The Role of Probability in Scientific and Mathematical Research: A Workshop on Current and Emerging Research Opportunities

A Brief Overview

Don Dawson

Theme: Identify the intrinsic nature of probability theory which makes it a rich area of stimulating mathematical inquiry in its own right - as opposed to being simply a subfield of analysis or a necessary tool for statistics or other applications.

Outline

- The place of probability in mathematics and science
- The nature of probability theory
- Major Landmarks
  - Pre 1960's
  - 1960-1980
- Rich mathematical objects
- Elegant but powerful ideas and methods
- Impact on other areas of mathematics and science
- Some views on the widening impact of probabilistic ideas
- The intellectual spectrum today
- Current and Emerging Research Opportunities

The Place of Probability in Mathematics and Science

Probability theory has a long history with major landmarks in the works of Bernoulli, de Moivre, Laplace and Gauss but it was in the 20th century that it blossomed into one of the core areas of mathematics. This development was fed from two sides.

- On the one hand the development of Lebesgue integration and measure theory was an essential ingredient for the development of the subject as we know it today.
- On the other hand the transformation from the deterministic world view to a probabilistic viewpoint in physics - Brownian motion, quantum physics, the role of probability in evolutionary biology and population genetics and the probabilistic modelling of social phenomena including Bachelier's now celebrated introduction of probabilistic models of financial markets served as a powerful catalyst for this development.

The ***central position of probability in science today*** results from

Probability links the mathematical world of theoretical models and the empirical world of data collection.

Probability also plays a key role in developing a rigorous foundation for the new mode of scientific investigation using computer simulation.

The development of probability theory owes much to this external stimulus but just as in the earlier developments in number theory, geometry and analysis probability now has its own distinct structure and methodology.

Probability also plays a rather special and increasingly central role in mathematics in that it forms a natural link between:

the continuum world of analysis and classical physics and the discrete world of combinatorics, computer science and bioinformatics.

With the advent and rapid development of theoretical computer science the previously neglected realm of discrete mathematics is now in ascendancy. This inevitably will futher enhance the position of probability at the center of the mathematical universe.

The nature of probability

- probabilistic models of natural phenomena
- probabilistic intuition and ideas
- rich mathematical objects
- powerful mathematical techniques
- interactions with other areas of mathematics and science

Some Major Landmarks

1960 Perspective

I will begin with a look at the state of the subject around 1960 when I was a graduate student. For a graduate student at that time there were a number of classical developments

- Kolmogorov's measure-theoretic foundations
- Limit Theorems for sums of IID r.v.
- The Lévy-Khinchine representation of inf. div. r.v.
- Construction of Wiener measure
- Markov chains and simple random walks
- Feller's semigroup approach to Markov processes

To give one perspective on the conceptual advance these represent I would like to quote from one of the great mathematical expositors.

Mark Kac From "**Enigmas of Chance**"

Independence is the central concept of probability theory and few would believe today that understanding what it meant was ever a problem. To most mathematicians, once a concept is defined, it is automatically understood. But such a super-formal view was distasteful to Steinhaus and he felt as I do, to an even greater extent, that accepting it is a sort of cop-out. Our work began at a time (1935-1938) when probability theory was emerging from a century of neglect and was slowly gaining acceptance as a respectable branch of pure mathematics. This turnabout came as a result of a book by the great Soviet mathematician A.N. Kolmogorov on foundations of probability theory, published in 1933. It appeared to us awesomely abstract. Even the much more accessible work of Stan Ulam and Zbigniew Lomnicki, who were the first to connect independence with so-called product measures, was not quite what we were searching for.

Kac then describes his work with Steinhaus on the "independence of cosines" and the fact that the proportion of time which the graph of the function

$$x_n(t) = \frac{\cos\lambda_1 t + \cos\lambda_2 t + \ldots + \cos\lambda_n t}{\sqrt{n}}$$

stays in the horizontal strip between a and b approaches as $n \to \infty$ the area under the normal curve. Here the $\{\lambda_i\}$ are linearly independent over the field of rationals. He then remarks that

“Even before the climax of our search for the meaning of independence was reached, it became abundantly clear why tout le monde was justified in believing in the loi des erreurs. It proved to be both un fait d’observation and une théorème de mathématiques.

In 1960 some key topics that were relatively new and exciting for a graduate student were

- Martingale theorem's - Doob's book
- Feller's classification of 1-d diffusions
- Lévy's work on Brownian motion - in particular local time
- Ito's theory of stochastic integrals and SDE
- Dynkin's general theory of Markov processes
- Ideas of probabilistic potential theory - Hunt's papers
- Shannon's information theory
- Kolmogrov-Sinai work on entropy in ergodic theory
- Ideas of weak convergence of probability measure - Prohorov and Skorokhod

Each of these introduced fundamental concepts and tools that led to enormously fruitful development and are now fully developed subfields of probability This spanned notions that are simple in retrospect but very powerful such as stopping times to the first indications of the beauty and depth of Brownian motion.

Some Landmarks:

1960's

- Theory of semimartingales and stochastic integration - the Strasbourg school (Dellacherie-Meyer)
- General theory of Markov processes and Hunt processes
- Diffusion Processes: Ito-McKean book
- Erdos-Renyi theory of random graphs
- Gibbs random fields - the DLR condition

1970's

- Regularity of Gaussian random functions
- Strong approximation - the Hungarian construction
- Probability in Banach space
- Subadditive processes
- Martingale inequalities - Burkholder, Davis Gundy
- Tightness criteria for measures on path space
- Random walks on groups
- Stroock-Varadhan theory
- Interacting particle systems
- Mathematical population genetics - the coalescent and Ewen's sampling formula
- Donsker-Varadhan theory of large deviations
- Stochastic control
- Nonlinear filtering theory

Rich Mathematical Objects

1. Brownian Motion

Beginning with the amazing insights of Paul Lévy, the investigation of Brownian motion has uncovered a mathematical object of great beauty and depth which continues to amaze. For example a recent major advance in the understanding of scaling limits of some basic probabilistic objects in two dimensions is related to the conformal invariance of the hull of planar Brownian motion stopped at the first hitting of the unit circle. An excellent exposition is found in the two volumes of Mark Yor's "Some aspects of Brownian motion". He remarks that "while preparing these notes, I was extremely astonished at the number of very natural questions which have escaped attention until very recently". One of the questions he raises is "To understand better the ubiquity of Brownian motion in a great number of probabilistic problems".

The Wiener sausage comes from Brownian motion but is an important mathematical object in its own right. For example the analysis of moderate and large deviations for the volume of the Wiener sausage plays an important role in the study of heat conduction and trapping in a random media and in the analysis of the spectrum of random Schrodinger operators.

2. Random Graphs.

Random graphs are random combinatorial objects that have a rich structure. The theory of random graphs was founded by Erdös and Rényi (1959-1961). They demonstrated that it is possible to show the existence of graphs having prescribed properties by showing that the probability that a random graph has such a property is positive. In his introduction to his book on random graphs, B. Bollobás remarks "The greatest discovery of Erdös and Rényi was that many important properties of graphs appear quite suddenly. If we (consider the sets of graphs with n labelled vertices and M edges and) pick a function $M = M(n)$ then in many cases either almost every graph in the nth set has property Q or else almost every graph fails to have property Q (as $n \to \infty$). In this vague sense we have a 0-1 law. The transition from a property being very unlikely to it being very likely is usually very swift."

3. The Ising Model and Glauber dynamics.

These serve as the canonical model of phase transition and critical behavior in statistical physics. In recent years exciting progress has been made in understanding in the context of the two dimensional Ising model the phenomena of metastability, surface tension and droplet formation (e.g. the work of Dobrushin, Kotecky, Schonman and Shlosman).

4. Super Brownian motion.

Superprocesses arose as a stochastic model for near critical populations. It became a powerful tool in pde's–witness Mselati's solution of the Dynkin program in classifying all solutions of $\Delta u = u^2$ on a region in R^d. The recent limit theorems which produce SBM as rescaled limits of particle systems (voter model, contact process, Lotka-Volterra systems) and of many of the fundamental models in statistical physics (lattice trees go to ISE, oriented percolation rescales to SBM, and percolation in high dimensions also approaches ISE–the latter still not complete) were never expected. The beautiful connections with excursion theory discovered by Le Gall and connections with random trees (Aldous and Le Gall), and current connection between spectrally positive Lévy processes and general continuous state branching mechanisms (Le Gall, Le Jan) producing long sought after extensions of the Ray-Knight theorem are examples of the rich an beautiful mathematical structures which are interacting within probability thanks again to a stochastic model for near critical growth and evolution. (Ref. E.A. Perkins).

5. Poisson-Dirichlet measures.

These have appeared in apparently unrelated areas such as group representations, prime factorizations, population genetics and Bayesian statistics. They now form part of a rich theory related to Fleming-Viot processes, the infinitely many alleles model and

Gamma measures. The related GEM random probability is invariant under size-biased sampling and the extension of this notion to a two parameter family by Pitman and Yor demonstrate that it is natural and important probabilistic object.

Examples of Simple but powerful Probabilistic Ideas and Methods

1. Coupling:

   The simple idea of coupling two processes on a common probability space has turned out to be a fertile idea with countless application in probability including:

   - percolation theory
   - exact simulation
   - infinite particle systems
   - Poisson approximation.

2. The Probabilistic method.

   Existence results based on probabilistic ideas can now be found in many branches of mathematics, especially in analysis, the geometry of Banach spaces, number theory, graph theory, combinatorics and computer science.

3. Concentration of measure.

   This is based on the intuitive idea that a function that smoothly depends on the influence of many independent variables satisfies Chernoff-type bounds. (Talagrand AP paper) Results based on this idea have been used in bin packing problems, traveling salesman problem, percolation theory, spin glass, random trees, etc.

4. **Martingale Problems.**

A large class of processes can be constructed via strong solutions of stochastic equations driven by Wiener or Poisson processes. However in singular situations and in infinite dimensional systems such as measure-valued processes this method is frequently not applicable. The idea of Stroock and Varadhan to reformulate the search for constructions of such processes as a martingale problem whose solution is the probability law of the desired object has made possible the development of entirely new classes of stochastic processes.

5. **Scaling, Self-similarity and Renormalization Methods.**

The renormalization group method at a heuristic level was introduced by the physicist K. Wilson around 1970. Since then a circle of ideas including self similarity, renormalization and universality has emerged to become an important conceptual tool and a rigorous theory is emerging. In particular rigorous results in this direction have recently been obtained using lace expansion methods in high dimensions for percolation, oriented percolation, branched polymers, self-avoiding walk, etc. These ideas also play a role in dynamical systems, random combinatorial structures and random matrix theory.

Impact on Other Areas of Mathematics & Science

- Analysis
  PDE, harmonic analysis, functional integration.
- Number Theory
  - Probabilistic number theory
  - Riemann zeta function:
    - relations with Brownian motion
    - Gaussian Unitary Ensemble: Pair correlation function for large random Hermitian matrices- relation with spacing of zeros of Riemann zeta function
- Banach space theory.
  Probabilistic ideas underlie some of the dramatic advances in the geometry of Banach space.
- Geometry.
  M. Gromov has developed a research program linking geometry and probability involving concentration ideas going back to the work of Paul Lévy.
- Incompressible Navier-Stokes equations.

Probabilistic proofs of the existence of weak solutions have been obtained that eliminated the smoothness assumptions on initial conditions previously required.

- Dynamical systems.

  The introduction of probabilistic ideas into dynamical systems theory has had a major influence including the construction of SRB(Sinai, Ruelle Bowen) measures and ideas of stochastic stability.

- Quantum and statistical physics.

  A whole range of probabilistic tools and ideas play a role here including function space integration, Feynman Kac formula, random fields, Gibbs distributions,etc.

- Computer Science.

  Probabilistic methods now play a central role in the analysis of algorithms and also in studying the asumptotic properties of random search trees and other combinatorial objects

Some Views on Nature and Widening Impact of Probabilistic Ideas

1. B. Bollobas.

   Probabilistic methods have become an important part of the arsenal of a great many mathematicians. Nevertheless, this is only a beginning: in the next decade or two probabilistic methods are likely to become even more prominent. It is also likely that in the not too distant future it will be possible to carry out statistical analyses of more complicated systems.

2. M. Csörg .

The nature of applications of, and doing applied research, in probability are also fundamentally different from doing the same kind of things in analysis. Take, for example, the interplay of theory and applications of probability, statistics, computer science, genetics, economics, etc.; an all together distinct activity as compared to doing research in applied analysis, proper.

3. Persi Diaconis.

Probabilistically, new ideas like coupling and stopping time techniques give 'pure thought' solutions to previously intractable problems. In theoretical computer science, a slew if intractable (#-p complete) problems like computing the permanent of a matrix or the volume of a convex polyhedron) have provably accurate approximations in polynomial time because simple Markov chains can be constructed and proved to converge rapidly.

4. Ed Perkins.

Probability finds itself at the nexus of biology, statistical physics, pde, analysis, statistics, quantum field theory, geometric measure theory,.... It can be motivated by problems in one area which modify techniques from another and then find application in a third. The mathematical objects created in probability often find application far beyond the original intent–this is true in all areas of good mathematics but probability's proximity to so many pure and applied areas make it especially prone to this.

5. M. Talagrand. "To explain (at least qualitatively the unconventional magnetic behavior of certain materials, the physicists have been led to formulate and to study simple mathematical models. The concepts and methods they developed appear to apply to a number of important random combinatorial optimization problems, for which they have proposed remarkable formulas. Their discoveries point towards a new branch of probability theory. Finding rigorous arguments to support their conjectures is a formidable challenge and a long range program."

6. Anatoly Vershik.
The intrusion of probabilistic methods into the study of algebraic objects elucidates algebraic properties and at the same time provides completely non-traditional examples of probabilistic situations.

7. M. Viana.
We are again trying to develop a global picture of Dynamics recovering in an new and more probabilistic formulation, much of the paradigm of finitude and stability for most systems that inspired Smale's proposal about four decades ago. This is at the core of a program (of Palis) that also predicts that statistical

properties of such systems are stable, under small random perturbations.

The Probabilistic Universe Today
Advances: 1980-2002

In this period many subfields of probability underwent remarkable development and a number of new subfields emerged. These include:

- Regularity of Gaussian processes
- Probability on Banach spaces
- Exchangeability
- Concentration of measure and isoperimetric inequalities
- Empirical Processes
- Excursion theory
- Malliavin calculus
- Diffusions on manifolds
- Stochastic flows
- Diffusions on fractals
- Dirichlet forms
- Stochastic mechanics
- Lévy processes
- Random Schrodinger operators
- Random media
- Percolation theory
- Self-avoiding walks
- Stochastic Loewner Evolutions

- Phase transitions and critical behavior
- Hydrodynamic limits of particle systems
- Stochastic partial differential equations
- Fleming-Viot processes, particle constructions
- Branching processes, catalytic and mutually catalytic branching
- Superprocesses, Brownian snakes, relation with semilinear PDE
- Historical processes and historical stochastic calculus
- Statistical mechanics and random fields
- Stochastic analysis and Malliavin calculus in mathematical finance
- Point processes - applications to queueing theory and survival analysis
- Random trees and other combinatorial structures
- Queueing networks
- Probabilistic analysis of algorithms
- Randomized algorithms
- Bootstrap methods
- Markov chain Monte Carlo

Trends & Challenges &

Emerging Research Opportunities

- Acceleration of the creative dynamics in which new scientific challenges lead to new probabilistic structures and methodologies.

- Continuing discovery of "Universality Phenomena" and study of the deep mathematical properties of the associated "generic" probabilistic objects.

- Search and discovery of new probabilistic models and ideas to give insights and produce probabilistic tools to address the challenge of large complex systems — neither deterministic nor random — but the interaction of both

<u>Eg.</u> • massively interconnected communications networks

- spin glass
- ecological systems
- evolutionary biology
- genetics of complex disease
- cell dynamics
- financial markets
- macroeconomics

References

1. B. Bollobás. Random Graphs, 2nd ed., Cambridge Univ. Press
2. M. Csörg . Personal communication.
3. M. Kac (1985). Enigmas of Chance, Harper and Row.
4. E.A. Perkins. Personal communication.
5. M. Yor (1997) Some Aspects of Brownian Motion, I,II, Birkhauser.
6. Proceedings of the ICM, Volume 1, Plenary lectures, Berlin, 1998.
   - P. Diaconis. From shuffling cards to walking around the building: an introduction to modern Markov chain theory.
   - M. Talagrand. Huge random structures and mean field models for spin glasses.
   - M. Viana. Dynamics: a Pobabilistic and Geometric perspective.
7. B. Engquist and W. Schmid (2001) Mathematics Unlimited 2001 and Beyond, Springer-Verlag.
   - Joel Spencer Discrete Probability
   - Daniel Stroock Probability Theory: A Somewhat Narrow Perspective

- Rick Durrett Probability Theory
- Anatoly Vershik Randomization of Algebra and Algebraization of Probability: An attempt at prediction

8. J.-P. Pier (2000) Development of Mathematics: 1950-2000, Birkhäuser.

- K.D. Elworthy: Geometric aspects of stochastic analysis
- G.R. Grimmett: Percolation.
- J.-Fr. Le Gall Processus de branchement, arbres et superprocessus.
- L. LeCam Le statistique mathematique depuis 1950
- P.A. Meyer Les processus stochastiques de 1950 à nos jours.
- M. Smorodinsky Information, entropy and Bernoulli systems.
- M. Yor Le mouvement brownien: quelques développements de 1950 à 1995.

9. École d'été de Probabilités de Saint-Flour 1971 - 2002, Springer Lectures Notes in Math.