Probability and Statistics: Five Questions (ed. Alan Hájek & Vincent F. Hendricks), Automatic Press/VIP, 2009, p.1-11

FIVE QUESTIONS

N. H. BINGHAM

Question 1. Why were you initially drawn to probability theory and/or statistics?

My first love in mathematics was geometry. This goes back to secondary school (Tadcaster Grammar School, 1955-63), where I discovered that mathematics was not just arithmetic, but also contained algebra (which I liked) and geometry (which I loved on sight). What I still remember is noticing that the mathematics teaching staff picked up that Bingham was good at geometry – and that (and this is the point – I was an insignificant new boy, an underage 10-year old) this was important. I caused a minor sensation when I got 100% in a geometry exam (Easter 1957) – and repeated it in the O Level mock (Easter 1960). At Oxford, I remember saying in my final year that I wanted to do research in geometry – and being laughed at. I was told that geometry was now well understood, and that the world of research had moved on. I only discovered much later that this is nonsense, and that geometry and topology (opposite ends of the same lollipop, as one of my friends put it) are still as vital and important as any other area of mainstream pure mathematics. Later still, I got involved with the geometric side of probability – which is fascinating. It's also fun to see how everything is linked to everything else.

My first exposure to probability was in the Sixth Form at school. It put me off for some years, but not for life, thank goodness. It consisted of an overdose of elementary combinatorial problems – balls of various colours in urns, etc. Mathematically I am naturally an analyst rather than a discrete mathematician or combinatorialist, so looking back I'm not surprised I reacted negatively. I love the combinatorial side of probability now. I overcame my initial dislike, partly through having to teach the material, partly through being won over by excellent accounts in textbooks – Feller [Fel1], and Chung [Chu].

Some time as an undergraduate, I saw in lectures (by J. S. (Jack) de Wet, as I recall) the proof of Weierstrass' approximation theorem by Bernstein polynomials. If $f \in C[0,1]$, and

$$f_n(x) := \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(k/n),$$

then

$$f_n(x) \to f(x) \qquad \text{uniformly in } x \in [0,1].$$

The proof reduces (apart from f being bounded and uniformly continuous) to the weak law of large numbers. What struck me immediately, and dramatically, was that if one knows some probability, one can do (at least some kinds of) analysis better than analysts can. I resolved to make the interface between analysis and probability my natural academic habitat.

One did two special subjects in one's final year at Oxford in my day. I wanted to do applied analysis – analysis because I loved it, applied because I felt that analysis was so vast that one would need a focus to avoid getting lost in it. Besides, I wanted to use mathematics, as well as do it for its own sake. I opted (no doubt partly under the influence of my tutor, John Hammersley – J. M. Hammersley FRS, 1920-2004) for statistics (as it was then called – it was mainly non-measure-theoretic probability) and numerical analysis. The second I enjoyed while learning it, but became bored by during revision – except for Gaussian quadrature, which I found fascinating. This survives in my lifelong love of orthogonal polynomials – which came into my thesis, and play a major role in a number of my papers. The first, probability, was the reason why I transferred to Cambridge to do my PhD under David Kendall (D. G. Kendall FRS, 1918-2007). Hammersley and Kendall had been colleagues at Oxford in the 1950s, and had a warm regard for each other. The move was arranged in one long phone call, while I walked round the University Parks. I never looked back.

As well as turning me into a probabilist, my time in Cambridge (1966-69) turned me into an analyst too. I learned measure theory and functional analysis; I went to all the Part III courses on analysis or probability over my three years there. My involvement in statistics came later, through teaching it. I was advised by Jef Teugels that I should grab the first chance I had to teach statistics. My then boss James Taylor used to say that there are only two ways to understand anything in mathematics properly: one is to do research in it, one is to teach it. Having agreed to teach a statistics course in the late seventies, I learned statistics from Keeping's book (as did Anthony Atkinson, I later learned). I am and always have been a mainstream probabilist, but the statistics dimension to my life has grown naturally into something I value greatly. It has also made me a much better probabilist. Indeed, I see the two subjects as the two sides of the same coin, the mathematics of randomness, and I see knowing both as necessary for balance.

Question 2. What examples from your work (or the work of others) illustrates the use of probability theory and/or statistics for foundational studies and/or applications?

Since you have twenty-five other contributors to this volume (with me the first author only alphabetically), it is safe and sensible for me to give a personal answer.

As I mentioned above, I found my natural academic habitat early in the interface between probability and analysis. The two areas there to which I have contributed most and most usefully are Tauberian theorems and regular variation. The two go together. In my thesis, I dealt with Kingman's theory of random walks with spherical symmetry [Kin]. There, the characteristic function in d dimensions is essentially a 1-dimensional rather than a d-dimensional integral transform, because of the spherical symmetry. If one radializes the Fourier transform in d-space, one obtains a type of Hankel transform, whose kernel involves a Bessel function. I obtained a domain-of-attraction condition in my thesis in terms of regular variation of this transform at the origin. I was then unable to translate this into a condition of regular variation of the tail of the probability law, but was able to do so in 1972, by using special methods (Abel's integral equation – and then essentially fractional calculus). I spoke on this at the British Mathematical Colloquium in 1972. My friend Milne Anderson (J. M. Anderson, the complex analyst at UCL) kindly said to me afterwards that one must be able to do this by Wiener Tauberian theory. I knew the Wiener theory (from Widder's book), realized he was right, went away and did it. This was my way in to Tauberian theory, which has become one of my specialities. I bring to it the viewpoint and motivation of a probabilist. This is valuable because, while all probabilists have to be analysts as well, this doesn't work in reverse. I am very proud that I was able to be of real service to Jaap Korevaar, when he was preparing his magisterial book on Tauberian theory [Kor], in suggesting coverage of the probabilistic dimension here. Of course, it is no accident that Norbert Wiener (1894-1964) was a great probabilist as well as a great analyst, and a mathematical genius.

Turning to regular variation: the limit theorems of probability – the soul of the subject – rest on regular variation, if one demands full generality. Take the weak law of large numbers, mentioned above – one formalization of the Law of Averages of the man in the street. Existence of the mean is sufficient – but not necessary. What is necessary and sufficient is slow variation of the truncated mean. Or again, take the central limit theorem – the formalization of the Law of Errors of the physicist in the street. Existence of the variance is sufficient – but not necessary. What is necessary and sufficient is slow variation of the truncated variance ([Fel2], XVII.5). In the early 1970s, it was clear that a properly organized synthesis of the theory and applications of regular variation was overdue and necessary. This became my first book [BGT] of 1987, with Charles Goldie and Jef Teugels. It is no accident that this book, formally an analysis book on 'a chapter in classical real-variable theory', has as its three authors three probabilists.

There is a coda to this story. A decade is a decent shelf-life for a mathematics book; [BGT] is still going strong after two decades plus. I think the bulk of the book stands the test of time well, but there were two glaring weaknesses, one at each end. The weakness at the beginning (p.11, §1.2.5) concerns the right smoothness condition to impose to eliminate pathologies – finding the right common generalization of measurability and the Baire property. The weakness at the end (Appendix 1) concerns the focus in [BGT] on the theory in one dimension, at the expense of generalizations to higher dimensions, topological groups etc. With A. J. (Adam) Ostaszewski, I have recently (2007 on) successfully attacked both problems. They are linked – which 'short-circuits' the two glaring weaknesses in [BGT]. The bulk of the book survives intact, but the context of the whole subject is now different. Meanwhile there are three new books on regular variation in higher dimensions, all motivated by extreme-value theory [BE], [FdH], [R], and so all written by probabilists. Look no further for what probability and analysis have to offer each other!

My next example concerns mathematical finance. I am a child of the post-war consensus in the UK, and politically left-wing. I had some interest in economics when young, because of the political dimension, and the work of Keynes and Galbraith, but none in finance. Following the work of Black and Scholes in 1973, it became clear that mathematical finance not only was (or would become) a field in its own right, but that it rested on probability theory and stochastic processes. I think I first learned this from seminars by Mark Davis, but it was made obvious in the 1981 paper by Harrison and Pliska [HP]. For maybe fifteen years, I avoided this mathematics of filthy capitalism as best I could (I certainly had plenty else to do!). Then in 1995 I joined Birkbeck College, University of London as Professor of Statistics (my first time sailing under statistics colours, and my first time as head of department). My predecessor had kindly drawn up the teaching schedule for my approval. He had me down for a topics course for the MSc, which I gladly approved. I was then told of an audience request. That particular cohort of MSc students consisted largely of people from the City, working during the day as traders and coming to Birkbeck (the College of the University of London specializing in evening teaching) during the evening to find out what was lurking in the background of their day's trading. I was asked to teach these people the relevant theory, mathematical finance. I swallowed hard, and agreed in real time: what else could I do? ('Or what man is there of you, whom if his son ask bread, will give him a stone? Or if he ask fish, will give him a serpent?' Matthew 7:10-11.) I duly learned the stuff – and promptly fell in love with it. It is such beautiful mathematics. I mourn the days when the UK economy actually made things – steel, ships, locomotives, volume cars, etc. – rather than pieces of paper such as financial derivatives, and had industries, trade unions and the like. But that was then and this is now; meanwhile, the revolution in financial mathematics has made much that was arcane pure mathematics when I learned it as a student the stuff that young folk joining the City need to have under their belt. I have in mind such things as Girsanov's theorem on change of measure (the guts of the Fundamental Theorem of Asset Pricing and risk-neutral valuation [BK]), Lévy processes of infinite activity (infinitely many jumps in finite time – modelling the 'jitter' of price movements of heavily traded stocks under normal market conditions, as supply and demand balance under the impact of a large number of small trades), etc.

Question 3. What is the proper role of probability theory and/or statistics in relation to other disciplines?

Mathematics is the common core of science. Science is what distinguishes between the modern world and the Middle Ages. Without science, we're back in the Middle Ages, burning witches.

Modern mathematics dates roughly from the 1680s, with the introduction of calculus. As analysis, which grew from calculus, developed, it proved a wonderful tool for probability and statistics. Thus measure theory, which is what one needs to do probability theory properly, dates from around 1900, and measure-theoretic probability from say Kolmogorov's *Grundbegriffe* in

1933 [Kol]; on the statistics side, the great flowering of mathematical statistics in the hands of first Galton, and then Pearson, Edgeworth, Student and Fisher, took place around 1880-1930, drawing on not only calculus and analysis, but also linear algebra and geometry. Now that probability and statistics are mature, they are able to repay their debts and feed back into mathematics. Some of the feedback from probability is touched on above. For statistics: it is often said, rightly, that the two most important developments of recent years have been Markov chain Monte Carlo (MCMC) and wavelets. Now MCMC is explicitly concerned with limit theorems in probability theory, but its vital statistical importance has had the effect of flooding parts of probability theory with an endless supply of interesting and statistically meaningful problems. Wavelets originated in work by French geophysicists in the 1980s (the motivation was searching for underwater oil fields). The subject was put onto a proper mathematical footing by Yves Meyer and Ronald Coifman in the 1990s, and shown to be a variant on Fourier analysis. Its power was dramatically demonstrated when the FBI fingerprint data bank was digitized – by Coifman. Here one has the full spectrum from the very pure to the very applied, but with everything feeding back on everything else (incidentally a fine example of why 'pure v. applied' is a false dichotomy).

We live in a random world. We have known since 1926 or so not only that the quantum picture is right, but that (and the first person to state this explicitly was Max Born (1882-1970): [Gre]) the quantum world is essentially random. Thus probability is an essential ingredient, not only of mathematics, but of physics (and hence of the other sciences). A fine example of a quintessentially probabilistic physical theory is provided by quantum electrodynamics (QED). See the splendid account by Feynman [Fey], where QED is presented as a logical development of a very classical physical theory – Newton's account (from his *Opticks* of 1704 [New]) of partial reflection of light from a mirror. Incidentally, I can recall first reading this, and puzzling on how a photon knew on striking the mirror whether to be reflected or refracted. I realized at length that this question, natural though it may be, is about as sensible, or as silly, as asking how a coin knows whether to fall heads or tails.

Question 4. What do you consider the most neglected topics and/or contributions in late 20th century probability theory and/or statistics?

I would say the general area of *feedback effects.* We know from mathematical finance how to model the uncertainty in new price information (tomorrow's, say) by a stochastic process, regarded as a driving noise – the basic case being Brownian motion. Now Brownian motion originates in physics (modelling the movement of pollen particles in suspension, buffeted by the surrounding water molecules). But prices are human constructs, and people are notoriously more complicated than physical particles. When economic and financial agents are taking decisions, they take into account not only their own perceptions of the relevant parts of the economic and financial scenery, but also their assessments of how these things will be perceived by others. The big danger is of herd behaviour or panic ('Gadarene swine'), which can have devastating effects on confidence, and on the stability of the economy or financial institutions, at national or global level. Hence the prime duty of the financial regulators, to intervene appropriately to stabilize matters when necessary, and to provide a regulatory regime that acts to promote stability and prevent crises. International agreements ('Basel I' and 'Basel II') have been put in place for this purpose. Some academic opinion at the time held that, while done with the best of intentions, some of the framework of Basel II might well prove more destabilizing than stabilizing. Recent financial crises (credit crunch, sub-prime mortgages, etc.), and the mixed record of success of the three kinds of institution representing the public interest (government, central banks and financial regulatory authorities) have tended to bear out these forebodings. One is reminded of Sir Isaac Newton, who at the height of his fame lost some twenty thousand pounds (a vast sum in those days) in the South Sea Bubble of 1720; he remarked bitterly that he could more easily control the stars in the heavens than men and markets.

Changing from finance to physics: the two great triumphs of twentieth-century physics are relativity theory and quantum mechanics. Each is clearly right rather than wrong, and each has been enormously successful and productive. However, general relativity and quantum mechanics are known to be incompatible. The great challenge of physics in the twenty-first century is to reconcile the two, and/or to unify gravity with the other three fundamental forces of nature – electromagnetism, and the strong and weak nuclear forces (that hold the nucleus together, and control radioactivity) – to produce a Grand Unified Field Theory, or 'theory of everything'. Great hopes are pinned in this regard on the forthcoming opening of the Large Hadron Collider at CERN. It may be that 'little black holes' – black holes (completely gravitationally collapsed objects) on a very small scale, where quantum effects, and so uncertainty and randomness, dominate – will be crucial here.

In ecology: everything to do with global warming, both prediction and policy, is clearly of vast importance. There is everything for the probability and statistics community here: lots of data, lots of uncertainty, problems of vast human importance, feedback effects between prediction and policy, feedback effects in the complicated dynamics of global climate, etc.

Question 5: What are the most important open problems in probability theory and/or statistics, and what are the prospects for progress?

In probability theory, the accelerating interchange of ideas and methods between probability and analysis provides an inexhaustible source of problems for workers in both fields. We merely list a few key areas:
(i) *Martingale methods in analysis.* Certain aspects of analysis – for example, the Litlewood-Paley theory (1931-37), the Calderòn-Zygmund theory of singular integrals dating from the 1950s, and functions of bounded mean oscillation (BMO – see e.g. [Gar], VI) – are so closely related to martingales in probability theory that the relevant analysis and probability are best studied together; see e.g. [Str] VI. Similarly in functional analysis, particularly Banach spaces: the Radon-Nikodym theorem and martingale convergence theorem are the same theorem, in the sense that if one holds on a Banach space B, so does the other (Uhl, 1969), and if $p \in [1, 2)$, the Marcinkiewicz-Zygmund strong law of large numbers holds for pth moments iff the space B is of type p (de Acosta, 1981; see e.g. [LeTa] §9.3).
(ii) *Long memory and long-range dependence.* What makes probability and statistics work is cancellation. Independent errors tend to cancel – hence the strong law of large numbers, central limit theorem, law of the iterated logarithm, etc. This extends also to weak dependence (a whole hierarchy of mixing conditions), martingale dependence, and to a degree Markov dependence. Typically one measures the strength of dependence by the rate of decay of correlations. As this decay slows, there is an abrupt transition from the classical regime to a non-classical one, in which the remote past (in the setting of time) dominates the present rather than being decently forgotten (one speaks here of long memory), or remote places dominate behaviour where one is (in the setting of space – one speaks here of long-range dependence; relevant here are physical phenomena such as phase transitions). There has been a great deal of interest in such matters (see e.g. [Ber], [Gri]). A recent impetus has been provided by the theory of orthogonal polynomials on the unit circle (OPUC), and the associated spectral theory [Sim]; see [I],

and ongoing work by the author, Akihiko Inoue and Yukio Kasahara.

In statistics, this century may correct the balance of the last one in which Bayesian statistics was largely (though not entirely) developed in a parametric rather than a non-parametric setting. More generally, data mining and the storage and retrieval of information, particularly in real time, is clearly so important in ordinary life that it will increasingly come to dominate the technical agenda. One thinks of such things as bank, computer and financial fraud (insider trading etc.), credit screening and the like, or of the impact of biometric data on security at airports and passport control and the fight against terrorism. In science, cosmology throws up endless theoretical challenges in the interpretation of vast and growing amounts of astronomical data (the search for gravitational waves, etc.). So too does particle physics; no doubt the LHC at CERN will prove a fertile source of both data and problems.

One of the areas of interest of the editors of this volume is philosophy, a word I have not mentioned so far. In my youth I had a positive distaste for philosophy, which I considered a sorry substitute for mathematics and science. I like proving theorems, and hate arguing – which is why I became a mathematician, where one knows when one is right or wrong, and doesn't have to argue, whereas in philosophy there is no such finality. But as one learns more, and grows older, one becomes (or at least I become) more interested in the why as well as the how, and in the broader context in which one's knowledge of different things fits together, and so in the philosophical side of one's field, in some sense. I look forward to reading what my distinguished co-authors may have to say about such matters – and indeed, about everything else.

References

[BE] A. A. BALKEMA and P. EMBRECHTS: *High-risk scenarios: A geometric approach*. Zurich Lectures in Advanced Mathematics, European Mathematical Society, 2007.

[Ber] R. BERAN: *Statistics for long-memory processes*. Chapman and Hall, 1994.

[BGT] N. H. BINGHAM, C. M. GOLDIE and J. L. TEUGELS: *Regular variation*, 2nd ed., Cambridge University Press, 1989 (1st ed. 1987).

[BK] N. H. BINGHAM and Rüdiger KIESEL: *Risk-neutral valuation: Pricing and hedging of financial derivatives*, 2nd ed., Springer, 2004 (1st ed. 1998).

[Chu] K.-L. CHUNG: *Elementary probability theory with stochastic pro-*

cesses (3rd ed.). Springer, 1979 (1st ed. 1974, 2nd ed. 1975).
[Fel1] W. FELLER: *An introduction to probability theory and its applications*, Volume 1, 3rd ed., Wiley, 1968.
[Fel2] W. FELLER: *An introduction to probability theory and its applications*, Volume 2, 2nd ed., Wiley, 1971.
[Fey] R. P. FEYNMAN: *QED: The strange theory of light and matter*. Princeton University Press, 1985.
[Gar] J. B. GARNETT: *Bounded analytic functions*. Academic Press, 1981.
[Gre] N. T. GREENSPAN: *The end of the certain world: The life and science of Max Born, the Nobel physicist who ignited the quantum revolution*. Wiley, 2005.
[Gri] G. R. GRIMMETT: *Percolation*, 2nd ed. Grundl. math. Wiss. **321**, Springer, 1999.
[dHF] L. de HAAN and A. FERREIRA: *Extreme value theory: An introduction*. Springer, 2006.
[HP] J. M. HARRISON and S. R. PLISKA: Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Applications* **11** (1981), 215-260.
[I] A. INOUE: AR and MA representations of partial autocorrelation functions with applications. *Prob. Th. Rel. Fields* **140** (2008), 523-551.
[Kin] J. F. C. KINGMAN: Random walks with spherical symmetry. *Acta Mathematica* **109** (1963), 11-53.
[Kol] A. N. KOLMOGOROV: *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, 1933.
[Kor] J. KOREVAAR: *Tauberian theory: A century of developments*. Grundl. math. Wiss. **329**, Springer, 2004.
[LeTa] M. LEDOUX and M. TALAGRAND: *Probability in Banach spaces*. Ergeb. Math. **23**, Springer, 1991.
[New] I. NEWTON: *Opticks, or A treatise on the reflections, refractions, inflections and colours of light*. Dover, New York, 1952 (based on the 4th ed., London, 1730).
[R] S. I. RESNICK: *Heavy tail phenomena: Probabilistic and statistical modelling*. Springer, 2006.
[Sim] B. SIMON: *Orthogonal polynomials on the unit circle. Part 1: Classical theory; Part 2: Spectral theory*. AMS Coll. Publ. **54**.1,2, American Math. Soc., 2005.
[Str] D. W. STROOCK: *Probability theory: An analytic view*, rev. ed. Cambridge University Press, 1994.

Department of Mathematics, Imperial College, London SW7 2AZ, UK
n.bingham@ic.ac.uk