L'ouvrage dont je vais vous rendre compte aujourd'hui est le deuxième que Ian Hacking a consacré au Calcul des probabilités. Son premier ouvrage, l'émergence de la probabilité, est paru en 1975, et je vous en avais parlé l'année dernière. Une question majeure était de savoir pourquoi cette émergence avait été si tardive, au XVIIème siècle?
Sa réponse, en résumant beaucoup, est que la théorie ne pouvait apparaître avant que le concept de probabilité, au sens ou nous l'entendons, ne fût lui même apparu. Selon Hacking, les conditions qui entourèrent son avènement déterminent sa structure fondamentale, la dualité qui existe entre ses deux aspects: statistique et épistémique, objectif et subjectif. D'un côté la probabilité se fonde sur la mesure de fréquences empiriques, de l'autre sur l'hypothèse a priori de chances égales (équiprobabilité).
Par une analyse inspirée par celles de Michel Foucault, il montre que la probabilité est fille des « basses sciences » , médecine, alchimie notamment ou l'opinion (la doxa) comptait bien plus que la déduction logico-mathématique (par opposition avec l'astronomie, la mécanique). Aux XVIème et XVIIème siècle, le concept d'évidence factuelle se substitue progressivement à celui de signe, encore imprégné de magie, et de croyance en l'autorité des textes antiques. L'évidence factuelle se base sur les choses elles mêmes, et non plus sur le recueil d'opinions autorisées (qu'on disait alors « probables »).
Nous avions cité quelques pages du célèbre traité de Laplace et notamment son introduction qui présente la vision mécaniste d'un monde où futur et passé se trouvent déterminés par les conditions présentes. Pas de place pour le hasard dans ce monde, sinon celui provenant de notre ignorance des véritables causes. La probabilité paraît donc devoir être entièrement subjective, mais, quelques chapitres plus loin, quand il est question de ce que nous nommerions maintenant l'inférence statistique, la détermination approchée d'une loi de probabilité à partir de données d'expérience, Laplace est bien obligé d'objectiver la probabilité, qu'il nomme alors plutôt possibilité.
Ainsi, la dualité et l'ambiguité demeurent au coeur de ce traité célèbre, comme de ceux qui lui succèderont, si l'on met à part les tentatives de philosophes, notamment de Carnap, qui sont restées sans écho dans le monde scientifique.
On peut néanmoins dire que jusqu'au XVIIIème siècle le point de vue subjectif était premier, mis en avant. Les choses changent au XIXème siècle et c'est l'objet du second livre de Hacking que je vais maintenant vous présenter. Ce deuxième ouvrage, « The Taming of Chance », (que l'on pourrait traduire par « la chance apprivoisée » si l'on veut bien voir la chance comme une sorte de mégère) est paru en 1990.
Hacking résume son propos comme suit: La physique du XXème siècle a découvert que le monde n'est pas déterministe. Auparavant, « durant le XX ème siècle, il devint possible de voir que le monde pouvait être régulier sans pour autant être soumis à des lois universelles de la Nature. Un espace fut dégagé pour le hasard ».
Dans l'age de la raison (le XVIIIème siècle), le hasard ne peut être du qu'à l'ignorance des véritables causes et « le concept cardinal de la psychologie de cette époque est la nature humaine ». Mais au XIXème siècle, « la société devient statistique ». On assiste à une « avalanche de chiffres » produits par les recensement des populations et de leurs comportements. L' « homme normal » va devenir le concept central.
Par ailleurs, les jeux de hasard, dont l'étude raisonnée marque les commencements du calcul des probabilités, se réduisent en théorie au moins au déterminisme: le lancer des dé est un problème de mécanique. Les faits statistiques bruts sont apparus dans le domaine des sciences sociales. Ceslois psychologiques, sociologiques, biologiques, sont de nature statistique et neanmoins ont un caractère inexorable. Un effet de rétroaction est à noter dans le cas des comportements humains dans la mesure où la plupart des individus cherchent à éviter d'être considérés comme déviants. Cette avalanche de chiffre et cette évolution de la psychologie sont bien illustrées, de manière ironique, par Honoré de Balzac. Hacking évoque son essai intitulé (en référence à Brillat-Savarin),

« Physiologie du mariage » . La première édition date de 1826. En voici deux citations:

« Aujourd'hui, en morale comme dans les sciences exactes, ce siècle veut des faits, des observations. Nous en apportons ».

« L'Administration s'est occupée, depuis vingt ans à chercher combien le sol de la France contient d'hectares de bois, de prés, de jachères. Elle ne s'en est pas tenue là, elle a voulu connaître le nombre et la nature des animaux. Les savants sont allés plus loin, ils ont compté les stères de bois, les kilogrammes de boeuf, les litres de vin, les pommes et les oeufs consommés à Paris.
Mais personne ne s'est encore avisé, soit au nom de l'honneur marital, soit dans l'intérêt des gens à marier, soit au profit de la morale et de la perfectibilité des institutions humaines, d'examiner le nombre de femmes honnêtes ».

Ces mots ouvrent la deuxième « méditation » de l'ouvrage intitulée « Statistique conjugale ». Balzac s'amuse ensuite, au cours de cette méditation statistique, à pallier au manque qu'il vient de révèler, évidemment à l'aide d'estimations plus ou moins fantaisistes.

Hacking décrit ce changement de point de vue advenu au XIXème siècle à travers une série d'exemples. Cette histoire, selon lui, a peu de héros. Les transformations conceptuelles, les changements de style de raisonnement ne sont pas produits par l'intervention d'individus isolés, fussent-ils géniaux.
Il se produit en même temps, selon Hacking, une lent érosion du déterminisme mécaniste dont le chantre le plus remarquable avait sans doute été Laplace.
Au XVIIIème siècle, Hume ne mettait pas en doute la nécessité:

« It is universally allowed that nothing exists without a cause of its existence, and that chance when strictly examined is a mere negative word, and means not any real power which has anywhere a being in nature ».

Il mettait cependant en doute notre capacité à la connaître. Newton était apprécié par lui en ces termes:
« While he seemed to draw off the veil from some of the mysteries of nature, he showed at the same time the imperfections of the mechanical philosophy and thereby, restored the ultimate secrets to that obscurity, in which they ever did and will ever remain ».

On note une nette différence de point de vue avec Laplace.

Le hasard, dans un monde régi par la nécessité, a été souvent conçu, depuis Aristote jusqu'à Cournot, comme l'effet de la rencontre de deux séries causales indépendantes (le pot de fleurs qui tombe sur la tête du passant). Indépendance ici signifie qu'on de peut décler de cause commune dans un environnement proche.

Voyons maintenant comment cette « avalanche de chiffres » est apparue. Au XVIIIème siècle, il y avait encore peu de données statistiques: quelques tables de mortalités servant aux calcul des rentes, quelques statistiques officielles généralement tenues secrètes.
Hacking note cependant que Kant, dans un petit essai sur l'histoire, reconnaît l'existence de tendances générales précise, à l'échelle des populations (mortalité, mariages) alors que les choix individuels sont libres et n'obéissent donc à aucune loi.
Les recensements systématiques furent d'abord effectués dans les colonies (Pérou, Virginie, Quebec, Irlande). Un recensement décennal fut prévu dans la constitution américaine.
Sur le conseil précoce de Leibnitz, un bureau chargé de recenser la population selon divers paramètres sera créé en Prusse dés le XVIIIème siècle. C'est d'ailleurs à Göttingen qu'apparaîtra le

terme « Statistik » : Collecte de faits remarquables concenant l'Etat. Plus tardivement, on le réservera aux études numériques, en laissant de côté les observations de nature plus géographique.
C'est aussi à la fin du XVIIIème siècle qu'apparait, notamment avec Malthus, ce qu'on pourrait appeler la biopolitique. Selon Malthus le problème démographique vient de ce que les populations tendent à croître géométriquement alors que la production ne croît qu'arithmétiquement. Le problème démographique est souvent posé sous la forme d'une opposition nous/les autres. En France, ce fût par contraste avec l'Angleterre et la Prusse, puis l'Allemagne. En Angleterre, on était plus sensible à l'opposition interne entre classe supérieure et classe laborieuse. En Prusse, une attention particulière fut attachée au recensement des juifs.
Au cours du XIXème siècle, des bureaux de statistiques sont créés dans toute l'Europe, au niveau des administrations centrales, des grandes villes etc... On assiste à une véritable explosion de la quantité de données disponibles.
Mais si l'existence de tableaux statistiques est une condition nécessaire à l'apparition de lois statistiques, elle n'est pas suffisante. On observe à cet égard, selon Hacking une différence notable entre l'est et l'ouest de l'Europe.
A l'est, (en Prusse, en Russie), les régularités sont vues comme l'expression d'une culture collective qui transcende les individus. A l'ouest (Angleterre, France), on y voit plutôt l'effet moyen des choix d'un grand nombre d'individus indépendants.
L'idée selon laquelle le comportement humain était soumis à des lois naturelles était déjà bien présente en France et en Angleterre durant le XVIIIème siècle. C'est un des aspects de la philosophie des Lumières illustré notamment par Condorcet dans son « Esquisse d'un tableau historique des progrés de l'esprit humain » . Cette idée de « lois morales » (c'est à dire relatives aux moeurs) n'était pas présente au même degré en Prusse.
A l'époque de Condorcet, la « mathématique sociale », c'est à dire ce que nous appellerions maintenant les statistiques, est encore débutante, mais elle existe. Quand l'avalanche de données arrivera, les lois statistiques seront naturellement décelées et les statistiques serviront de nouveau fondement au ? sciences morales ?.
Hacking présente de façon détaillée plusieurs exemples des circonstances dans lesquelles ces données ont été recueillies.
Je dois abréger mais je mentionnerai tout de même de façon quelque peu détaillée la question des suicides.
Au début du XIXème siècle, le suicide est considéré comme une maladie. On s'interroge, notamment en France, sur ses causes, sur les différences supposées entre les pays (rivalité France/Angleterre). On dissèque des suicidés.
Dés 1823, des statistiques sont faites à Paris: Classification par age, sexe, statut, méthode employée (défenestration, armes à feu, armes blanches, noyade...)Elles montrent une stabilité remarquable et des variations sinusoïdales selon les saisons (plus de suicides en été). On tente une classification selon les causes: Amour déçu, dettes, dénuement extrême... Guéry est un personnage marquant de cette époque, admiré en Angleterre, notamment pour sa pratique des statistiques comparatives.
Il utilisa pour la première fois les statistiques de rang pour inférer que l'éducation ne réduit pas nécessairement le crime. A cette fin, il évaluait le niveau d'instruction des départements à partir des questionnaires remplis par les conscrit et les comparait ensuite avec les données recueillies par le ministère de la justice. D'où il ressortait que la criminalité avait tendance à croître avec le niveau d'instruction.
Attention, il y a une cause commune envisageable: la richesse.

Guéry décrivait son activité comme de l' « analyse morale »: une science positive, seulement soucieuse d'établir des faits, sans jugement de valeurs.

Hacking considère que dans les années 1830, on passe progressivement d'un point de vue principalement subjectiviste à un point de vue principalement objectiviste sur les probabilités.
Il note que c'est à cette époque que Poisson et Cournot comprirent pleinement la logique de ce qui

deviendrait plus tard la statistique mathématique, notamment les notions de test et d'intervalle de confiance. Le point de vue de Laplace sur ces questions était plutôt bayesien.
Rappelons qu'un intervalle de confiance (en langage commun, une fourchette) est un intervalle aléatoire dans lequel la vraie valeur du paramètre a par exemple 99% de chances de se trouver, auvu de la méthode d'échantillonnage employée.
Un propagandiste marquant de la « révolution statistique » fut un astronome belge, Adolphe Quetelet. Il connaissait bien la loi des erreurs de Gauss selon laquelle les mesures d'une quantité donnée, effectuées avec le même dispositif, suivent une loi normale. On peut attribuer ce phénomène à la superposition de diverses petites erreurs indépendantes comme dans un jeu de pile ou face.
Quetelet étendit cette idée aux sciences morales. En gros, si la population considérée est homogène, il existe un « individu type » qui la représente, dont les paramètres sont les valeurs moyennes mesurées dans la population. Les observations se répartissent par rapport à lui de façon gaussienne. Les peuples, les races seront ainsi caractérisées par un individu type. (Nous pensons toujours en ces termes quand nous parlons de la durée de vie moyenne des français, des japonais etc...). Plus tard Galton superposera des photos anthropométriques de criminels violents, d'officiers etc..., à la recherche de types.
La constance des statistiques criminelles conduit à s'interroger sur la notion de responsabilité pénale. Quetelet a écrit: « C'est la société qui prépare le crime, le coupable n'est que son instrument ». Quid alors du libre arbitre, sur lequel se fonde le droit? Cette contradiction demeure irrésolue et revient (de façon lassante) à chaque procès, quand il est question d'évaluer les circonstances atténuantes.
Rappelons que le conflit ente liberté et nécessitarisme est aussi vieux que la philosophie. (Voir dans l'exposé précédent notre citation du De Fato).
Plus récemment, certains, au vu des résultats de Bell qui semblent démonter que la mécanique quantique suppose un hasard irréductible dans le cours des choses, ont estimé que cela remettait en selle le libre arbitre menacé par le déterminisme laplacien.

Mais l'existence de lois statistiques n'est pas a priori incompatible avec le libre arbitre et Quetelet lui même décrivait les actes libres comme de petites causes indépendantes dont les effets se compensent. Et l'existence de lois statistiques n'a pas été considérée par leurs premiers avocats comme impliquant une forme de fatalisme.
Voici un extrait, cité par Hacking, du discours de bienvenue au 4ème congrès international de statistique, qui fut tenu à Londres, en 1860, sous la présidence du Prince Albert:

« The statistical discoveries of one nation are the lights of all nations.
Despite the accidents of conflagrations, the unstableness of winds, the uncertainties of life and the variations in men's minds and circumstances, on which fires, wrecks and deaths depend, they are subject to laws as invariable as gravitation and fluctuate to within certain limits, which the calculus of probability can determinate beforehand.
This holds of crimes, and other acts of the will, so that violation itself is subject to law.
Shall a system of fatalism be built upon this foundation?
No, for statistics has revealed also a law of variation.
Introduce a new system of ventilation into unventilated mines, and you substitute one law of accidents for another.
These events are under control.
Some races, however, commit crimes of violence in greater proportion to other races.
Some classes are more dangerous.
But as men have the power to modify their race, they have the power to change the current of human actions within definite limits, which statistics can determine. »

Ce discours montre que l'on pense déjà à dépasser le « fatalisme statistique ». Farr, en Angleterre,

partage avec Quetelet l'idée que les gouvernements peuvent et doivent créer les conditions qui feront changer les paramètres de ces lois statistiques. Ceci conduit aussi bien à la philanthropie qu'à l'eugénisme. Notons en passant que Galton et Pearson, en Angleterre, fonderont le mouvement eugéniste, qui fut très important en Europe et en Amérique, avant la guerre (ce qui conduisit à de nombreuses stérilisations forcées).
Galton, auteur en 1869 de l'ouvrage « The hereditary Genius » était cousin de Darwin. Le mouvement proposait de sélectionner une élite génétiquement supérieure. Il créa un laboratoire d'anthropométrie judiciaire (empreintes digitales, photos d'identité) à University College qui devint le premier département de Statistique moderne. Il inventa la notion de corrélation, de droite de régression. Son élève Pearson, qui inventa le test du chi carré et créa les journaux Biométrika et Annals of Eugenics (devenu en 1954 Annals of Human Genetics) voyait dans la corrélation une révolution dans la notion physique de causalité.
Un des aspects remarquables de la pensée de Galton, selon Hacking est son attention portée à la variance, à la corrélation et aux valeurs extrêmes. Pour lui, la moyenne est médiocre. Les extrèmes peuvent être pathologiques mais aussi admirables.
Hacking souligne encore, à cet égard, la différence entre le point de vue individualiste de l'Ouest, qui croit en la main invisible et aux lois statistiques, et le point de vue de l'Est, culturaliste, pour qui ces lois ne sont pas de véritables lois (Gesetz), mais de simples régularités (Regelmässigung).
Un personnage important, à l'Est, fut Engel, chef du bureau prussien des statistiques, qui joua un rôle important dans la social-démocratie. L'état providence est né à la fin du XIXème siècle, dans l'empire allemand, en opposition au libéralisme de la main invisible. Hacking établit un lien entre ces faits historiques et cette conception différente des faits statistiques. Engel travailla beaucoup sur l'économie, le budget moyen des familles dans les différentes classes de la société. La loi de Engel, en économie, dit que la proportion de dépenses consacrées à la nourriture est en raison inverse du niveau de vie.
Cette importance croissante des statistiques ne manqua pas de susciter des réactions négatives. Hacking mentionne Auguste Comte (qui créa la sociologie) , Claude Bernard (physiologie) et Le Play (anthropologie). Bernard, comme Le Play, s'intéresse aux individus et non aux moyennes. La tache du médecin est de soigner un être humain, de déterminer les causes de ses maladies et de les traiter.
Hacking cite aussi cette boutade attribuée à Disraeli: « There are three kinds of lies: lies, damned lies, and statistics ».
Il évoque aussi Nietzche mais ses réflexions sont à la fois trop brèves et trop complexes pour que je les reproduise ici. Disons seulement qu'Hacking relève l'affirmation que hasard et nécessité sont appariés au sens que le concept de l'un ne peut exister sans celui de l'autre. Nietzche reprend aussi, dans certains textes, la vieille idée épicurienne selon laquelle existent, dans l'infinité de l'espace et du temps, une infinité de mondes et que la nécessité qui régit celui-ci est aussi le fruit du hasard. Cette conception existe apparemment aussi dans l'hindouisme. Le jeu de dés, sans qu'on sache d'ailleurs très bien en quoi il consistait, joue d'ailleurs un rôle central dans le Maharabatha, l'immense épopée qui est au coeur de la culture hindoue.

Hacking achève son ouvrage par l'évocation d'une figure assez particulière de la philosophie: Charles Sanders Peirce (1839-1914). Son influence sur la philosophie analytique anglo-saxonne a été considérable. Il n'eut pratiquement jamais de poste universitaire mais fut employé durant trente ans au « Coast Survey », un service public de géodésie qui était dirigé par son père. Son travail visait notamment à améliorer les méthodes de la géodésie. Après la mort de son père, le service fut fermé suite à un rapport du congrés qui concluait à une « overdedication to worthless science ».
Peirce, selon Hacking, est le premier philosophe à « voir le hasard couler dans toutes les avenues du sens ». Son monde est un monde de probabilités. Il fait un examen critique de la doctrine de la nécessité. L'observation ne peut établir la causalité mécanique mais seulement des régularités. Comme Epicure et Lucrèce, il pense qu'il y a de la spontanéité dans le monde et que notre libre arbitre en est un élément. L'univers évolue *et ses lois également*.

Peirce est aussi connu pour avoir distingué trois types d'inférence:
La déduction logique (mathématiques)
L'induction (les sondages)
L'abduction appelée encore inférence à la meilleure explication ou encore simplement hypothèse.
Quelle est cette troisième forme d'inférence? Débarquant en 1820 dans un port turc, il voit arriver un personnage à cheval entouré de quatre serviteurs portant un dais au dessus de lui. Il en infère que ce doit être le gouverneur.
Selon Hacking, Peirce est le premier philosophe à avoir clairement expliqué en quoi consiste la méthode statistique. Nous concevons des hypothèses et nous les testons par induction. Nous rejetons une hypothèse par une méthode qui ne tombera en défaut que dans une faible prportion des cas où elle est employée.

Reste un point à éclaircir. Ce qui m'intéresse en fait, quand je fais une inférence, c'est de savoir à quel point celle ci est fiable, et non de savoir si la méthode employée conduit le plus souvent à la vérité.
« An individual inference must be either true or false, and can show no effect of probability; and therefore, in reference to a single case considered in itself, probability can have no meaning »

La réponse donnée par Peirce est assez remarquable.
« It seems to me that we are driven inexorably to this, that logicality inexorably requires that our interests shall not be limited. They must not stop at our own fate, but must embrace the whole community. This community, again, must ***not*** *be limited, but must extend to all races of beings with whom we can come into immediate or mediate intellectual relation... Thre is nothing in the facts to forbid our having a* ***hope****, or calm and cheerful wish, that the community may last beyond any assignable date ».*

Ainsi, pour ce solitaire obstiné, le lien social est un présupposé du raisonnement.

« The very origin of the notion of reality shows that this conception essentially involves the notion of community without definite limits and capable of indefinite increase of knowledge »?
Cette position est à l'inverse de la méthode cartésienne, qui fonde la réalité sur une introspection rigoureuse. Peirce pense que cette méthode ne peut servir de base à la science et à la logique moderne. Hacking voit dans cette attitude l'influence de son expérience de la géodésie et de son appartenance à cette communauté internationale de scientifiques.

Je voudrais ajouter, en guise de conclusion, quelques mots sur le contexte de recherche plus large dans lequel se situe Hacking. Il considère son travail comme contribuant à l'histoire et à la philososphie de la connaissance. Il se réfère, dans son introduction, à Alistair Cameron Crombie (1915-1996), historien oxfordien dont l'oeuvre maîtresse s'intitule « Styles of Scientific Thinking in the European Tradition. The History of Argument and Explanation especially in the Mathematical and Biomedical Sciences and Arts ». Hacking se propose de modifier la liste établie par Crombie dont le point de vue est plus historique et moins philosophique que celui de Hacking. Le point essentiel est qu'un style de raisonnement ne peut être mis en défaut. C'est en quelque sorte un monde clos sur lui même. Peut être vous souvenez vous du texte de Coetzee que j'avais cité au début de mon premier exposé...

« Les propositions probabilistes constituent un petit monde à part. Ce qui est exprimé en termes probabilistes ne peut s'interpréter qu'en termes probabilistes. Si vous ne pensez pas déjà en termes probabilistes, les prédictions qui émanent du monde probabiliste semblent vides de sens ».