A Short History of Probability

A gambler's dispute in 1654 led to the creation of a mathematical theory of probability by two famous French mathematicians, Blaise Pascal and Pierre de Fermat. Antoine Gombaud, Chevalier de Mere, a French nobleman with an interest in gaming and gambling questions, called Pascal's attention to an apparent contradiction concerning a popular dice game. The game consisted in throwing a pair of dice 24 times; the problem was to decide whether or not to bet even money on the occurrence of at least one "double six" during the 24 throws. A seemingly well-established gambling rule led de Mere to believe that betting on a double six in 24 throws would be profitable, but his own calculations indicated just the opposite.

This problem and others posed by de Mere led to an exchange of letters between Pascal and Fermat in which the fundamental principles of probability theory were formulated for the first time. Although a few special problems on games of chance had been solved by some Italian mathematicians in the 15th and 16th centuries, no general theory was developed before this famous correspondence.

The Dutch scientist Christian Huygens, a teacher of Leibniz, learned of this correspondence and shortly thereafter (in 1657) published the first book on probability; entitled *De Ratiociniis in Ludo Aleae*, it was a treatise on problems associated with gambling. Because of the inherent appeal of games of chance, probability theory soon became popular, and the subject developed rapidly during the 18th century. The major contributors during this period were Jakob Bernoulli (1654-1705) and Abraham de Moivre (1667-1754).

In 1812 Pierre de Laplace (1749-1827) introduced a host of new ideas and mathematical techniques. Before Laplace, probability theory was solely concerned with developing a mathematical analysis of games of chance. Laplace applied probabilistic ideas to many scientific and practical problems. The theory of errors, actuarial mathematics, and statistical mechanics are examples of some of the important applications of probability theory developed in the l9th century.

Like so many other branches of mathematics, the development of probability theory has been stimulated by the variety of its applications. Conversely, each advance in the theory has enlarged the scope of its influence. Mathematical statistics is one important branch of applied probability; other applications occur in such widely different fields as genetics, psychology, economics, and engineering. Many workers have contributed to the theory since Laplace's time; among the most important are Chebyshev, Markov, von Mises, and Kolmogorov.

One of the difficulties in developing a mathematical theory of probability has been to arrive at a definition of probability that is precise enough for use in mathematics, yet comprehensive enough to be applicable to a wide range of phenomena. The search for a widely acceptable definition took nearly three centuries and was marked by much controversy. The matter was finally resolved in the 20th century by treating probability theory on an axiomatic basis. In 1933 a monograph by a Russian mathematician A. Kolmogorov outlined an axiomatic approach that forms the basis for the modern theory. Since then the ideas have been refined somewhat and probability theory is now part of a more general discipline known as measure theory.

＃ Andrei Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, 1933.

Following that work, a mathematician would answer the question of what is probability by saying: Anything that satisfies the axioms. Probability is a normalized denumerably additive measure defined over a σ-algebra of subsets of an abstract space. Something is lost with this answer, however. For if the space is finite, the answer also shrinks down to saying: *Probabilities are numbers between 0 and 1 such that two events cannot occur simultaneously, the probability of either one of them occurring is the sum of the probability of the first and the probability of the second.* The formalist approach to the question does not address the meaning of probability, but only the difference in form between classical and modern probability.

以下に確率の数学的な歴史と問題を扱う。古典時代、物理学、今世紀の確率の数学的形式化の三つの段階について確率概念の扱われ方を考察する。（内容は主に Creating Modern Probability, Jan von Plato, Cambridge University Press 1994 **に依る。）**

1 Shift from classical probability

Early probability theory was concerned with a finite number of alternative results of a trial. The rule for the computation of probabilities was very simple in principle. A composite event consists of several elementary events; its probability is the sum of the probabilities of the elementary events. To determine the probabilities of composite events, the elementary events themselves must have some probabilities. Computational schemes were based on treating the elementary events as symmetric. This results in giving each of a number *m* of elementary events the same probability 1/*m*. Symmetry in the results of a game means the very idea of *fairness* of the game. The finitary classical calculus of probability is based on the ‘classical interpretation of probability’: There is supposed to be a finite number *m* of ‘possible cases.’ They are judged ‘equipossible,’ hence equiprobable, if ‘there is no reason to think the occurrence of one of them would be more likely than that of any other.’(Laplace’s rule of insufficient reason, indifference argument)

The real world does not possess the absolute symmetries of the classical theory’s ‘equipossible cases.’ Classical probability is not sufficient for frequentist applications that show unfair results in a game, so that a conceptual change was required. This change, from about 1900 on, transformed the classical calculus of probability into a mathematically deep subject. The theory began using mathematical infinities in an essential way, specifically, in the form of infinitary events. Lebesgue had created this theory, at the instigation of Borel, at the turn of the century. The measure theoretic study of sets of real numbers and of real functions, and the asymptotic properties of sequences of natural numbers are connected through the identification of a real number with such a sequence, as in a decimal expansion. Thus the probabilistic problem of the limiting behavior of relative frequency, for example, could be formulated as a problem about the measure of a set of real numbers.

(Gyldén’s problem)

The intuitive idea of rationals being very ‘special’ cases of real numbers was made precise by Henri Poincaré. In his study of the three-body problem, he proved a ‘probability 1’ result, later called Poincaré’s recurrence theorem. Counterexamples to the theorem’s recurrent motion are not impossible, but exceptional. As a similar case, a real number’s being rational is an exceptional ‘probability 0’ property. Hence with an infinity of possible results, probability 0 does not always mean impossibility, and probability 1 not certainty.

Emile Borel’s measure theory of 1898 provided a conceptual basis for Gyldén’s and Poincaré’s intuitions. He was the first to consider explicitly and systematically the probabilities of events whose occurrence depends on an infinity of elementary events. Borel formulated his theory in a very special context, namely, in a denumerable context. It was supposed to be a branch of probability lying between the finite and the continuous, or geometric.

2 Physics

（#）非決定論と量子力学（量子力学はどのような意味で非決定的な理論であるか。）

トリチウム（三重水素）は水素の同位元素であり、その半減期は 12.32 年である。*n* 個のトリチウム原子はその半数が 12.32 年以内に崩壊し、残りの半数はそのままである。*n* 個の原子のうちのいずれが崩壊し、いずれが崩壊しないかは原理的にわからない。それでも、それぞれの原子がそれぞれ特定の時刻に崩壊しているはずであり、あわよくばそのような法則を見つけることができると考えたくなる。これが不可能な望みであることは量子力学が本質的に非決定論的な理論であることを肯定することから得られる。では、どのような意味で量子力学は非決定論的なのか。量子力学に踏み込む前に、上の放射性同位元素の崩壊がどのように扱われているかを確認しておこう。どの原子がいつ崩壊するかわからないので、非決定的な崩壊のモデルは確率を使ったものになる。非決定性を確率によって解釈する。トリチウムの原子の多数の集団を考えると、一つの原子についてその崩壊時刻を知ることはできないが、時刻 *t* にまだ崩壊していない原子の個数 *N*(*t*)を近似的に予測することはできる。これを利用して崩壊の確率モデルをつくることができる。

非決定性は統計・確率概念によって表現される。非決定性と確率は深い結びつきをもっているとはいえ、それは論理的な結びつきではない。この結びつきには経験的な確証が必要である。

1　量子力学的な非決定性と波動－粒子の二重性

量子レベルでの非決定性というとハイゼンベルグの不確定性原理(uncertainty principle)を思い浮かべる人が多い。それは「電子のような粒子の位置と運動量を同時に正確に知ることができない」とよく述べられる。この表現の

なかの「知ることができない」理由はそのような値が存在しないからである。そのような値が存在しないので、電子の未来の振舞いは予測できないだけでなく、非決定的である。この非決定性(indeterminacy)が量子力学の領域全体を支配している。では、このような非決定性はどのように見出され、何を意味しているのだろうか。それを考えるために波動－粒子の二重性を軸に量子力学の歴史を織り交ぜながら考えよう。

波動－粒子の二重性をもっとも具体的に示してくれるのは光である。光は物理的には極めて興味深い対象である。私たちの周りにあふれて存在しながら、その正体は長い間曖昧なままであった。この曖昧さは光の本性に関するニュートン以来の論争が見事に証明してくれている。ニュートンは光は粒子の束であると考えたが、ホイヘンスらは光は波であると考えた。不思議なことに光はマクロなレベルでも粒子の束であるという性質と波であるという性質を両方とももっている。波の性質の代表は干渉である。この干渉現象はヤングの二重スリットの実験で見事に示された。更にマックスウェルの電磁気学の理論は光が波であることの理論的な支柱となった。こうして今世紀初頭までの状況は光が波であるという考えがはるかに優位に立っていた。しかし、古典的な電磁気学に問題がなかったわけではない。黒体放射がプランクの量子論をつくらせることになったのは 1900 年 12 月だった。アインシュタインによる光量子仮説、コンプトン効果がこれに続く。三者に共通するのは、電磁放射における放出、伝播、吸収は量子として、つまりはエネルギーの局所的な束として行われることを示したことだった。だが、これらのいずれも光が波であることを傷つける結果ではない。光はある場合には波の性質を、別の場合には粒子の性質を示す。これが光の波動－粒子二重性(wave-particle duality for light)である。これを更に物質的な粒子にも拡大したのがド‐ブロイである。(wave-particle duality for massive particle)

$p = h/\lambda$ はド‐ブロイが見出した関係である。運動量は波長の値に反比例している。この関係を念頭において光の二重性、粒子の二重性の関係を考えてみよう。波は空間に局在しておらず、したがって、位置をもっていない。これは粒子を波と見た場合も同じで、正確な運動量をもつ粒子は決定的な位置をもっていない。波長の異なる複数の波を重ねあわせると干渉が起こる。これをうまく利用して、異なる波長の波の重ねあわせからコンパクトな波束をつくりだすことができる。これを十分に局所化することができる。ただ、この波束は多くの異なる波長の波からなっており、波束の運動量は一定の幅をもつことになってしまう。つまり、波束の運動量は非決定的となる。これは次の関係を示している。

運動量の決定性－位置の非決定性　　位置の決定性－運動量の決定性

（位置と運動量だけではなく、エネルギーと時間の間にもこのような関係が成立する。位置と運動量はその決定性に関して反比例の関係にあるのである。）

ハイゼンベルグの不確定性原理は上の決定性－非決定性の関係を正確に物理的に表現したものである。

The maximal joint precision possible for position and momentum taken together at any particular time is limited.

2 非決定論と量子力学

ド‐ブロイの波動という考えはシュレーディンガーの波動力学に引き継がれる。彼の理論によれば、電子のような量子力学的なシステムは波動関数 $\phi(q, t)$ によって完全に記述される。ここで q はそのシステムの位置を、t は時間を表している。古典力学の運動方程式と同じように、そのシステムの時間的な変化は波動方程式によって特徴づけられる。以前のシステムの状態がわかれば、その後のシステムの状態もわかるという決定論的な関係をこの波動方程式は表している。しかし、古典力学とは異なって、この後の状態は位置と運動量の両方の正確な値をもっていない。以前のシステムの位置と運動量も正確な値をもっていないのであるから、このことは決定論的な方程式の性格に違反しているわけではない。波動方程式は、もし時点 t のシステムの状態（位置と運動量）が正確に与えられれば、それを使ってそれ以後のシステムの状態が正確に計算できる、予測できることを主張しているだけである。

＃決定論の主張：ある時点 t でのシステムの完全な記述から、それ以後の時点 t' のそのシステムの完全な記述が演繹できる。

シュレーディンガーの波動方程式はこの決定論の主張を正確に満たしている。1925 年にハイゼンベルクのマトリックス力学が出ると、シュレーディンガーはその力学が自分の波動力学に論理的に等価であることを示したが、ハイゼンベルクの力学が彼の不確定性原理を含んでいたことから、不確定性原理はシュレーディンガーの力学においても主要な地位を占めるはずである。

ここで注意しておかなければならないのは、ド‐ブロイの波動が通常の物質的世界の対象であるのに対し、シュレーディンガーの波動は抽象的な数学的空間の対象である点である。では、どのようにこの抽象的な波動を物理的な実在に適用したらよいのか。1926 年にボルンが答えを見出した。波動関数の２乗が電子が与えられた領域に観測の結果として見出される確率であるというのがその答えである。

これまでわかったことをまとめると次のようになる。

Even if the Schrödinger φ-function provides a complete description of whatever quantum mechanical system we happen to be investigating, and if the Born interpretation affords the correct way to apply Schrödinger's wave equation to the physical world, then quantum mechanics implies that the world is objectively indeterministic. This result is not from our lack of knowledge, but from a genuine lack of determinacy in the world. The antecedent state of a physical system does not determine a unique later outcome; instead, the initial state may lead to any of a number of alternative later results.

（1）物質のミクロな描像が粒子と波動の両面を備えている。粒子と波動というモデルは、物理的には互いに排除の関係にあるにも拘わらずである。
（2）統計的なアンサンブルについてはその確定記述を完璧に行うが、個々の事象の記述を考えた場合、波動関数の振る舞いを決める何かが欠落していると感じられる。これが「不確定性関係」である。そこをコペンハーゲン解釈は「波動関数の収縮」で解釈するが、そこが測定、観測、認識といった対象自体でない要素で補完されている点は、理論として不完全に思える。
（3）シュレーディンガーの猫のパラドックスが示すように、測定と過程、対象系と測定器の境界、分離の曖昧さがコペンハーゲン解釈にはつきまとう。
（4）EPR 実験やベルの不等式で議論されているように、存在しているものを局所的に自立した実体と考えることができない。そう考えると矛盾がある。
（5）量子的対象のイメージが明確でない。「イメージ」として人々が要求するものは物質、本質、実体、存在、実在、実存、秩序、情報、などなどいろいろなものがある。

Unlike in quantum mechanics, the use of probabilistic reasoning in classical physics is not based on the existence of indeterminacy or objective lawlessness. In the very foundation of classical physics we find the assumption that, given the precise state of the world in one instance, the laws of physics determine its future states completely. Then how are we to interpret statements from statistical physics whose abstract form is "The probability of A is p."? Do probabilistic statements form an integral part of the description of the physical world? Do they merely reflect our ignorance with respect to the precise state of the world? Can they be deduced from non-probabilistic statements? Can they be understood in terms of frequencies?

(Maxwell's demon) Maxwell, using the demon argument, concluded that the truth of the second law is a statistical and not a mathematical truth, for it depends on the fact that the bodies we deal with consists of millions of molecules and that we never can get a hold of single molecules.

Quantum statistical mechanics contains two types of probabilities that cannot be given a unified presentation. The first type of probabilities are related to the Schrödinger wave function. Probabilities of the second type are defined exactly the same way as in the classical theory.

（古典物理と確率についての二、三の文献）

Sklar, L.: *Physics and Chance*, Cambridge U. P., 1993
Ehrenfests: *The Conceptual Foundations of the Statistical Approach in Mechanics*, Dover, 1990.
Guttmann, Y. M.: *The Concept of Probability in Statistical Physics*, Cambridge U. P., 1999.
Callender, C.: "Reducing Thermodynamics to Statistical Mechanics: The Case of Entropy", *Journal of Philosophy*, XCVI, 1999, 348-373.

The development of physics has had a profound influence on probability. This influence stems from two sources, quantum mechanics and statistical physics. Quantum mechanics destroyed the deterministic doctrine of classical physics. But the conceptual changes statistical physics brought into the modern worldview do not have the dramatic character associated with quantum mechanics. Statistical physics took its first steps in the 1850s with the formulation of the mechanical theory of heat. There was always a tension between the classical mechanics that was supposed to be valid on the level of the atomic motions, and the macroscopic behavior of matter. Specifically, while mechanical processes are reversible and symmetric in time, heat processes obviously have a preferred direction, namely, toward the equalization of temperature differences. The problem of irreversibility became the crucial one for the kinetic theory. Probabilistic arguments were invented for reconciling the two levels with each other, that is, the levels of the mechanical molecular processes and of the macroscopic observable ones.

The conceptual basis of physics in itself provided a natural habitat for some of the essential features of modern probability theory. It provided: 1 a continuous state space, and 2 continuous time. Nothing was more natural than to apply measure theory to the state space of a statistical mechanical system. This state space was a Euclidean space R^n

for which Borel and Lebesgue had created their measure theory. The second concept provided by physics took a long time to be incorporated into a concept essential for modern probability theory. Whereas continuous random quantities had been studied in probability since Newton, random events following each other in continuous succession formed an entirely new concept. On the mathematical side, the first random processes systematically developed were Markov chains, where time is discrete. The Markov property says that the probability of the next state, given the present state, is independent of previous history. The Markov property for these processes was a probabilistic rendering of the most characteristic feature of classical mechanics: Given the law of motion and the present state, the future evolution is determined. The probabilistic analogue was obtained by replacing an exact future state by a probability distribution. When the mathematical theory was developed, the physical literature on the topic first remained unknown, except for Einstein's papers on Brownian motion. The theory of *stochastic processes* was the most profound thing statistical physics gave to probability theory.

(ergodic hypothesis) (Brownian motion) (radioactivity)

When did physicists become indeterminists? Probably around or after 1925-27. After the von Neumann's book, the following consensus was established among the physicists. Quantum mechanics is an *irreducibly acausal* theory of elementary processes. His view of probability in physics has become standard. In classical physics probabilities are basically epistemic additions to the physical structure, while quantum physics has probabilities which stem from the chancy nature of the microscopic world itself. Epistemic probability is a matter of 'degree of ignorance' or of opinion. The quantum mechanical probabilities are computed out of the ψ-function so that no place seems to be left over at which the knowing subject could inject his ignorance. The standard view is that these two kinds of probabilities, the classical-epistemic and the quantum mechanical-objective, besides the frequency probability, exhaust all possibilities. (Popper's propensity interpretation of probability)

After Heisenberg's opening, quantum mechanics emerged in a few years. The general formulation of his theory was given in the form of matrix mechanics in 1925. Soon there followed a quite independent development: Schrödinger's wave mechanics in the first half of 1926. Then the concept of probability in quantum mechanics arose through Born's work in the middle of 1926, from the soil of Einstein's and Heisenberg's transition probabilities, de Broglie's matter waves, Einstein's gas theory, and Schrödinger's wave function. Heisenberg's uncertainty relation from early 1927 showed that quantum mechanics does not permit the kind of exact concept of physical state as classical mechanics, thereby certainly reinforcing the indeterminism of elementary processes Born envisaged in 1926. In 1927 the quantum mechanical building of the new physics was approaching completion in the works of Paul Dirac. The systematic mathematical formulation of the theory was being developed in Weyl's 1927 work on group theory and quantum mechanics and in the general formulation in terms of Hilbert spaces and their operators by von Neumann. The philosophy of quantum mechanics was dominated in the early years by the 'Copenhagen spirit of quantum theory.' Its central doctrine was Bohr's 1928 idea of complementarity, a philosophical generalization of the uncertainty relations.

But the impact quantum physics on probability is largely restricted to the indeterminism-probabilism of Heisenberg-Born. For in quantum theory, the set of events typically becomes discrete because of quantization, in contrast to the continuous state spaces of classical statistical mechanics. From the latter, modern probability got its continuous time random processes, while quantum mechanics gave no technical contribution of comparable magnitude. From these technical and conceptual points of view, modern probability owes more to classical statistical mechanics than to quantum mechanics. From a foundational and philosophical point of view, it seems to be the other way around.

3 The final stage 1919-1933

1 Von Mises

In 1919 von Mises published two long works on probability. The first one was a survey of the mathematics of probability. The second one was concerned with the foundations of probability. His foundational system is based on the following ideas. There is a sample space of possible results, each represented by a number. An experiment is repeated indefinitely. The resulting sequence of numbers is called a *collective* if the following two postulates are satisfied:

1. Limits of relative frequencies in a collective exist.
2. These limits remain the same in *subsequences* formed from the original sequence.

Probability is a concept that applies only to collectives. It is a *defined* notion, the limit of relative frequency. The second condition above is a postulate of randomness. The subsequent development of the theory of collectives has centered on the proper definition of randomness. If the limit of the relative frequency of 0's in the collective differs from 0 and 1, there exist subsequences containing only 0's, say. Surprisingly many people thought this trivial observation shows the impossibility of defining randomness. From von Mises' expositions it is clear that the nonexistence of subsequences in the randomness postulate is intended in a way different from the unlimited notion of existence such as one has in set theory. Several ideas were pursued towards clarifying the principles of choice of subsequences that can be considered admissible. For von Mises, the randomness postulate was an expression of his indeterminism. He formulated a program toward developing a purely probabilistic theory of statistical mechanics in 1920. He assumed that there is a finite set of discernible macroscopic states, with probabilistic laws of transition between them. The states together with their transition probabilities form what later became to be called *Markov chains*. Their characteristic property is that the probability of transition from state *i* to state *j* does not depend on previous history of the process. But it was limited to treating only discrete time processes.

2 Kolmogorov

Kolmogorov's first paper on measure theoretic probability appeared in 1929. At the time, he viewed measure theoretic probability as a way of incorporating probability theory into pure mathematics. The first result behind this view was Borel's strong law of large numbers. Kolmogorov started his probabilistic research with a joint paper with Alexander Khintchine in 1925. They gave a sufficient condition for the convergence of a sum of random variables in the probability 1 sense. The exceptions have Lebesgue measure 0.

In a long paper of 1931 on continuous time random processes, measure theoretic probability is needed for the treatment of problems of statistical physics. That is also the stated main motivation in his *Grundbegriffe*. He gave the formulation in 1933. Probability theory in this formulation is the theory of normalized σ-additive measures on abstract spaces. It would be of no particular interest were it not for the special structure given to the basic space Ω. Kolmogorov suggested that the space Ω is the set of elementary events, members of *F* are events, Ω is the certain event, and ϕ the impossible one.

3 De Finetti

De Finetti published his subjectivist program for the foundations of probability in the early 1930s. Probability was to be interpreted as a primitive concept in an account of human behavior under uncertainty, through the betting ratios one would be willing to accept for the occurrence of the event one is uncertain about. Corresponding to these ideas, de Finetti created a theory of *qualitative probability*, and showed how the usual properties of numerical probability can be derived from the notion of *coherent bets*. There are four axioms for the qualitative probability relation $E \geqq E'$, to be read as 'the event E is at least as probable as the event E'.' The axioms are:

1. $E \geqq E'$ or $E' \geqq E$ for any events E and E'.
2. $A > E > B$ for A certain, B impossible and E neither of these. (Here $A > B$ means that $A \geqq B$ and $A \neq B$.)
3. $E \geqq E'$ and $E' \geqq E''$ implies $E \geqq E''$.
4. If E_1 and E_2 are both incompatible with E, $E+E_1 \geqq E+E_2$ iff $E_1 \geqq E_2$. Specifically $E_1 = E_2$ iff $E+E_1 = E+E_2$.