From Gambling to Random Modelling

MARTA SANZ-SOLÉ

Among mathematical fields, probability stands out because of its short history and peculiar origins. Starting with the analysis of games of chance, probability developed as the mathematical theory of uncertainty, and nowadays probabilistic methods impregnate models of random phenomena. Going through the scientific career of the Japanese mathematician Kyoshi Itô (1915–2008), we analyze his theory of stochastic analysis from the perspective of a symbiotic relationship between purely curiosity driven and applied research interests.

Introduction

This aim of this article is twofold, to introduce non specialists to some aspects of stochastic analysis — a main achievement in probability theory of the 20th century — and to give tribute to its founder, Kyoshi Itô (1915–2008). First we set Itô's work in a historical context, then we describe some of his groundbreaking ideas and early contributions in the 40's and 50's; finally, we present some recent developments and highlight the importance of his research in pure mathematics and in stochastic modelling.

In comparison with other mathematical fields, like geometry or arithmetic, probability theory has a very short history. Its origins are in the late Renaissance period, when gambling was a common leisure activity in circles of nobility and learned people, especially in Italy and France. Speculative discussions and conjectures on games of chance attracted the interest of scientists like Gerolamo Cardano (1501–1576), Galileo Galilei (1564–1642), Pierre de Fermat (1601–1655) and Blaise Pascal (1623–1662), who tried to give mathematical answers to challenging questions posed by astute and experienced gamblers.

Until the beginning of the 19th century, probability developed very slowly, possibly because of the absence of suitable mathematical tools (in particular, of combinatorial algebra), moral or religious barriers in the society to the development of the idea of randomness and chance, and the weight given to the connections of probability with philosophy. The publication in the year 1812 of the treaty *Théorie analytique des probabilités* by Pierre-Simon Laplace was a turning point in the early consolidation of the field. For the first time, there was a successful attempt to build a theory of randomness and to give some unity to the subject. However, the independence and blooming of probability only happened at the beginning of the 20th century, thanks to giant figures like Andrei A. Markov, Andrei N. Kolmogorov and Paul Lévy, with stochastic processes in their centre of interest. Around these dates, when he was a student at the University of Tokyo (1935–1938), Itô met stochastic processes. In his own words [4]:

> [...] I was fascinated by the rigorous arguments and the beautiful structures seen in pure mathematics, but also I was concerned with the fact that many mathematical concepts had their origins in mechanics. Fiddling around with mathematics and mechanics, I came close to stochastic processes through statistical mechanics.

This was the beginning of a long and fruitful journey. With his deep mathematical contributions and insight, Itô laid the foundations of stochastic models used now in many fields, like statistical physics, population genetics and mathematical finance. His influence extends to areas far beyond his imagination, as he declared when he was awarded the Gauss Prize in 2006.

Stochastic processes and Markov processes

A stochastic process is a measurable mapping $X : \Omega \times I \longrightarrow \mathbb{R}^d$, where Ω is the set consisting of the random arguments, called the sample space, and I is the set of indices, usually a subset of $\mathbb{R}^k$ or $\mathbb{Z}^k$. By fixing $\omega \in \Omega$, the deterministic mapping $X(\omega) : I \longrightarrow \mathbb{R}^d$ defined by $X(\omega)(t) = X(\omega, t)$, corresponds to an observation of the random evolution described by the process X. The deterministic functions $X(\omega)$ are called the sample paths or trajectories of the stochastic process.

Markov processes are prominent examples of stochastic processes. They are characterized by the lack of memory. If $I = [0, \infty)$, this means that, for any $t \in [0, \infty)$, the future of the process, $X_s, s > t$, is conditionally independent of the past, $X_s, s \leq t$, knowing X_t. In other words, the past information gathered by the process is summarized at the present time. Introduced and studied by A.A. Markov in 1905, memoryless random dependence already appears in the work by F. Galton and H.W. Watson (1874) on evolution of populations.

Diffusions and Kolmogorov's equation

Diffusion processes are a fundamental class of Markov processes. They provide models for particles moving randomly in a fluid, like the Brownian motion. Kolmogorov (1931) described diffusions by specifying the behavior of the conditional average of increments of the process over an interval of length h, and of covariances. More specifically, a d-dimensional diffusion $X = \{(X_t^1, \ldots, X_t^d), t \geq 0\}$ satisfies,

$$\begin{aligned} &\mathbb{E}\left(X_{t+h} - X_t | X_s, 0 \leq s \leq t\right) = b(X_t)h + o(h), \\ &\mathbb{E}\left[(X_{t+h} - X_t - b(X_t)h)\,(X_{t+h} - X_t - b(X_t)h)^{\top}\right] \\ &\qquad = a(X_t)h + o(h). \end{aligned} \tag{1}$$

The functions b and a are termed the drift and the diffusion coefficients, respectively.

For fixed $t \geq 0$ and $x \in \mathbb{R}^d$, define $P(t, x, A) := \mathbb{P}\left(X_{t+r} \in A | X_r = x\right)$, that is, the probability that, if at time r, the process is at point x, after a further t units of time, it visits the set $A \subset \mathbb{R}^d$. Assume that this probability has a density, meaning the existence of a nonnegative function, $p_t(x, \cdot)$, called the transition probability density, such that $P(t, x, A) = \int_A p_t(x, y)\ dy$. Kolmogorov proved that transition probability densities of diffusions are solutions to partial differential equations defined by the partial differential operator $\mathscr{L} = \frac{1}{2} a^{i,j}(x)\partial^2_{i,j} + b^i(x)\partial_i$, and by its dual $\mathscr{L}^*$. These are the backward and forward Kolmogorov equations,

$$\begin{aligned} &\frac{\partial}{\partial t} p_t(x, y) = \mathscr{L} p_t(x, y),\ \forall y \in \mathbb{R}^d, \\ &\frac{\partial}{\partial t} p_t(x, y) = \mathscr{L}^* p_t(x, y),\ \forall x \in \mathbb{R}^d, \end{aligned} \tag{2}$$

respectively.

Existence of Markov processes

In the 1930's, the existence of Markov processes was a central question of study. Using tools of measure theory, Kolmogorov (1931) proved the existence of Markov processes with continuous sample paths. Later on, Feller (1936) proved the existence of Markov processes with jumps.

Itô, who went deeply through the methods of these works, envisioned a differential approach to the problem focussing on the sample paths of the processes, rather than on their probability laws. For this, he viewed the Kolmogorov transition probabilities $P(t, x, \cdot)$ (defined above) as a flow of probability measures; then he introduced a suitable notion of tangent and obtained $P(t, x, \cdot)$ by integration of the tangent. Finally, using this representation, he found a realization of the integral flow $\{P(t, x, \cdot), (t, x) \in [0, \infty) \times \mathbb{R}^d\}$ on the path space of the Markov process, (called nowadays the Itô map). The tangent to the flow should be the best linear approximation. In the probabilistic context, and because of results of Paul Lévy, the role of lines is played by processes of independent increments. In this way Itô (1942) deduced that the sample paths of a diffusion (as defined in (1)) leave the initial state x in the same way as the paths of a Brownian motion B having instantaneous mean and variance $b(x)$ and $a(x)$, respectively. Hence, the probabilistic dynamics of the sample paths of a diffusion is given by the stochastic differential equation

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt,\ \ t > 0,\ \ X_0 = x, \tag{3}$$

where $\sigma(x) = [a(x)]^{1/2}$.

In Itô's words [4]: "It took some years to carry out this idea".

The birth of stochastic calculus

Having achieved a pathwise description of the dynamics of Markov processes, Itô faced two tasks, namely to give a rigorous meaning to (3), and to recover the transition probabilities of the Markov process solution to (3), thereby establishing the connection with Kolmogorov's theory. Itô's stochastic calculus with respect to the Brownian motion was created to solve both questions. With the stochastic integral (1944), equation (3) is rigorously formulated, and using the change of variables formula, called Itô's formula (1951), it is proved that the transition probabilities of $X(t)$ satisfy Kolmogorov's partial differential equations (2).

Let f be a $\mathscr{C}^2$ function and B a Brownian motion. The simplest version of Itô's formula says

$$f(B_t) = f(0)+\int_0^t f'(B_s)dB_s+\frac{1}{2}\int_0^t f''(B_s)ds. \quad (4)$$

From this equality, we see that stochastic calculus does not follow the same computation rules as ordinary calculus. Moreover, the Itô map is not continuous with respect to the Brownian motion and therefore equation (3) lacks a suitable stability property. These facts produced perplexity, especially in applied circles of scientists, motivating further investigations (Rubin and Fisk, 1955; Stratonovich, 1966). Such questions were at the origin of *rough path analysis*, a theory initiated by Terry Lyons in the late 1990's.

Itô's equation and modelling

The connections of equation

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \; t > 0, \; X_0 = x, \quad (5)$$

with modelling go back to the notion of Brownian motion, since for $\sigma = 1$, $b = 0$ and $x = 0$, equation (3) defines the Brownian motion. Let us mention some examples. In 1900, Louis Bachelier proposed a model of Brownian motion while deriving the dynamic behaviour of the Paris stock market. Fluctuation of prices in the stock market is due to the selling and buying activity of many agents, producing a similar effect as the movement of pollen particles in the experimental description of Brownian motion. In 1908, Paul Langevin described the velocity v of a Brownian particle of mass m by the equation

$$m\; dv(t) = -\lambda v(t)dt + B(dt), \; t > 0,$$

where B is a Brownian motion independent of the particle. Building on Bachelier's work, McKean and Samuelson (1965) showed that

$$dX_t = \mu X_t dt + \sigma X_t dB_t \; t > 0.$$

is a good model for stock price variations. The process (X_t) is called a *geometric Brownian motion*.

Itô's influence and parallel developments

Generalizations of the Itô stochastic integral (1944) to more general processes than Brownian motion came already in the early 50's. Motivated mainly by the study of the striking connections between stochastic calculus and potential theory, the extensions concern processes with conditionally orthogonal increments (J.L. Doob, 1953), submartingales (P.-A. Meyer, 1962), local martingales (K. Itô and S. Watanabe, 1965), and semimartingales (C. Doléans-Dade and P.-A. Meyer, 1975). In parallel, with very little interaction with other countries, the Russian school made impressive advances in stochastic analysis. Dynkin, Gikhman, Skorohod, Wentzell, Freidlin, Krylov, are in the group of main contributors. Gradually, with the break down of isolation, the groundbreaking work of colleagues in the former Soviet Union became visible and influential.

From finite to infinite dimensions: stochastic partial differential equations

It took several decades to lift Itô calculus to infinite dimensions, in particular, to develop a theory of *stochastic partial differential equations* (SPDEs). At present, SPDEs form a large and blooming field of stochastic analysis, where Itô also made some contributions in relation to measure-valued processes. In this article, we will focus on a class of SPDEs closely related to Itô's equation.

The equation (5) can be understood as an ordinary differential equation with a random forcing nonlinear term $\sigma(X_t)dB(t)$. Some classes of SPDEs obey a similar principle. Consider the heat equation on $\mathbb{R}$,

$$\frac{\partial}{\partial t}v(t,x) - \frac{\partial^2}{\partial x^2}v(t,x) = f(t,x), \; t > 0, \; x \in \mathbb{R}, \quad (6)$$

with a given initial condition, describing the evolution of the temperature along a metal bar under some external influence f. If f contains a stochastic source then (6) gives rise to a stochastic heat equation. For example, f may be a space-time white noise, $\dot{W}(t,x)$ — an infinite dimensional version of a Brownian motion. The linear stochastic heat equation driven by a space-time white noise is one of the most basic examples of an SPDE.

There are good motivations to study SPDEs, as the following two examples illustrate.

The *parabolic Anderson model* is a Cauchy problem for the heat equation with random potential. It has connections with motions in random potentials, trapping of random paths and spectra of random operators, among others. An interesting example of potentials are Gaussian processes. If the choice is a space-time white noise, as before, we have the *stochastic parabolic Anderson model* described by the SPDE

$$\frac{\partial}{\partial t}u(t,x) - \frac{\partial^2}{\partial x^2}u(t,x) = u(t,x)\dot{W}(t,x),\ t>0,\ x\in\mathbb{R}. \tag{7}$$

This equation is related with the Kardar–Parisi–Zhang equation (KPZ, in short) that Martin Hairer successfully solved in 2013,

$$\frac{\partial}{\partial t}h(t,x) - \frac{\partial^2}{\partial x^2}h(t,x) - \left[\frac{\partial}{\partial x}h(t,x)\right]^2 = \dot{W}(t,x). \tag{8}$$

Indeed, the stochastic process obtained via the Cole–Hopf transformation

$$u(t,x) = \exp\left(h(t,x)\right),$$

solves (formally) equation (7).

The KPZ equation is a new universality class (similar as the Gaussian law of the Brownian motion in much simpler situations), to describe phenomena like one-dimensional interface growth processes, interacting particle systems and polymers in random environments, which display characteristic scalings and new statistics or limiting behaviors.

Very fundamental questions on SPDEs, like giving a rigorous meaning to the equations, and proving existence and uniqueness of solution, can be very challenging and generate a great deal of wonderful mathematics. A paradigmatic example is the theory of regularity structures of Martin Hairer (2014), motivated by the well-posedness of the KPZ equation.

A question concerning sample paths of SPDEs

Numerical simulations of SPDEs show that their sample paths are complex and intriguing mathematical objects. We next describe a specific problem where the nature of the sample paths plays a very important role.

Let $v = \{v(x), x\in\mathbb{R}^m\}$ be a $\mathbb{R}^d$-valued stochastic process, solution to a system of SPDEs. How likely is it that the sample paths of v visit a deterministic set A? This fundamental question in probabilistic potential theory is clearly related to the regularity of the sample paths and geometric-measure properties of A.

Having upper and lower bounds for the *hitting probabilities*

$$\mathbb{P}\{\omega : v(\omega)(I)\cap A \neq \emptyset\} \tag{9}$$

in terms of notions of geometric measure theory, like the *capacity* or the *Hausdorff measure* of the set A, provides an insight into the problem.

A first result in this direction, proved by S. Kakutani in 1944, states that, up to positive constants, for a d-dimensional Brownian motion, the hitting probability (9) is bounded from above and from below by $\mathrm{Cap}_{d-2}(A)$, the capacity of dimension $d-2$ of the set A. In particular, this implies that a d-dimensional Brownian motion hits points if and only if $d=1$.

Hitting probabilities for the sample paths of solutions to SPDEs have been in the focus of research in the last fifteen years. Since in most cases, the random field solutions to SPDEs fail to have a "suitable" Markov property, Kakutani's method, and further generalizations to different types of Markov processes, cannot be applied. The new successful approach for SPDEs relies on the study of densities of the random field solutions at fixed points, using as mathematical background Malliavin calculus.

Closing the circle: Kolmogorov, Itô, Hörmander, Malliavin

We finish this article with some touches on Malliavin calculus, not only because of its role in the study of hitting probabilities for SPDEs but especially, to close a beautiful circle of ideas that started with Kolmogorov, continued with Itô and Hörmander, and ended with Malliavin.

Recall the forward Kolmogorov equation for densities (in the distribution sense) of diffusions starting from x,

$$\left(\frac{\partial}{\partial t} - \mathcal{L}^*\right)p_t(x,\cdot) = 0. \tag{10}$$

In the theory of partial differential operators, there is the notion of *hypoellipticity*. It tell us that, if $\frac{\partial}{\partial t} - \mathcal{L}^*$ is *hypoelliptic* then the solution to the PDE

$$\left(\frac{\partial}{\partial t} - \mathcal{L}^*\right)\alpha = 0,$$

is a smooth function. That is, the solution to (10) in the distribution sense is a smooth function $p_t(x,y)$.

In a seminal paper, L. Hörmander (1967) gave sufficient conditions of geometric type for a partial differential operator in quadratic form to be hypoelliptic (see also Kohn (1973), Oleinik and Radkevič (1973)). Hörmander's theorem applies to the operator $\mathscr{L}$, therefore to Kolmogorov's equation.

Malliavin envisioned giving a probabilistic proof of Hörmander's theorem, thereby having an exclusively probabilistic understanding of (10). Starting with Itô's stochastic differential equation (5) he proved that, by expressing Hörmander's assumptions in terms of geometric properties of the coefficients σ and b, the law of the random vector X_t, for any $t > 0$, possesses a smooth density. Existence of densities are obtained by integration by parts formulas on the Wiener space, that rely on a stochastic calculus of variations given the name of *Malliavin calculus*.

Malliavin theory, introduced in [3], was further investigated and expanded by Bismut, Stroock, Ikeda, Watanabe, Bouleau, Hirsch, Meyer, Kusuoka, Shigekawa, Nualart, Bell, Mohammed, Ocone, Zakai, among others. Its scope goes far beyond the initial project of giving a probabilistic proof of Hörmander's theorem. In particular, integration by parts formulas provide explicit expressions for densities of random vectors defined on abstract Wiener spaces. Explicit formulas for densities, along with the Malliavin calculus toolbox, are the basic ingredients to address the problem of hitting probabilities for systems of SPDEs with multiplicative noises (see work of R. Dalang, D. Khoshnevisan, C. Mueller, E. Nualart, M. Sanz-Solé, R. Tribe, C. Tudor, F. Viens, Y. Xiao, L. Zambotti).

Final remarks

By changing the approach to Markov processes of Kolmogorov and Feller, and by giving the sample paths a priority, Itô paved the way to stochastic modelling. His mathematical work consolidated the theory of stochastic processes, that was still in its infancy when he was a student at the University of Tokyo, expanded the theory, by the creation of stochastic calculus, boosted connections with other mathematical fields, initiated the use of stochastic models in sciences, and is still a source of new and challenging problems in probability.

Acknowledgements

This article is based on the London Mathematical Society Lecture delivered by the author at the ICM in Rio de Janeiro on August 7th, 2018. The author would like to express her deep gratitude to the LMS for the invitation.

FURTHER READING

[1] H. Föllmer, On Kyoshi Itô's work and its impact, in: Proceedings of the ICM 2006. EMS Publishing House, pp. 109–123.
[2] L.E. Maistrov, Probability Theory, a historical Sketch, Academic Press, 1974.
[3] P. Malliavin, Stochastic calculus of variations and hypoelliptic operators, in: Proc. Int. Symp. on Stochastic Differential Equations. Research Institute of Mathematical Sciences, Kyoto Univ., pp. 758–784, Wiley, 1978.
[4] Kiyosi Itô Selected Papers. D.W. Stroock and S.R.S. Varadhan (Eds.), Springer Verlag, 1987.
[5] D. Williams. To begin at the beginning . . ., Stochastic Integrals, in: Proc. LMS Durham Symposium 1980, Lecture Notes in Mathematics 851, pp. 1–55.

Marta Sanz-Solé

Marta Sanz-Solé is Professor of mathematics at the University of Barcelona. Her research interests are in stochastic analysis, in particular stochastic differential and partial differential equations. Sanz-Solé was awarded with a Narcis Monturiol Medal of Scientific and Technological Excellence by the Generalitat of Catalonia in 1998. She was elected Fellow of the Institute of Mathematical Statistics in 2009, and was President of the European Mathematical Society from 2011–2014.