Kolmogorov se trouva en temps et lieu pour répondre à cette exigence. Une déclaration de Staline, qui date de 1939, illustre bien cette ambition: "Nous avons maitenant une intelligenstia nombreuse, nouvelle, populaire, *qui se différencie de l'ancienne intelligentsia bourgeoise*[1] non seulement par sa composition, mais aussi par son orientation socialo-politique."[2]. On peut d'ailleurs noter avec intérêt que la vie professionnelle du mathématicien a pour ainsi dire épousé la période soviétique: il avait 19 ans à la fondation de l'URSS et est mort trois ans avant son éclatement.

[1] Je souligne.

[2] Cité dans le compte-rendu du XVIIIème Congrès du PC(b), *Изв.Акад.Наук СССР*,1939

## 2.1 Aperçu biographique d'un savant soviétique

Andrei Nikolaievitch Kolmogorov est né le 25 avril 1903 à Tambov, à 500 km environ au sud-est de Moscou, lors d'un déplacement de sa mère de retour de Crimée. Sa mère, Maria Yakovlevna, mourut des suites de l'accouchement et son fils fut confié à ses deux sœurs à Tunoshna près de Iaroslavl. D'après le tableau qu'en fait [Shiryaev,1989], il semble que ces deux femmes, ainsi que le père d'Andrei Nikolaievitch, agronome scientifique (c'est à dire statisticien) reconnu, aient eu alors des activités sociales et politiques d'opposition au régime tsariste. Il faut peut être un peu tempérer le recensement naturellement complaisant qui a été fait de cette donnée pendant la période soviétique. Dans ces années où le régime impérial se crispait en une autocratie policière, crispations qui devaient conduire aux événements de 1905 et à la timide parlementarisation du régime, toute l'Intelligentsia russe bouillonnait. Quoi qu'il en soit, il est certain que le jeune Kolmogorov a grandi dans un milieu ouvert et qu'il a dû accueillir avec une certaine sympathie le renversement de l'ancien régime. C'est à Moscou, où Vera Yakovlevna et son fils adoptif avaient aménagé en 1910, que ceux-ci vécurent le tournant de 1917. En 1920, Andrei Nikolaievitch entre à l'Université de Moscou. [Shiryaev,1989] rapporte aussi que Kolmogorov s'inscrivit en parallèle pour des études d'histoire et également dans le département de Métallurgie de l'Institut Mendeleev. Dans ces années post-révolutionnaires, il était en effet fort bien vu, et en tout cas encouragé, de ne pas se cantonner à une discipline scientifique théorique: le jeune pouvoir des Soviets savait qu'il allait avoir besoin de cadres techniciens pour consolider son économie et survivre aux défis militaires en cours. Il n'y a aucune raison de douter que Kolmogorov ait suivi ce chemin par conviction personnelle.

Dès son entrée à l'université en 1920, Kolmogorov fait connaissance de Nikolai Nikolaievitch Luzin en suivant son cours de fonctions analytiques. En 1921, les cours de P.S.Alexandrov le font s'intéresser à la théorie descriptive des ensembles dans la lignée de Borel, Baire, Lebesgue et Souslin, domaine dans lequel il obtient plusieurs résultats intéressants dès 1922 (publiés à partir de 1925). Il devient entre temps un étudiant attitré de Luzin dans son groupe Luzitania (voir *infra*); celui-ci l'a remarqué après l'obtention en 1921 d'un joli résultat sur les séries de Fourier qui sera publié en 1923 et constitue le premier article de sa longue bibliographie. Dans cette atmosphère, il se familiarise rapidement avec les techniques de la théorie de la mesure et de l'intégration, domaines dans lesquels il publie plusieurs papiers entre 1925 et 1928. A partir de 1924, il commence à s'intéresser de près à la théorie des probabilités sur laquelle un premier travail (en collaboration avec Alexandr

Yakovlevitch Khinchin (1894-1959), leur seul papier commun d'ailleurs) paraît dès 1925. Cette année là, il obtient son diplôme universitaire (la thèse ne fut introduite en URSS qu'en 1934). En 1929, P.S.Alexandrov le fait nommer (contre le vœu d'Egorov) comme chercheur titulaire à l'Institut de Mathématique et Mécanique de Moscou. Kolmogorov et Alexandrov, de huit ans son aîné, vont dès ce moment devenir des amis intimes. En mars 1931, au retour d'un voyage de neuf mois en Allemagne et en France en compagnie de P.S.Alexandrov, Andrei Nikolaievitch est nommé professeur à l'Université de Moscou, et devient directeur de l'Institut de Mathématique en 1933. Il est élu à l'Académie des Sciences de l'URSS en 1939 et reçoit le prix Staline en 1941.

Dès les années 30, son importance à l'intérieur de la communauté mathématique soviétique ne cesse de s'accroître et les signes de déférence de la part du pouvoir se multiplient. En 1935, il est un des rédacteurs spécialisés en mathématique de la Большая Совиетская Энциклопедия(Grande Encyclopédie Soviétique), gigantesque entreprise de rédaction d'une encyclopédie aux couleurs de l'idéologie bolchevique mise en route en 1928 pour le dixième anniversaire de la révolution. A partir de 1937, il en devient l'unique spécialiste cité et rédige en 1938 le grand article "Mathématique" en 47 colonnes. On peut noter au passage que l'article "Probabilités" qui date de 1928 (Вероятность précède dans l'ordre alphabétique Математика) avait été rédigé par Khinchin; les démêlés de ce dernier pendant l'affaire Luzin ont dû jouer dans sa sortie de l'équipe rédactionnelle. Kolmogorov resta quasiment en permanence en contact avec l'Occident, en tout cas avec la France, ( sauf naturellement pendant la Seconde Guerre Mondiale) comme le montrent quelques lettres du fonds Fréchet de l'Académie des Sciences de Paris, même après que ce soit abattu la première vague de terreur de la Ejovchtchina en 1937. Il est en particulier impressionnant de voir dans ces archives des lettres de 1939, époque où tout contact était quasiment rompu entre les communautés scientifiques. Au passage, on peut quand même noter que le ton de certaines lettres à partir de 1934 montre un certain désarroi quand la nasse stalinienne commença à se faire de plus en plus étanche: de tels mouvements d'inquiétude chez quelqu'un d'aussi choyé que Kolmogorov donne une idée de l'atmosphère de l'époque.

Il semble cela dit que Kolmogorov se montra toujours plutôt docile avec le régime (ce qui n'est pas une critique), même s'il n'eut pas les engagements directs que purent avoir d'autres scientifiques. Il était un membre de la Nomenklatura et en avait privilèges et distinctions, voyageant quasiment sans entrave, et accédant aisément à la littérature technique la plus récente. On a mis en avant avec une certaine facilité depuis la fin de l'URSS que de vagues origines nobles lui auraient imposé une certaine prudence et un peu de zèle. C'est par exemple le cas de [Vucinich, 2000], qui d'ailleurs présente tout au long de son article Kolmogorov sous un jour particulièrement favorable, vantant par exemple son comportement pendant l'affaire Luzin où il aurait essayé de détourner les foudres qui étaient dirigées vers son ancien vé) (on peut noter cependant que [Демидов и Левшин, 1999] est plus réservé à ce sujet). Nous verrons par la suite quelle fut son attitude dans les débats idéologiques autour des mathématiques et du calcul des probabilités en particulier, mais il est aussi instructif de voir comment il intervint dans un débat qui n'était pas encore tranché [3], à savoir celui autour de la génétique, qui secoua la société scientifique soviétique à la

[3]Ce point atténue un peu la présentation de [Shiryaev,1989] qui semble souligner un exceptionnel

fin des années 30 avant d'être réglée idéologiquement en 1948 par Staline lors de l'Affaire Lysenko ([Krementsov, 1997]). En 1940, Kolmogorov publie un papier [4] où l'utilisation des méthodes de test statistiques lui permet d'affirmer que les expériences de l'élève de T.D.Lysenko, N.I.Ermaloyeva, sur l'hérédité des caractères acquis démontrent sans ambiguïté la validité des lois de Mendel. A partir de 1948, quand Staline aura décidé de faire de Lysenko le paradigme du savant soviétique, ce papier disparaît de ses publications comme le signale non sans ironie [VonPlato, 1994]. Autre signe, beaucoup plus tardif, de la docilité de Kolmogorov, l'article qu'il cosigne à 71 ans avec P.S.Alexandrov dans la Pravda du 15 février 1974 pour féliciter le gouvernement soviétique de la déchéance de Soljenitsyne de la nationalité soviétique[5].

Ceux qui l'ont connu, et naturellement avant tout ses étudiants qui forment une cohorte impressionnante parmi les plus grands mathématiciens soviétiques de l'après-guerre, soulignent combien Kolmogorov fut la figure d'un intellectuel complet, curieux de tout bien au delà des limites de la science, passionné d'art et des cultures du monde, parlant couramment français et allemand, entre autres. E.Mourier, qui l'a rencontré quelquefois dès la fin des années 50 m'a confirmé cette impression de distinction et d'intelligence prodigieuse qui se dégageaient de lui ainsi que l'intérêt très marqué qu'il portait aux jeunes mathématiciens en formation. A la fin de sa vie, il s'intéressa d'ailleurs de plus en plus aux questions d'éducation et fut en particulier très impliqué dans la rénovation des programmes de mathématiques dans le système scolaire soviétique. Il fonda en outre une école spéciale pour les jeunes doués en mathématique, dans la grande tradition de sélection étroite qui était la conception centrale de l'éducation soviétique.

## 2.2 L'école probabiliste russe avant la Révolution

### 2.2.1 Saint-Pétersbourg

Le fondateur de l'école probabiliste russe fut incontestablement Pafnuty Lvovich Chebyshev (1821-1894), professeur à Saint-Pétersbourg où il avait été recruté sur l'insistance de Bouniakovsky après son échec à Moscou. Avec lui, c'est là que se concentra pendant cinquante ans l'essentiel des recherches en mathématiques en Russie, et tout spécialement sur les probabilités. Chebyshev, dès le début de son travail mathématique s'intéressa aux probabilités, et en particulier au fait de rendre plus satisfaisantes les démonstrations des théorèmes limites fondamentaux (loi des grands nombres et théorème de la limite centrale)

courage *politique* (le courage, ou en tout cas l'honnêteté scientifique ne faisant, lui, aucun doute) dans le fait d'avoir publié ce papier. On lira cependant avec intérêt la réponse acerbe de T.D.Lysenko lui-même (*По поводу статьи академика А.Н.Колмогорова,*(Au sujet d'un article de l'académicien A.N.Kolmogorov)Докл.Акад.Наук СССР,XXVIII,9,1940).

[4] *Об одном новом подтверждении законов Менделя* (Sur une nouvelle confirmation des lois de Mendel), Докл.Ак.Н., XXVII,1,1940

[5] *La trahison ne se pardonne pas.* "C'est avec un sentiment de grande satisfaction que nous accueillons la nouvelle que le Presidium du Soviet Suprême a privé A.Soljenitsyne de la nationalité soviétique [...]. Le sentiment qui habite avant tout l'Intelligentsia soviétique est celui du devoir devant le peuple et l'Etat, de la fierté pour les grandes réalisations des Soviétiques qui construisent le communisme .[...] Dans ses œuvres, publiées en Occident, A.Soljenitsyne noircit notre société, offense la mémoire de ceux qui sont tombés dans les combats de la Grande Guerre Patriotique[...]" (traduction personnelle)

et de préciser les hypothèses sous lesquelles ils étaient applicables. Le préambule de sa thèse, soutenue en 1844 à Moscou, donne une idée de son but: "Donner la possibilité de croire à toutes ces conclusions par une analyse simple et précise, accessible à la majorité des étudiants, c'est un grand pas vers un exposé élémentaire de la théorie des probabilités". Chebyshev . L'arrière-plan des recherches de Chebyshev en probabilités fut toujours orienté vers la justification des méthodes exploitées dans des applications. [Юшкевич, 1968] expose la façon dont ses intérêts pour la balistique dans les années 1860 et pour la mécanique furent à l'origine de la découverte de la célebrissime inégalité qui porte son nom, et de la méthode des moments. Il n'est pas sans intérêt de citer Kolmogorov lui-même, dans son rapport de 1947 [А.Н.Колмогоров, 1947] même s'il faut lire avec un certain filtre ce qui s'écrivait dans cette année qui précède de peu le déclenchement des campagnes de purges de Zhdanov, particulièrement dirigées contre les milieux intellectuels souvent suspects de cosmopolitisme. Le nationalisme pan-russe était alors en pleine explosion et Chebyshev fut élevé au rang d'emblême. Kolmogorov écrit[6]: "[L]e principal, c'est que Chebychev a cherché en permanence à obtenir des estimations précises de l'écart avec les lois limites [théoriques], qui soient valables pour un nombre éventuellement grand, mais *fini*[7] d'expériences [...]. Plus, Chebychev le premier a pris en compte et employé dans toute sa force la signification de "variable aléatoire" et d' "espérance mathématique"(valeur moyenne) d'une variable aléatoire. Le sens de ces expressions était connu auparavant et apparaissait comme un prolongement de la compréhension des expressions "événement" et "probabilité". Mais les variables aléatoires et leurs espérances mathématiques prenaient un sens bien plus précis par un algorithme pratique et souple." Comme on le voit, la théorie des Collectifs, d'après Kolmogorov, germait déjà dans l'école probabiliste de Pétersbourg!

Les élèves de Chebyshev, Andrej Andrejevitch Markov (1856-1922) et A.M.Lyapunov (1857-1918) reprirent le flambeau de leur maître. Chebyshev fit recruter Markov dès 1886 à l'Académie Impériale des Sciences. Seneta [Seneta, 2001] observe avec justesse que A.A.Markov, qui vécut jusqu'en 1922, eut pendant le régime tsariste une courageuse attitude de résistance face à certaines des mesures les plus répressives prises par le gouvernement, ce qui naturellement, ne pouvait qu'attirer un œil bienveillant du régime bolchevique sur lui, et par ricochet sur l'école de Petrograd (nom pris par Saint-Pétersbourg pendant la Première Guerre Mondiale). Markov et Lyapunov continuèrent l'étude des théorèmes limites en affaiblissant notemment les hypothèses concernant l'identité des lois dans le théorème de la limite centrale . On doit à Lyapunov le premier développement systématique de la théorie des fonctions caractéristiques que Lagrange et Laplace n'avaient qu'ébauchée. C'est en cherchant un contre-exemple à l'hypothèse d'indépendance pour la loi des grands nombres que Markov formula en 1907 un premier modèle des chaînes auxquelles il allait donner son nom. [Seneta, 2001] affirme qu'il le fit en réaction à un article de Nekrassov qui voyait dans l'observation usuelle dans la vie courante des phénomènes de moyennisation, adjointe à la nécessité de l'indépendance pour assurer la convergence dans la loi des grands nombres, une preuve de l'existence du libre-arbitre. Quoi qu'il en soit, il ne fait pas de doute qu'à l'éclatement des événements de 1917, Petrograd était le grand centre de mathématiques

[6] traduction personnelle

[7] Je souligne.

appliquées en Russie.

## 2.3 Moscou

C'est au début du vingtième siècle que l'Ecole mathématique de Moscou, l'autre capitale de l'Empire, prit de plus en plus d'importance et apparut comme une rivale pour Saint-Pétersbourg. Les noms emblématiques de ce groupe étaient Dmitri Fedorovitch Egorov (1869-1931) et son élève Nikolai Nikolaievitch Luzin (1883-1950) qui furent à l'origine d'un mouvement formaliste extrêmement puissant autour de la théorie des fonctions et des espaces fonctionnels abstraits dans la lignée des travaux de l'école française de Borel, Baire, Lebesgue et Fréchet. On doit à cette période les grands travaux sur l'utilisation de la théorie de la mesure pour la théorie des fonctions. Cette recherche formelle n'allait d'ailleurs pas sans lien avec de profondes considérations philosophiques et religieuses. Luzin et Egorov entretenaient avec les milieux théologiques orthodoxes des liens étroits, notamment avec l'étonnant Pavel Alexandrovitch Florensky(1882-1937), ami intime de Luzin[8] qui mena de front des études de mathématiques à l'Université de Moscou et de religion à l'Académie de Théologie où il soutint en 1912 une thèse cherchant, dans la lignée de V.Soloviev, une interprétation mystique des sciences. Luzin lui même avait quelque peu étudié la théologie. Par ailleurs, sa formation mathématique avait été influencée par son voyage en Allemagne à Göttingen où l'école formaliste autour de Hilbert règnait en maître. De retour à Moscou, où il soutint sa thèse en 1914, il fonda en 1917 avec l'aide d'Egorov un groupe de recherches de pointe sur les mathématiques fondamentales (auquel fut donné le surnom affectueux "Luzitania"[9]) auquel participaient ses étudiants, dont Kolmogorov à partir de 1922.

## 2.4 Deux outsiders: Bernstein et Slutsky

Deux autres fortes personnalités apparaissent dans le tableau que je dresse à grands traits.

Serge Natanevitch Bernstein (1880-1968), a joué un rôle majeur dans l'émergence des probabilités modernes. Présent à Paris à partir de 1898 où il entreprend une thèse à la Sorbonne sur les équations elliptiques qu'il soutient en 1904, il a pu côtoyer tout le creuset d'où allait sortir la conception contemporaine de la discipline. Il a été un point de contact essentiel entre les communautés mathématiques russes et françaises. À partir de 1911, à Kharkov où il a enseigne depuis 1907, il commence à travailler dans le domaine des probabilités et propose en 1917 un essai d'axiomatisation des fondements [Бернштейн, 1917] sur laquelle je reviendrai plus loin. Vucinich [Vucinich, 1999] fait remarquer l'attitude peu conciliante de Bernstein après la Révolution envers les exigences idéologiques du nouveau régime. C'est sans doute sa réputation internationale bien établie dès le début de l'ère soviétique qui lui permit de traverser sans encombre les années les plus noires de la période, la rétorsion se limitant en gros à le dissimuler partiellement derrière Khinchin et Kolmogorov.

Enfin, comme le signale [Юшкевич, 1968] il est nécessaire de mentionner ici le rôle joué par Evgeni Evgenievitch Slutsky (1880-1948). Dans les années qui précèdent la Révolution,

[8] La correspondance entre Luzin et Florensky a fait l'objet d'une étude intéressante dans [Демидов, Паршин и Половинкин, 1989].

[9] Probablement en allusion au Lusitania coulé en 1916 qui fut prétexte à l'entrée en guerre des USA.

il publie à Kiev, où il a difficilement terminé ses études en 1911 en raison de différents problèmes politiques, quelques travaux autour des méthodes mathématiques pour les statistiques, et notemment une monographie [Слуцкий, 1912] qui en propose un exposé très moderne. Dans les années 20, il se lancera dans une œuvre pionnière en calcul stochastique dont les travaux seront reconnus comme tels par Kolmogorov et Khinchin. Ses recherches furent cependant peu connues hors des frontières de l'URSS. On observe néanmoins dans la correspondance entre Lévy et Fréchet ([Barbut, Locker et Mazliak, 2003]) que le grand voyageur que fut Fréchet, un des seuls à maintenir le contact avec l'URSS, permit d'en faire connaître quelques unes en France à la fin des années 30. On pourra se reporter avec avantage à [Locker, 2001] pour des informations détaillées à ce sujet ainsi qu'à l'élogieuse nécrologie de Slutsky écrite par Kolmogorov en 1948 (Усп.Мат.Наук,III,4,142-151).

## 2.5 Nouveaux enjeux des mathématiques et des probabilités

La tourmente révolutionnaire et la nouvelle idéologie règnant à partir de 1917 obligèrent les mathématiciens russes de l'ancienne école à modifier progressivement le regard sur leur discipline et sur les stratégies de recherche à suivre. Ces remises en question favorisèrent l'émergence d'une nouvelle génération naturellement plus en phase avec l'idéologie du régime. Comme le signale [Vucinich, 1999], les années 20 furent une période où l'on tenta de mettre en place une conception des mathématiques adéquate, la difficulté venant du fait qu'il n'existait pas chez les Pères (Marx,Engels ou Lénine) de textes vraiment pertinents sur cette question. Cependant, dans *Matérialisme et Empiriocriticisme* (1909) qui fut érigé en évangile pour les scientifiques, Lénine ne cache pas une certaine défiance envers une discipline qui cède facilement à un idéalisme petit bourgeois, comme l'illustre la fracassante citation suivante: "Hermann Cohen[10], ravi, comme nous l'avons vu, de l'esprit idéaliste de la physique nouvelle, en arrive à recommander l'enseignement des mathématiques supérieures dans les écoles, cela afin de faire pénétrer dans l'intelligence des lycéens l'esprit idéaliste évincé par notre époque matérialiste. Ce n'est là assurément que le rêve absurde d'un réactionnaire: en réalité, il n'y a, il ne peut y avoir là qu'un engouement momentané d'un petit groupe de spécialistes pour l'idéalisme. Mais il est significatif au plus haut point que les représentants de la bourgeoisie instruite, pareils à un naufragé qui s'attache à un brin de paille, recourent aux moyens les plus raffinés pour trouver ou garder artificiellement une place modeste au fidéisme engendré au sein des masses populaires par l'ignorance, l'hébétude et l'absurde sauvagerie des contradictions capitalistes."([Lénine, 1970],V,8). C'est certes plutôt la crise des fondements de la physique qui occupe bruyamment le devant de la scène intellectuelle de ces années prérévolutionnaires, mais les mathématiques contemporaines sont comme on l'a vu concernées. Les physiciens empiristes (Mach, cible principale de Lénine) les voient comme une simple idéalisation de l'expérience sensible alors que les idéalistes ( à la suite de Poincaré) leur attribuent une valeur synthétique *a priori* (qui néanmoins être passée au crible de l'expérience). Les deux tendances se retrouvent en tout cas en porte à faux avec le matérialisme dialectique.

[10]H.Cohen (1842-1918), philosophe allemand, chef de lécole de Marburg.

On peut raisonnablement penser que les conceptions très abstraites de l'Ecole de Moscou autour d'Egorov et Luzin soient devenues peu favorisées aux yeux du régime et que les nouvelles instances idéologiques auraient été satisfaites de pouvoir donner plus d'influence à d'autres sphères de recherche plus proches de slogans léninistes qui réclamaient, comme le signale [Graham, 1993], une unification de la théorie et de la pratique. En ce sens, la phrase lapidaire qui inaugure le grand article de Kolmogorov [Колмогоров, 1938], "La mathématique est la science qui traite des formes et des relations formelles et quantitatives du monde réel", donne le ton. On pourra aussi se reporter à l'instructif résumé du XVIIIème Congrès du Parti publié en introduction du Volume de 1939 des *Известия Академии Наук СССР*. Il est intéressant à ce sujet de voir une description de la situation des mathématiciens soviétiques de la fin des années 20 dans la bouche de Luzin lui-même. Quand il dut se défendre avec l'énergie du désespoir au milieu de la tempête qui s'abattit sur lui en juillet 1936, il eut à répondre de l'accusation de soumission à l'Occident dont témoignaient ses multiples publications dans des journaux occidentaux. Dans la sténographie des houleuses réunions de la Commission de l'Académie des Sciences qui nous a été miraculeusement conservée, Luzin s'écrie[11]: "Pouvais-je ignorer l'opinion du camarade Kolman, qui occupait une haute position et qui me prenait à partie pour l'abstraction théorique et la nuisance de mes travaux? Non, bien sûr. Mon but était de ne pas nuire à ma patrie. Et si ces personnes qui ont une vision claire de ces choses trouvaient qu'une activité théorique était nuisible, eh bien qu'elle ne nuise plus. Mais de moi sortaient ces théorèmes. Que pouvais-je faire? Quoi, rentrer en moi-même? Et j'ai décidé que tout ce en quoi je pouvais aider mon pays dans une direction appliquée, je le publierai chez nous. Et dans cette direction qu'on déclarait nuisible dans mon activité(...), de publier à l'étranger. Car je devais tenir compte de la situation. Jugez vous mêmes si continuer avec entêtement à publier des travaux théoriques qui non seulement n'ont pas d'application à l'heure actuelle, mais dont l'éventualité d'une future application n'est pas prévisible n'est pas une obstruction de notre édition scientifique."

Comme on sait, les années 20 voient se développer une opposition au formalisme hilbertien sous la forme de l'intuitionisme défendu par Brouwer, qui refuse de voir dans la simple consistence logique un principe d'existence. Une telle théorie était naturellement bien mieux en conformité avec les affirmations (pseudo-)scientifiques de Lénine et on comprend aisément qu'elle ait trouvé les faveurs de la nouvelle idéologie dominante, du moins après qu'en soient gommés les aspects idéalistes (accusés par la langue de bois soviétique de "psychologisme") comme ceux formalisés par Heyting en 1930 [12]. Significatif est le fait que le livre de 1948 [*Математика, 1948*] survolant les immenses réalisations mathématiques de la période soviétique depuis 30 ans accorde une place importante et primordiale à ces questions de logique formelle [13] sous la plume très autorisée de S.A.Yanovskaya (1896-1966), qui depuis le début des années 30 était le Fouquier-Tinville de la communauté mathématique

[11] Voir [Демидов и Левшин, 1999]. Traduction personnelle.

[12] "Les objets mathématiques atteignent sans intermédiaire l'esprit pensant. De ce fait, la connaissance mathématique ne dépend pas de l'expérience"[Heyting, 1930], une déclaration inacceptable dans l'optique léniniste.

[13] Et pourtant ce livre a été écrit avant la grande tempête qui devait suivre l'affaire Lyssenko où les attaques contre le formalisme devaient prendre une tournure encore plus dramatique

(elle fut un fer de lance de l'affaire Luzin d'après [Vucinich, 2000])[14]. C'est le premier thème abordé: sur près de trente pages, une dizaine est consacrée, dans la plus pure rhétorique de la période stalinienne, à un exposé détaillé de la pensée marxiste-léniniste.

Il est remarquable que ce soit les deux futurs leaders probabilistes (en 1925 Kolmogorov [Колмогоров,1925] et en 1926 Khinchin [Хинчин, 1926]) qui successivement mettent en place dans deux articles l'interprétation orthodoxe de la question des fondements et du contenu des mathématiques, interprétation qui sera érigée en dogme pour toute la période soviétique et pour laquelle une vision pratique du calcul des probabilités pouvait servir de paradigme. En effet, puisque les créateurs de l'Ecole russe avaient développé le calcul des probabilités dans une optique plutôt pratique, on devait pouvoir y contourner le problème de l'idéalisme assez aisément. Affirmer que la discipline restait à l'interface avec le monde "réel" pouvait s'appuyer sur l'autorité de Chebyshev, Markov et Lyapunov. De ce fait, la faveur exceptionnelle dont le calcul des probabilités a joui à travers la période soviétique (dans [*Математика, 1948*], 51 pages sont consacrées à cette thématique) n'est pas étrangère à ce caractère spécifiquement appliqué [15]. Dans la partie du rapport de Vavilov[Вавилов,1946] consacrée aux problèmes scientifiques fondamentaux auxquels l'Académie des Sciences devra se consacrer dans le prochain plan quinquennal, les probabilités sont la première discipline mathématique citée, spécifiquement pour leur aspect applicatif en prise avec les sciences de la nature ou les sciences humaines comme l'économie.

Il faut cependant nuancer une vision trop heureuse de la situation des mathématiques de l'aléatoire. Comme toute période mouvementée, celle-ci fut traversée de courants contradictoires. Dans son étude déjà citée, [Vucinich, 1999] Vucinich fait remarquer que l'aléatoire en tant que tel se prétait mal à une interprétation du type matérialisme-dialectique, et de ce fait restait assez suspect aux yeux de certains qui se chargeaient dès les années 20 de formuler les contours du nouveau dogme. Vucinich cite par exemple une contribution au symposium consacré aux relations entre le matérialisme-dialectique et les mathématiques organisé en 1931 sous l'égide de l'Académie Communiste: l'auteur y regrette l'intérêt porté par l'Ecole de Leningrad aux fluctuations des variables aléatoires qu'il trouve plus en rapport avec le comportement anarchique de l'économie capitaliste qu'avec celui de l'économie planifiée socialiste.

Quoi qu'il en soit, à partir de 1924 les probabilités vont devenir un des thèmes majeurs de l'école mathématique soviétique sous l'impulsion de ces deux fortes personnalités choyées par le régime et l'URSS pendant vingt ans le leader incontesté de la discipline, s'emparant du flambeau détenu par la France à la suite de Poincaré et de Borel.

[14] Au moment où j'écrivais cette phrase, j'ai découvert la nécrologie consacrée à S.A.Yanovskaya par son élève B.A.Trakhtenbrot (*Modern Logic*,7,2,1997,160-187), qui en donne naturellement un portait plus contrasté.

[15] qui, de manière ironique, est aussi la cause de son peu de considération par l'école structuraliste française... [Bourbaki, 1999]

# 3 Quelques étapes de l'œuvre probabiliste

Comme je l'ai déjà mentionné, l'immensité de l'œuvre probabiliste de Kolmogorov rend impossible d'en faire ici un exposé exhaustif. On se reportera naturellement à [Shiryaev,1989] pour un passage en revue de l'ensemble des articles du mathématicien, ainsi que pour la liste complète de leurs références. Par ailleurs, comme également mentionné, la séparation des aspects probabilistes du travail du savant soviétique des autres domaines d'activité est souvent artificielle. Cela dit, dans cette partie, je vais essayer de présenter quelques unes des étapes les plus remarquables de ses recherches, dans une présentation chronologique. Par ailleurs, l'étendue de la vie productive de Kolmogorov oblige là encore à restreindre le champ étudié, et je vais en gros poser deux limitations assez arbitraires. En premier lieu, une contrainte de temps: les travaux considérés s'étendront dans la période qui commence en 1925 pour s'achever dans les premières années qui suivirent la mort de Staline en 1953. En second lieu, je n'évoquerai pas les travaux essentiels de Kolmogorov qui touchent à des phénomènes physiques ou mécaniques comme la théorie de la turbulence ou la théorie des perturbations (théorie de Kolmogorov-Arnol'd-Moser). Au sujet des phénomènes turbulents, on pourra consulter le très complet dossier formé par les deux articles de B.Bru [Bru, 1998] et J.Bass[Bass, 1999] parus dans MATAPLI. Pour ce qui concerne la théorie KAM, voir le paragraphe 3.3 *infra*.

## 3.1 lois des grands nombres et processus de Markov

### 3.1.1 Convergence des séries

Comme signalé, le premier article de Kolmogorov dans le champ probabiliste est son seul papier commun avec Khinchin. Il date de 1925 et s'intitule *Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden*(Sur la convergence de séries dont les termes dépendent du hasard) (Mat.Sb., 32, 1925,658-667) et étudie de façon approfondie les conditions de convergence pour une série de variables aléatoires. Le recensement en termes modernes fait par [Shiryaev,1989] est le suivant: pour une suite $(\xi_n)_{n\geq 0}$ de variables aléatoires indépendantes,

(i) Si $\sum_n E(\xi_n)$ et $\sum_n \mathrm{Var}(\xi_n)$ convergent, la série $\sum_n \xi_n$ converge p.s.

(ii) Si les $\xi_n$ sont des variables uniformément bornées, la condition est aussi nécessaire.

(iii) Si $c > 0$ est une constante donnée, et si $\xi_n^c = \xi_n \mathbb{1}_{|\xi_n|\leq c}$, la convergence des trois séries $\sum_n E(\xi_n^c), \sum_n \mathrm{Var}(\xi_n^c), \sum_n P(\mid \xi_n \mid \geq c)$ suffit à assurer la convergence p.s. de $\sum_n \xi_n$ (résultat aujourd'hui connu et enseigné sous le nom de "théorème des trois séries").

### 3.1.2 Loi des grands nombres

On a vu que le grand centre d'intérêt de Chebyshev et de Markov a été l'étude approfondie de conditions pour la validité des lois limites. À la suite de Markov et Lyapunov, Kolmogorov réétudie la loi des grands nombres et formule en 1927 dans *Über die Summen durch den Zufall bestimmter unabhängiger Grössen*(Sur les sommes de quantités indépendantes dépendant du hasard), Math.Ann.,99,309-319, une loi des grands nombres généralisée. Soit $(\xi_n)_{n\geq 0}$ une suite de variables aléatoires indépendantes et $S_n = \xi_1 + \ldots + \xi_n$. La condition

nécessaire et suffisante pour que $\frac{S_n - E(S_n)}{n}$ converge en probabilité est qu'il existe ce qu'on appelle aujourd'hui un tableau triangulaire (voir [Billingsley, 1986]), soit une suite double $(\overline{\xi}_{n,k})_{n\geq 1, 1\leq k\leq n}$ de variables aléatoires, indépendantes pour chaque $n$, telles que

$$\sum_{k=1}^{n} P(\xi_k \neq \overline{\xi}_{nk}) \to 0, \frac{1}{n}\sum_{k=1}^{n}[E(\xi_k) - E(\overline{\xi}_{nk})] \to 0, \frac{1}{n^2}\sum_{k=1}^{n} \mathrm{Var}\overline{\xi}_{nk} \to 0.$$

Il réobtient ainsi la loi ("semi"-faible) des grands nombres obtenue par Khinchin en 1926 (donnant la convergence en probabilités des moyennes arithmétiques d'une suite de variables indépendantes et de même loi, sous l'hypothèse d'une espérance finie), et finira en 1930 par énoncer dans *Sur la loi forte des grands nombres*, CRAS Paris, 191,1930, 910-912, note présentée par Hadamard, la loi forte sous son aspect classique: si $(\xi_n)$ est une suite de variables aléatoires indépendantes de variance finie telle que $\sum_n \frac{\mathrm{Var}(\xi_n)}{n^2} < +\infty$, alors p.s., $\frac{S_n - E(S_n)}{n} \to 0$. La démonstration de Kolmogorov s'appuie sur l'inégalité classique $P(\max_{1\leq k\leq n} \mid S_k \mid \geq \varepsilon) \leq \frac{E(S_n^2)}{\varepsilon^2}$ qu'il a prouvée dans son article de 1927.

### 3.1.3 Loi du logarithme itéré

En 1924, Khinchin avait prouvé la loi du logarithme itéré pour un schéma de Bernoulli uniforme et indépendant. Reprenant ce résultat, Kolmogorov le généralise dans son article de 1929 (mais écrit en 1927), *Über das Gesetz des iterierten Logarithmus*(Sur la loi du logarithme itéré), Math.Ann., 101,1929, 126-135, sous la forme suivante. Soit $(\xi_n)_{n\geq 0}$ une suite de variables aléatoires indépendantes de moyenne nulle et de variance $\sigma_n^2$ finie. On pose $B_n = \sum_{k=1}^{n} \sigma_k^2$. Supposons qu'il existe une suite de constantes $(M_n)_{n\geq 0}$ telle que $M_n = o((\frac{B_n}{\ln\ln B_n})^{1/2})$ et p.s. $\mid \xi_n \mid \leq M_n$. Alors, p.s.,

$$\overline{\lim} \frac{S_n}{\sqrt{2B_n \ln\ln B_n}} = 1.$$

### 3.1.4 Processus de Markov en temps continu

Le début de l'année 1931 voit la publication de l'un des papiers les plus éblouissants de Kolmogorov *Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung*(Sur les méthodes analytiques du Calcul des Probabilités), Math.Ann.104,415-458 portant sur l'étude systématique des transitions des processus de Markov à temps continu. La partie technique de l'article est précédée d'une longue introduction (pp.415-417) où l'auteur expose sa conception distinguant le phénomène réel ("naturel ou social" écrit-il) et l'équation d'état qui n'est qu'un schéma de la réalité, la considération d'un processus stochastique n'étant justifiée que par le fait que l'état n'est pas parfaitement connu (par exemple lors de l'application de la mécanique (*classique* précise le mathématicien...)). De ce fait, il est loisible de supposer qu'on a une transition reliant la position $x$ à l'instant $t_0$ à un *ensemble* de positions envisageables à l'instant ultérieur $t$ réparties aléatoirement suivant une

distribution $P(t_0, x, t, .)$. Il est à noter que Kolmogorov cite explicitement les travaux de Bachelier, et notamment sa (maintenant célèbre) thèse (*Théorie de la spéculation*, Ann.ENS,17,1900,21), comme origine de ses idées sur les fonctions de transition dans le cas continu. À l'époque, comme on le sait, Bachelier était totalement méconnu en France, et ce n'est pas le moindre mérite du mathématicien soviétique d'avoir su voir tout ce qu'il y avait à tirer de ces superbes travaux.

Partant de l'équation intégrale du semi-groupe de transition (aujourd'hui appelée Equation de Chapman-Kolmogorov - elle avait été signalée par Chapman dans son mémoire de 1928 sur le mouvement brownien [Chapman, 1928]),

$$P(s,x;t,A) = \int P(s,x;u,dy)P(u,y;t,A), 0 \leq s < u < t,$$

dans le cas d'existence d'une densité (c'est à dire si on peut définir $f(s,x;t,y) = \frac{\partial F(s,x;t,y)}{\partial y}$ où $F(s,x;t,y) = P(s,x;t,]-\infty,y])$), il obtient les équations aux dérivées partielles satisfaites par $f$, constituant ainsi le premier cas d'introduction faible d'un processus (c'est à dire en loi) et de l'interprétation probabiliste systématique de la solution d'une famille d'EDP et ouvrant ainsi la voie au calcul stochastique d'Itô et, plus proche de nous, aux problèmes de martingales développés par Stoock et Varadhan. Kolmogorov introduit les caractéristiques instantanées du processus, jouant le rôle d'espérance et de variance instantanées, sous la forme des fonctions

$$A(s,x) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \int_{-\infty}^{\infty} (y-x) f(s,x;s+\Delta,y)dy, B^2(s,x) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \int_{-\infty}^{\infty} (y-x)^2 f(s,x;s+\Delta,y)dy,$$

et montre que sont satisfaites les deux EDP suivantes

$$(i) - \frac{\partial}{\partial s} f(s,x;t,y) = A(s,x)\frac{\partial}{\partial x} f(s,x;t,y) + \frac{B^2(s,x)}{2}\frac{\partial^2}{\partial x^2} f(s,x;t,y)$$

$$(ii) \frac{\partial f(s,x;t,y)}{\partial t} = -\frac{\partial}{\partial y}[A(t,y)f(s,x;t,y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[B^2(t,y)f(s,x;t,y)]$$

### 3.1.5 Application à la statistique mathématique

En 1933, est publié l'important article *Sulla determinazione empirica di una legge di distribuzione* (Sur la détermination empirique d'une loi de probabilité), Gior.Ist.Ital.Att.,4,83-91 qui marque en quelque sorte la prise de contact de Kolmogorov avec la statistique d'estimation, prolongement naturel des études sur la loi forte des grands nombres. Considérant une suite de variables aléatoires indépendantes $(\xi_n)_{n\geq 1}$ de même loi de fonction de répartition $F(x) = P(\xi_1 \leq x)$, on introduit la fonction de répartition empirique $F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\xi_k \leq x}$. Alors, on a

$$\lim_{n\to+\infty} P(\sqrt{n} \sup_{x\in\mathbb{R}} \mid F_n(x) - F(x) \mid \leq \lambda) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2\lambda^2 k^2}.$$

Ce résultat est naturellement à rapprocher du théorème de Glivenko-Cantelli stipulant la convergence p.s. vers 0 de $\sup_{x\in\mathbb{R}} \mid F_n(x) - F(x) \mid$. On peut d'ailleurs voir un indice de l'agitation des esprits en ces années cruciales autour de l'estimation statistique dans le fait que ce dernier résultat a été publié simultanément par Glivenko (cas où $F$ est continue) et Cantelli (cas général) dans deux articles portant le même titre que celui de Kolmogorov, et dans le même numéro du *Giornale dell Istituto Italiano degli Attuari*!

## 3.2 La formalisation comme aiguillon: les Grundbegriffe

Nous abordons maintenant la publication de Kolmogorov qui fit sans doute le plus pour sa réputation internationale. J'en veux pour indice le fait observable aujourd'hui, soixante-dix ans après, que pour la plupart des probabilistes, l'époque moderne du calcul des probabilités a commencé avec la parution en 1933 des *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Notions fondamentales du Calcul des Probabilités), au sens où tout ce qui avait été fait auparavant restait dans une imprécision complète de cadre formel. Or, comme cela a déjà été signalé, et même sans remonter au traité de Borel de 1909 où apparurent pour la première fois clairement les bénéfices que les probabilités pourraient retirer d'un traitement par le biais des outils de la théorie de la mesure (mais Borel n'eut pas comme on le sait, peut être par peu de considération pour la discipline, le désir de mener jusqu'au bout cette investigation), il y eut avant Kolmogorov plusieurs tentatives de normalisation et de formalisation, ce qui fait qu'une des questions centrales est de comprendre pourquoi c'est seulement la tentative de Kolmogorov qui eut le retentissement que l'on sait. Lors d'une conversation avec Bernard Bru, celui-ci m'a fait part de son étonnement du fait que l'on ait accordé une importance démesurée à l'axiomatisation par Kolmogorov des probabilités par rapport à d'autres de ses travaux qui lui semblent, à juste titre, beaucoup plus exceptionnels par leur profondeur et leur originalité. Je pense cependant qu'il ne faut pas sous-estimer le rôle pour ainsi dire politique qu'ont pris les *Grundbegriffe*. Kolmogorov a trente ans en 1933, et il revient de son long séjour en Occident (en particulier à Göttingen où règne l'esprit d'Hilbert). Il a sans doute mieux pris conscience à cette occasion de l'intérêt grandissant pour la discipline et l'attente d'une formalisation efficace. Peut-être aussi certaines rencontres comme celle de Fréchet, qui l'a invité pendant plusieurs semaines dans sa maison du Var pour parler de processus de Markov, lui ont montré que le moment était bien choisi pour publier, qui plus est chez un éditeur international, un cadre précis permettant de manipuler des processus à temps continu et des calculs conditionnels.

Avant de rentrer plus avant dans la monographie de Kolmogorov[16], quelques mots des deux tentatives importantes qui l'avaient précédée. En 1917, juste avant la Révolution d'Octobre, S.Bernstein avait publié un long article intitulé *Опыт аксиоматического обоснования теории вероятностей* (Essai de fondement axiomatique de la théorie des probabilités) Сообщ.Карьк.матем.о-ва,15,209-274,1917. Dans cet article, suivant l'explication de [Юшкевич, 1968] qui les tire de la monographie de 1927 de Bernstein *Теория вероятностей* (Théorie des Probabilités), Bernstein énonce trois axiomes fondamentaux[17] desquels il prétend déduire les

[16] On se reportera aussi avec intérêt à l'analyse détaillée que fait de ces questions [VonPlato, 1994] dans son chapitre 7.

[17] [VonPlato, 1994], p.275, fait remarquer la coïncidence entre ces axiomes et ceux que proposera une

propositions de la théorie.

Ce sont:

1) L'axiome de comparaison: si $a$ est un sous-événement de $A$, alors la probabilité de $a$ est inférieure à celle de $A$. Inversement, si la probabilité de $a_1$ est strictement inférieure à celle de $A$, alors il existe un sous-événement $a$ de $A$ tel que sa probabilité soit égale à celle de $a_1$.

2) L'axiome sur les événements incompatibles: si les événements $A$ et $A_1$ d'une part, et $B$ et $B_1$ d'autre part sont incompatibles, et que $A$ et $B$ d'une part, $A_1$ et $B_1$ d'autre part ont mêmes probabilités, alors la probabilité de l'événement ($A$ ou $A_1$) est la même que celle de l'événement ($B$ ou $B_1$).

De ces deux axiomes, se déduit en particulier la forme standard de la probabilité uniforme dans un univers fini, au sens où un événement composé de $m$ cas élémentaires parmi $n$ également possibles, se voit attribuer une probabilité qui est une fonction croissante de la fraction $\frac{m}{n}$: le plus simple est alors de choisir précisément $\frac{m}{n}$ pour cette probabilité. L'importance de cette justification de la probabilité uniforme est sans doute à cette époque d'une importance stratégique. On peut rappeler que dans la deuxième édition de son Cours de Probabilités (en 1912 donc, soit cinq ans seulement avant l'article de Bernstein), Poincaré ne fournit encore comme seule définition effective de la probabilité que le *rapport du nombre de cas favorables au nombre de cas possibles.*

3) L'axiome de "rencontre" (совмещение) des événements: la probabilité de la rencontre des événements $A$ et $B$ dépend exclusivement de la probabilité de $A$ et de celle attribuée à $B$ après la réalisation de $A$. S'en déduit naturellement la définition usuelle de la probabilité conditionnelle.

En 1919 à Vienne, Richard von Mises publie un article au titre fort voisin du futur traité de Kolmogorov, et sans nul doute lui aussi inspiré par Hilbert, *Grundlagen der Wahrscheinlichkeitsrechnung* (Fondements du Calcul des Probabilités), Mat.Zeit.,5,52-99. Comme il est bien connu, il s'agit là de l'acte fondateur d'une théorie fréquentiste des probabilités sous la forme de suites *imitant* le hasard, ou théorie des Collectifs, qui fut à l'origine de discussions acharnées dans les années 30: on sait par exemple que la thèse de Jean Ville réalisée sous la direction de Fréchet et soutenue en 1939 était intitulée *Critique de la théorie des Collectifs.* Il n'est pas question ici de développer cet immense sujet. Le lecteur pourra se référer à [VonPlato, 1994] pour une analyse détaillée du modèle proposé par von Mises, et aussi à [Barbut, Locker et Mazliak, 2003] ou [Cavaillès, 1940] pour lire des commentaires contemporains de ces discussions. On peut en tout cas observer que le système de von Mises a sans doute attiré l'attention des probabilistes russes puisque Khinchin en 1929 publie un article critique (*Учение Мизеса о вероятностях и принципы физической статистики*(Les études de Mises sur les probabilités et les principes de la statistique physique)Усп.Физ.Наук,9,141-166,1929).

Les *Grundbegriffe der Wahrscheinlichkeitsrechnung*, publiés en 1933 par Springer, se présentent comme un petit livre broché de 62 pages, qui surprend au premier abord par sa modestie. Dans l'introduction du premier chapitre consacré à la théorie élémentaire, Kolmogorov se place dans l'optique des *Grundlagen* de Hilbert: il s'agit pour lui de définir

quinzaine d'années plus tard de Finetti pour sa construction des *probabilités qualitatives.*

un système axiomatique pour faire des probabilités une discipline mathématique à l'instar de la géométrie. Kolmogorov prend soin de signaler que d'autres axiomatisations que la sienne sont possibles, et cite les articles de Bernstein et von Mises que j'ai mentionnés plus haut. On remarquera d'ailleurs avec intérêt que pour justifier l'énoncé des axiomes, l'auteur entend se placer d'emblée (cf. page 3, note 1) dans la lignée de la conception fréquentiste de von Mises: les propriétés de la probabilité élémentaire (en tant que fonction d'ensemble additive normée définie sur une algèbre de Boole) sont induites par celles des fréquences relatives des expériences répétées. [VonPlato, 1994] fait cependant remarquer que derrière cette façade se dissimule une différence importante: alors que von Mises faisait de l'hypothèse du Collectif un préalable à l'application du calcul des probabilités (il fallait, en quelque sorte, s'assurer de travailler sur une expérience qui soit vraiment "liée" au hasard), Kolmogorov qui a vraisemblablement en tête des applications à la physique ou à la théorie des nombres remarque que le lien entre sa proposition d'axiomatisation et la question du hasard n'est pas sa problématique. L'obligation de naviguer à vue entre une conception fréquentiste pure, plus facilement conciliable avec la doctrine matérialiste et l'intention d'applications beaucoup moins restrictives que celles de von Mises amènera ainsi Kolmogorov à une originale situation de "flou" intermédiaire, qui n'est probablement pas pour rien dans la curiosité de Fréchet dont témoigne une étonnante lettre (dont on remarquera la date exceptionnelle) du 3 août 1939 où Kolmogorov répond à un lettre de Fréchet qui l'interrogeait sur sa conception de la probabilité[18]

Le premier grand apport des *Grundbegriffe* a été l'énoncé clair du théorème (appelé depuis "le" Théorème de Kolmogorov) permettant de construire une probabilité sur un espace produit infini sous les conditions de compatibilité projective, et donc en particulier de donner corps aux lois des processus stochastiques qui commençaient comme on l'a vu à être le sujet central des préoccupations des probabilistes. Comme le fait remarquer [Dellacherie et Meyer, 1975], malgré un énoncé qui semble d'une généralité spectaculaire, il ne faut surestimer la portée pratique du théorème qui reste avant tout un résultat formel de non-vacuité, l'espace mesurable produit sur lequel se fait la construction étant généralement beaucoup trop pauvre pour pouvoir travailler. Mais on peut comprendre qu'il frappa cependant les esprits. C'est néanmoins le deuxième résultat qui s'avéra le plus important, à savoir une définition précise de la loi et de l'espérance conditionnelle dans le cinquième chapitre, pour laquelle Kolmogorov se fonde sur le théorème de Nikodym sur les mesures absolument continues, publié en 1930. Kolmogorov illustre immédiatement l'efficacité de la définition de la loi conditionnelle en levant le "Paradoxe de Borel" sur le choix aléatoire d'un point à la surface d'une sphère (voir [Billingsley, 1986],p.462). Dans la foulée, il définit l'espérance conditionnelle d'une variable aléatoire par rapport à une autre comme l'intégrale par rapport à la loi conditionnelle, notion dont il montre qu'elle coïncide avec l'approche plus directe qui consiste à discrétiser la variable conditionnante et à passer à la limite du pas de discrétisation.

La réception des *Grundbegriffe* créa sans doute un certain émoi dans la communauté probabiliste, bien qu'il faille se garder de croire, comme le mentionne [VonPlato, 1994], qu'elle

[18] "Vous avez raison aussi en m'attribuant l'opinion que l'axiomatique formelle doit être accompagnée par une analyse de sa signification réelle". Archives de l'Académie des Sciences, Fonds Fréchet.

fut un tremblement de terre. En ce sens, les citations que livre un peu complaisamment [Shiryaev,1989] p.885, et notamment la longue jérémiade de Lévy qui comme d'habitude explique qu'il savait tout avant tout le monde, donnent un tableau un peu trop *dramatic* de l'événement. En fait, les dates de traduction témoignent déjà d'une émergence plus longue que l'on ne croit usuellement: la traduction russe, sans doute ignorée en Occident, date de 1936 et la première édition anglaise date de 1950 seulement. [VonPlato, 1994] fait remarquer que les manuels universitaires de probabilités sont loin de s'être emparés de suite de l'axiomatisation russe. À l'exception du livre de Cramer de 1937 *Random Variables and Probability distributions*, qui d'ailleurs justifie cette axiomatisation uniquement par sa supériorité pratique sur ses alternatives (dont la théorie des Collectifs), ce n'est qu'après la Seconde Guerre Mondiale que progressivement va s'imposer le modèle de Kolmogorov.

## 3.3 Théorie de l'information

Le dernier sujet dque je voudrais mentionner ici constitua un des grands centres d'intérêt du mathématicien dans les années 50. Il s'agit de la théorie de l'information dont il s'occupa en parallèle avec ses recherches sur les systèmes dynamiques (en particulier l'importante théorie *KAM* (Kolmogorov-Arnol'd-Moser) sur les petites perturbations dont je ne parlerai pas[19]). Comme je sors légèrement avec elle du cadre temporel que j'avais fixé (Staline est mort en 1953 et c'est en 1956 que Kolmogorov publia son article commun avec Gel'fand et Yaglom *К общему определению количества информации* (Vers une définition générale de la quantité d'information),Докл.Ак.Наук СССР,111,1956,745-748), je ne vais ici que brièvement exposer la thématique en question en suivant de très près l'exposé de [Shiryaev,1989]. Je suppose cependant qu'il serait intéressant de placer ces questions d'une part dans le contexte de ces années de Guerre Froide et de conquête spatiale et d'autre part dans celui de la déstalinisation progressive du régime, car il est difficile de penser qu'il n'y ait pas eu de connexions fortes entre ces éléments. [Vucinich, 2002] pose des jalons dans ce sens et lie le développement accéléré de la Cybernétique en URSS avec les changements qui suivent le dégel progressif du pays à partir de 1953. Il est aussi révélateur que le grand exposé de Kolmogorov sur ces questions à l'Académie des Sciences de l'URSS en 1956 [Колмогоров, 1956] ait pris place dans le cadre d'une session spécialement destinée aux "problèmes scientifiques de l'automatisation de la production".

En 1948 paraît le livre de Wiener *Cybernetics or Control and Communication in the Animal and the Machine* simultanément avec les premières publications de Shannon sur la théorie de l'information ( le livre de Shannon *The mathematical theory of communication* date de 1952). Kolmogorov, cité par [Shiryaev,1989], rapporte lui même le très grand intérêt qu'il porta dès le début aux travaux de l'ingénieur de Bell auxquels il eut visiblement accès très tôt ainsi que Khinchin qui publia dès 1953 un papier [Хинчин, 1953] où il introduit la notion générale d'entropie de deux probabilités. Dans ce dernier, Khinchin s'appuie sur les recherches de Shannon (et cite en particulier son article fondateur de 1948 aux publications de la Bell Company *Mathematical theory of communication.* On notera avec intérêt que l'autre référence est un article de Fortet[20]. Ce dernier depuis quelques années s'intéressait

[19] On pourra, sur les conseils de [Shiryaev,1989], se reporter à [Abraham and Mardsen, 1978]

[20] *Processus stationnaires et de Markov et entropie.* La cybernétique, théorie du signal et de l'information.

lui aussi de près aux problèmes de transmission, en parallèle (ou en collaboration?) avec Felix Pollaczek [21] et il publia en 1948 deux Notes concernant la perte d'appels téléphoniques dans les centraux (CRAS Paris,226,1948 159-161 et 1502-1504).

Kolmogorov, avec ses collègues Gel'fand et Yaglom, va se fixer comme premier objectif de fournir une théorie mathématique de la quantité d'information. Shannon, dans le cas discret de messages aléatoires $\xi$ et $\eta$ pouvant prendre les valeurs $x_1, \ldots, x_n, \ldots$ (resp. $y_1, \ldots, y_n, \ldots$) avec les probabilités $p_1, \ldots, p_n, \ldots$, (resp. $q_1, \ldots, q_n, \ldots$) avait défini l'entropie du message $\xi$ par $H(\xi) = -\sum_i p_i \ln p_i$ et la quantité d'information contenue dans $\xi$ par rapport à $\eta$ sous la forme d'une entropie relative $I(\xi, \eta) = \sum_{i,j} p_{ij} \ln \frac{p_{ij}}{p_i q_j}$ (où $p_{ij} = P(\xi = x_i, \eta = y_j)$). Alors que Shannon met en avant l'entropie comme la notion fondamentale, Kolmogorov fait remarquer qu'il vaut mieux s'attacher à la quantité d'information pour étendre ces travaux au cas continu dans lequel les équivalents naturels de l'entropie sont infinis. De ce fait, par analogie, Gel'fand, Kolmogorov et Yaglom définissent la quantité d'information relative de deux tribus

$$I(\mathcal{U}, \mathcal{L}) = \sup_{\mathcal{U}_1 \subset \mathcal{U}, \mathcal{L}_1 \subset \mathcal{L} \text{ finies}} \sum_{A_i \in \mathcal{U}_1, B_j \in \mathcal{L}_1} P(A_i \cap B_j) \ln \frac{P(A_i \cap B_j)}{P(A_i)P(B_j)}.$$

Ils montrent alors la continuité faible de cette notion, au sens où si $\mathcal{U} = \sigma(\xi_n)$ et $\mathcal{L} = \sigma(\eta_n)$, et $(\xi_n, \eta_n) \to (\xi, \eta)$ en loi, alors $\underline{\lim} I(\xi_n, \eta_n) \geq I(\xi, \eta)$, ce qui permet d'énoncer un certain nombre de situations d'approximations où la quantité d'information de la limite va rester finie.

**Remerciements**: *Au moment de conclure ce bref tour d'horizon, je voudrais remercier très chaleureusement Marc Barbut et Bernard Bru pour m'avoir donné l'occasion d'aborder ces sujets passionnants en préparant un exposé pour le séminaire d'Histoire du Calcul des Probabilités du CAMS et du Centre Alexandre Koyré de l'EHESS.*

Revue d'optique théorique et expérimentale,1951,373-385.

[21] Felix Pollaczek fut le premier mari de la mathématicienne Hilda Geiringer qui devait épouser von Mises en 1943. La vie de celle-ci, véritable roman, est racontée sur http://www-gap.dcs.st-and.ac.uk/ history/Mathematicians/Geiringer.html.

# 4 Complément: tableau chronologique

| **Kolmogorov** *Vie Académique* | **Année** | **Russie et U.R.S.S.** |
|---|---|---|
| Naissance à Tambov | 1903 | |
| | 1905 | Mouvement révolutionnaire (Dimanche Rouge) |
| Installation à Moscou avec sa tante | 1910 | |
| | 1917 | Révolutions de Février et d'Octobre. Prise du pouvoir par les Bolchéviques. |
| | 1918 | Traité de Brest-Litovsk avec l'Allemagne. Début de l'intervention étrangère et de la guerre civile |
| Entrée à l'Université de Moscou. Fait la connaissance de Luzin | 1920 | |
| | 1921 | N.E.P. |
| Publication d'un premier travail sur les Séries de Fourier | 1922 | Fondation de l'U.R.S.S. |
| | 1924 | Mort de Lénine |
| Première publication probabiliste (avec Khinchin) | 1925 | |
| | 1928 | Exil de Trotsky. Abandon de la N.E.P. Début de la Collectivisation des campagnes. |
| Kolmogorov en Allemagne et en France. Nommé professeur à Moscou. Arrestation et mort d'Egorov. | 1930-31 | |
| Publications des *Grundbegriffe der Wahrscheinlichkeitsrechnung* | 1933 | |
| | 1934 | Congrès des Vainqueurs. Assassinat de Kirov |
| | 1935-36 | Tentative de prise en main par le pouvoir de l'Académie des Sciences. Affaire Luzin. |
| Kolmogorov rentre au comité d'édition de la Большая Совецкая Энциклопедия | 1937 | Début de la Ejovchtchina |
| Kolmogorov élu à l'Académie des Sciences | 1939 | Pacte Germano-Soviétique |
| Première phase de l'Affaire Lysenko | 1940 | |
| Kolmogorov, Prix Staline | 1941 | L'Allemagne attaque l'U.R.S.S. |
| | 1943 | Défaite allemande de Stalingrad. Contre-Offensive soviétique. |
| | 1945 | Conférence de Yalta. Fin de la 2ème Guerre Mondiale. |
| Affaire Lysenko | 1948 | Début de la Jdanovchtchina. |
| | 1953 | Mort de Staline |
| | 1956 | Rapport Khroutchev. Accélération de la déstalinisation. |
| Kolmogorov en France | 1958 | |

# References


[Abraham and Mardsen, 1978] R.Abraham and J.E.Mardsen :*Foundations of Mechanics*, 2nd Ed.,Benjamin/Cummings,1978.

[Barbut, Locker et Mazliak, 2003] M.Barbut,B.Locker et L.Mazliak: *Paul Lévy-Maurice Fréchet, 50 ans de correspondance.* A paraître chez Hermann, 2003.

[Bass, 1999] J.Bass: *A propos des modèles probabilistes de turbulence au début du XXème siècle,*MATAPLI,58,1999,49-56.

[Бернштейн, 1917] С.Н.Бернштейн:*Опыт аксиоматического обоснования теории вероятностией* (Essai de fondement axiomatique de la théorie des probabilités), Сообщ.Харьк.Мат.Об.,15,1917.

[Billingsley, 1986] P.Billingsley: *Probability and Measure,*2nd Edition, Wiley and sons, 1986.

[Bourbaki, 1999] *Bourbaki, une secte de mathématiciens*, Pour la Science, 1999

[Bru, 1998] B.Bru: *A propos des modèles probabilistes de turbulence au début du XXème siècle,*MATAPLI,55,1998,51-64.

[Cavaillès, 1940] J.Cavaillès: *Du Collectif au Pari*, Revue de Métaphysique et de Morale, 1940,139-163.

[Chapman, 1928] S.Chapman: *On the Brownian displacements and thermal diffusion of grains suspended in a non-uniform fluid*, Proc.Roy.Soc.Lond.,A,119,34-54,1928.

[Dellacherie et Meyer, 1975] C.Dellacherie et P-A.Meyer: *Probabilités et Potentiel*, I, Hermann, 1975.

[Демидов, Паршин и Половинкин, 1989] С.С.Демидов,А.Н.Паршин и С.М.Половинкин: *О переписке Н.Н.Лузина с П.А.Флоренским* (Sur la correspondance entre N.N.Luzin et P.A.Florensky), Ист-Мат.Иссл., 31,1989,116-124

[Демидов и Левшин, 1999] С.С.Демидов и Б.В.Левшин: *Дело академика Н.Н.Лузина* (L'affaire de l'académicien N.N.Luzin), РХГИ, 1999

[Fortet, 1948] R.Fortet: *Opinions modernes sur les fondements du Calcul des Probabilités* in *Les grands courants de la pensée mathématique*, F.Le Lionnais,ed.,Hermann, 1948

[Гнеденко, 1950] Е.Б.Гнеденко: *Курс теории вероятностией* (Cours de Théorie des Probabilités), Гос.Из.Тех-Теор.Лит., 1950

[Graham, 1993] Loren R.Graham: *Science in Russia and the Soviet Union*, Cambridge University Press, 1993

[Guiraldencq, 1999] P.Guiraldencq: *Emile Borel. L'espace et le temps d'une vie sur deux siècles*, Distribution Librairie Blanchard (1999)

[Maz'ja and Shaposhnikova, 1998] V.G.Maz'ja et T.Shaposhnikova: *Jacques Hadamard : A Universal Mathematician*,AMS-London Math.Society (1998)

[Heyting, 1930] A.Heyting: *Die formalen Regeln der intuitionistischen Logik* (Les règles formelles de la logique intuitionelle), Sitz.Preuss.Akad.Wiss.,1930, 42-56

[Хинчин, 1926] А.Я.Хинчин: *Идеи интуиционизма и борьба за предмет в современной математике*(Idée de l'intuitionisme et bataille pour le sujet dans la mathématique contemporaine),Вест.Комм.Акад.,16,1926,184-192

[Хинчин, 1953] А.Я.Хинчин:*Понятие энтропии в теории вероятностей* (Le sens de l'entropie en théorie des probabilités)Усп.Мат.Наук,8,1953.

[Колмогоров,1925] А.Н.Колмогоров: *О принципе tertium non datur*(Sur le principe du *tertium non datur*), Матем.Сб., 1925

[Kolmogorov, 1933] A.N.Kolmogorov: *Grundbegriffe der Warscheinlichkeitsrechnung* (Notions fondamentales du Calcul des Probabilités), 1933

[Колмогоров, 1938] А.Н.Колмогоров: *"Математика"*("Mathématique"), Большая Совиетская Энциклопедия, 1ое Изд, 1938

[А.Н.Колмогоров, 1947] А.Н.Колмогоров: *Роль русской науки в развитии теории вероятностей* (Le rôle de la science russe dans le développement de le théorie des probabilités), Уч.Зап.МГУ,91,1947

[Колмогоров, 1956] А.Н.Колмогоров:*Теория передачи информации* (Théorie de la transmission de l'information),АН СССР. Сессия по научным проблемам автоматизации производства.Пленарное заседание,1956.

[Krementsov, 1997] N.Krementsov: *Stalinist science*, UPC, 1997

[Lénine, 1970] V.I.Lénine: *Matérialisme et Empiriocriticisme*, Ed. du Progrès, Moscou; 1970

[Lévy, 1970] P.Lévy: *Quelques aspects de la pensée d'un mathématicien*, 1970

[Locker, 2001] B.Locker: *Paul Lévy: la période de guerre. Intégrales stochastiques et mouvement brownien*, Thèse d'Université Paris V, 2001

[*Математика, 1948] Математика в СССР за 30 лет* (La mathématique en URSS pendant 30 ans),Гос.Изд.Тех-Теор.Лит., 1948.

[Seneta, 2001] E.Seneta: *A.A.Markov*, in *Statisticians of the Centuries*, C.C.Heyde and E.Seneta (Editors), Springer, 2001

[Shiryaev,1989] A.N.Shiryaev: *Kolmogorov: Life and creative activities*, Ann.Prob., 1989

[Ширяев, 1993] А.Н.Ширяев (ред.): *Колмогоров в воспоминаниях* (Souvenirs de Kolmogorov), Наука,1993.

[Слуцкий, 1912] Е.Е.Слуцкий: *Теория корреляции и элементы учения о кривых распределения* (Théorie de la corrélation et éléments d'étude des courbes de répartition), Изв.Киев.Коммерч.Инст.,Киев,1912

[Вавилов,1946] С.Вавилов: *Советская наука на новой этапе* (La science soviétique à une nouvelle étape),Из-во. А.Н.СССР, 1946

[Vucinich, 1999] A.Vucinich: *Mathematics and Dialectics in the Soviet Union: The Pre-Stalin period*, Hist.Mathemat., 26,107-124, 1999

[Vucinich, 2000] A.Vucinich: *Soviet Mathematics and Dialectics in the Stalin era*, Hist.Mathemat., 27,54-76, 2000

[Vucinich, 2002] A.Vucinich: *Soviet Mathematics and Dialectics in the Post-Stalin era*, Hist.Mathemat., 29,13-39, 2002

[VonPlato, 1994] J.Von Plato: *Creating modern probability*, Cambridge, 1994

[Юшкевич, 1968] А.А.Юшкевич: *История математики в России до 1917 года. Гл.19: Теория вероятностей* (Histoire des mathématiques en Russie avant 1917. Chap.19: Théorie des Probabilités),Наука, Москва,1968

[Werth, 1995] N.Werth : *Histoire de l'Union Soviétique de Lénine à Staline*, Que sais-je 2963, 1995

[Zust, 2002] M.Zust : *L'expérience religieuse dans la théologie de Pavel Florensky*, Thèse d'Université Paris IV, 2002