Mathematical Conversations

Wilfrid Kendall : Dancing with Randomness >>>

Interview of Wilfrid Kendall by Y.K. Leong

Wilfrid Kendall followed in the scientific footsteps of a distinguished father (probabilist and statistician, David Kendall) and established himself as a well-known expert in probability theory who has made significant and wide-ranging contributions to random processes, stochastic geometry, stochastic calculus and perfect simulation. His interest in the use of computers in teaching and research has also led him to develop computer algebra software in statistics and probability.

He is on the editorial boards of numerous leading mathematics and scientific journals, among them *Annals in Probability*, *Statistics and Computing*, and the *London Mathematical Society Journal of Computation and Mathematics*. He has been invited to give lectures at major scientific meetings and conferences, and has served on the committees of international scientific organizations. He is a professor in the Department of Statistics of the University of Warwick.

He was the Chair of the Organizing Committee of the Institute's program on Markov Chain Monte Carlo (MCMC) held in March 2004. The Editor of *Imprints* interviewed him at the Institute on 17 March 2004. The following are edited and vetted excerpts of the interview, in which he talks about the early formation of his career interest, the role of randomness in computer simulation, the close connection between probability and statistics and his views about the place of computers in statistics and intellectual thought.

Imprints: I'd like to thank you for giving us this interview in spite of your busy schedule. Where did you do your PhD and what was it on?

Wilfrid Kendall: I got my PhD, or DPhil as it is called there, at Oxford. I was an undergraduate at Queens's and then a graduate student at Linacre College. My thesis ended up with the title "Three problems in probability theory", which was very naughty of me because I had been told that the PhD title should be informative about what the PhD is about. But I was so anxious to get it submitted that I forgot all about the title until the last minute. There were, of course, three problems in the thesis. One was to do with early work on the knotting of Brownian motion, one concerned contours in random fields and one related to work I had done with my father on the statistics of shape. They were probability or statistics topics and they all had some kind of geometry attached to them, which has continued a fairly common theme in all the work I have done. My supervisor was John Kingman. He, in fact, was almost supervised by my father. Well, he was supervised by my father, but he never got around to submitting his PhD – never needed to! So my father is also my academic grandfather except for that small technicality.

Wilfrid Kendall

I: Your father is a famous probabilist and statistician. How much were you influenced by your father?

K: It's a very interesting question to me. In one sense, enormously - the fact that he was a working mathematician, that research was clearly exciting and interesting for him. That had a great influence on me. On the other hand, he was very wise, and he knew then what I know now; that if you are following in your father's footsteps, then it can be a difficult path sometimes. And so he never pressed me at all. Occasionally he would tell me a little bit about mathematics but I never felt any compulsion from him to do mathematics or statistics. It was all a choice of my own free will. In fact, my free will was really well informed! At the time when I came to choose my subject for a PhD, my tutor at Oxford, whom I admire very much, warned me that it could be difficult to follow in one's father's footsteps. He gave me very sound, very helpful advice, and he said afterwards that I had listened to him very politely and then I went away and did just what I was going to do before. But I really did take what he said very seriously. However I found I hugely enjoyed doing not just mathematics but also probability and statistics. So I was led that way. I was doing it because it was interesting and engrossing. I didn't want to do it because it was something my father had done. I would be quite strong on that point to anyone in the same position as myself. You really must be sure that you are doing what you want to do because inevitably there are going to be times when it is difficult, and then you'll need to know that you made your own choice for yourself.

I: What is the difference between applied probability and statistics?

K: That's a tricky question! It's like asking what's the difference between strawberry and cream. They are very close, and it's really nice to have both of them together. In

Continued on page 9

Continued from page 8

statistics, the questions are different: you are saying that there are things you want to know about, so you estimate and you test your hypotheses and so on. In probability, you are saying, "The system is behaving randomly and I want to know how it's going to behave." It's a different kind of question. It's going the other way, if you like. There is not a clear cut line of division.

I: What are some of the hottest topics or developments in applied probability?

K: I asked people at lunch about this question and we all agreed that this is a very hard question! Certainly I can give a personal answer: what is hot for me is all the things that I like doing at the moment. The whole interface between probability and computing is very interesting. What we are doing now (in the Institute's program) is only a small part of that. There's a lot going on. Some of the work going on in random matrices is absolutely brilliant, and there's some lovely stuff to do with percolation theory. It is a difficult question to answer, particularly about applied probability. Some parts of science, and even of mathematics, are like a huge factory. You just have one or two products, and everybody involved is somehow working on the same products. It may take a long time before they eventually produce something really big. There are other parts of mathematics, and probability is one of these parts, which are extremely creative and vigorous, but there is no great master plan to which everybody tries to contribute a little. Instead, it's a very rich and fertile field and there are lots of different problems coming up all the time with a lot of premium on being original and trying to find your own way to do things. Temperamentally, I find that much more exciting. But it's difficult to say what the hottest development is in probability. You can say what you like to do right now but it's probably unwise and counterproductive to try to have much influence on what everybody else chooses to do.

I: Or shall we say, what are the central problems in applied probability?

K: Well, I think there are central problems that people are looking at and getting intrigued by. I'm not sure if you should talk about probability problems. They typically are problems to do with mathematical science generally. For example, at the moment some of my friends are extremely interested in random matrix theory because they think it might have something to do with the Riemann Hypothesis. Sometimes people think there is something there, and sometimes people think it's a mad dream. But it is interesting in its own right. There are other questions which have really been there a long time in statistical mechanics – whether there is some universal structure hiding behind things like percolation. There are people who have done some exciting work related to that. There are certainly big questions that people would like to think about. But I think that it's true generally that's what makes probability an attractive and vigorous subject, why a lot of people are attracted to it; there are lots of things to do and they are all very interesting. No one can quite tell what will be the next new development.

I: Could you tell us how the term "Markov Chain Monte Carlo" came about?

K: Monte Carlo refers to the process where you want to calculate something and it may be too difficult to do either by hand or by using a computer to find the integral directly, and you try to do it instead by doing a random experiment, which involves the probability of interest. It actually goes back a long way – the famous Buffon needle problem. You drop a needle onto a lattice of lines. Find the probability of overlapping the lattice. (Clue: it is related to pi.) But Monte Carlo itself is a term coined probably during the war because of computational demands in the development of the atomic bomb. Why Monte Carlo? Well, because the method had to do with the roulette wheel and randomness. Markov chain Monte Carlo is a particular way of doing Monte Carlo. If you like, you could read it as "Monte Carlo with Markov chains". So when you are doing these random experiments, the question is how are you going to do the randomness? For example, like tossing a dice, tossing a coin, or running a roulette wheel. You may do it indirectly, you may say let's build a stochastic system, a Markov chain, and let's design it so that it has an equilibrium distribution which is what we are interested in. Then you run it for a long time and you observe the outcome and that gives you a way of handling the calculations. The adjective "Markov chain" describes a way of doing a Monte Carlo.

This idea goes back a long way, one of the first ideas people were using. There are many complicated problems for which the quickest approach is to relate them to the probability of long-run behavior of Markov chains. There was a very famous paper by Metropolis and others which goes back to the 1950s, but almost certainly they were doing a lot before that. The physicists who have a lot of money to buy big computers have always been interested in computing and developed it. Relatively recently, statisticians started to persuade people to buy them computers too. And the computers got flown in and sit on the statisticians' desks. At that stage, a large number of statisticians started to get involved using computers. Once they have the computing power, then it started to become a more feasible way to solve problems. It is pretty effective and has a tremendous influence upon the way people are doing statistics now.

I: How much of the new developments in probability and statistics have been dictated or influenced by the advances in computer technology?

Continued on page 10

Continued from page 9

K: I think, a huge amount. Here is a very simple example: the sort of questions that I used to mark for undergraduate examination papers when I started lecturing have largely gone out of fashion because they had to do with hand calculations but now you simply use a statistical package. I think that had a very big influence on the sort of things one does because some things have become very easy. One no longer thinks about them. But then, that means you can pose much harder questions. Markov chain Monte Carlo is another kind of example; computations that would have been inconceivable without accurate computing power. And then there are other applications, not really applications, but problems stimulated by the presence of computers and computation. You get interested in different kinds of questions. Back at Warwick I have a number of people I like to spend time to talk to – many of them are in the statistics department, many in the mathematics department, and also a significant number in computer science because probability is now important if you want to understand how to analyze the behavior of computer algorithms.

I: What about the theoretical aspects? The computer is good for computations, but will it have any influence on the theoretical development in probability?

K: As soon as you know how to do something, that there is a possibility of an answer, then your theory changes because your theory is about how you do things and you have just acquired a whole new way of doing things. That means you need a whole new theory. You can trace that all through statistics. What people are interested in theoretically is very often driven by the things that they can already do on the computer, which suggests theoretical questions. And then people on the theoretical side are motivated to do new things.

I: There seems to be a prevalent faith in some kind of order underlying every random, if not chaotic, behavior. Do you see this as a new paradigm in science or even in mathematics?

K: It's a very old paradigm. For example, in the book of Genesis, God builds order out of chaos. I think the idea of order coming out of chaos is not particularly new! Indeed, the mathematical notion of chaos can be viewed as saying there is a randomness in the choice of initial conditions right at the beginning, but you only see it bit by bit as the system evolves. I don't think there is any real conflict between randomness and systems with a great deal of order. Adrian Smith once said that probability is about what you don't know. You make probabilistic statements about the things you don't know are happening. It is perfectly compatible with ordered complex systems. Some things you don't know about, maybe you'll find out bit by bit as the system evolves. You can even use probability to do it. In fact, we had a conference in Durham in the summer which was to do with Markov chains in the overlap in between many different areas. And one of the things that was very interesting to see is that the group of people using Markov chain Monte Carlo in statistics were often working towards the same end as people who study deterministic systems with no randomness whatsoever and who are finding that the theory of Markov chains is a good way to describe how the initial conditions propagate through the system.

I: Does probabilistic modeling require the design of a "perfect" random number generator or some similar "random process generator"? Is that achievable in practice?

K: The answer to the first part is "no", and the answer to the second is "probably not". Practically, what you need is something which generates random numbers which are good enough. You don't want a number generator that produces a periodic sequence 0, 1, 0, 1, …. That's not good enough. How good is "good" enough? It's good enough if it does what you want it to do. If it has done its job, then it's good enough. A lot of work goes into the design of an arithmetical random number generator. From time to time, it gets replaced by one that is thought to be better and sometimes one can indicate rigorously how good the properties of these random number generators are. Indeed we have just had an example in the workshop: someone was talking about the case where you can show, a bad choice of random number generator leads to errors in certain complicated Markov chain Monte Carlo calculations. So you have to be careful. There is no replacement for the computer in your head. You've got to think about these things.

Suppose you want to produce a perfect random generator. Maybe you go to quantum theory, but there are all sorts of ways that things can go wrong. For example, suppose you built it wrong. That's embarrassingly easy to do if it's of complicated design. I recall a friend of mine who tried to build random number generators using thermal noise. He said that it was going to be perfect. He set up the stuff which electronically converted the thermal noise into noughts and ones, and it had a subtle correlation in it. He showed it to me and we agreed "It's wrong. There is not enough random deviation." Eventually he traced the problem to some subtle kind of electrical feedback.

This morning, somebody was talking about the design of generators of random bits based on a Geiger counter but they failed to take into account the fact that the Geiger counter worked better at higher voltages and there was a 24-hour period fluctuation in the voltage supply to the Geiger counter. So in a technical sense it wasn't doing the job it set out to do, producing more random bits at some times than at other times. You have to realize that in the

Continued on page 11

Continued from page 10

black box you are using to produce a string of random numbers, there's probably going to be some factor there which you can't quite control and which you might have left out. When you take that perspective, then it doesn't seem so crazy, on the other hand, to use what we call a pseudo-random number generator using an arithmetical sequence because at least, you understand the properties of that. One of the criteria in the practice of random number generation is that you should prefer a random number generator whose defects you know to one whose defects you don't know. There is a nice quote about this. John von Neumann said back in the fifties, "Anyone attempting to generate random numbers by deterministic means is, of course, living in a state of sin." You have to do it, you are using a random number generator, at the back of your mind there may be something wrong with the generator, or maybe it's something wrong with the way you code the thing. One of us was just estimating coding error probabilities this morning. He reckons that the programs he writes have a 40 percent chance of being wrong in a first working draft. For my programs the chance is probably higher. Once you take that into account, you start looking for the bugs you *know* must be there!

I: Are there any limits to the levels of computer simulation? Do you think that computer simulation can shed some light on some of the mysteries of life such as the way the brain functions or even the origin of life itself?

K: The answer to the first question is: "Yes, there are limits". The answer to the second question is clearly yes and clearly no. The first question is interesting. My friends in computer science tell me about some very interesting theorems which show that there are practical limits to what we can do with computer simulation and which are related to algorithmic limitations to do with the phrase "NP-complete". You are looking at a world of problems of scale. In other words, when you double the size of the problem, does the work you do double or quadruple or worse ... or much, much worse ... and hence you can derive notions of hierarchies of difficulty of algorithmic problems. You can get the same sort of hierarchy for problems to do with computer simulation. So there appear to be logical limits as to how much can be done with computer simulation.

Now to your second question. Science certainly can shed much light on amazing things. Everyday, for example, I read about new progress in understanding and control of diabetes. On the other hand, you just have to look into the eyes of a new-born baby to realize that there is something about which science remains silent. If, by the mysteries of life you mean something like that, the answer is: No.

I: Could you tell us something about the role and position of computers in mathematics education at your university?

K: Our department was one of the early UK statistics departments to use the computer in a big way in teaching statistics, so we were early starters. At Warwick, we have a center which tries to encourage innovation in the use of computing and it has taken on a very practical strategy. It recognizes that there are people using computers in all sorts of different ways across the university. It produces a newsletter which reports on these ways. It encourages people to experiment a bit and to report what is useful. Now, for example, whenever I give a talk or a course, I make sure that my lectures have notes on the web which are highly hypertexted so that they have all sorts of links in them. Increasingly, people say they like them and find these helpful. But I think that, while innovation and experimentation are good things, it's important always to bear in mind that actually education is ultimately about what is going on in people's heads.

I: Is it compulsory for Warwick statistics students to do some computer programming?

K: As a matter of fact, it is. Our students are all exposed to a course using the computing package Mathematica. But the point I'm making is that in the end what matters is when people walk out of the classroom or computer room, have they changed their way of thinking? Have they actually learned anything? You don't need a computer to make a difference to that and sometimes the best thing we can do to help people learn is to put the computer in a quiet corner of the room and switch it off. What matters is what's going on in people's heads.

I find the computer a great aid in making illustrative material available to students when I talk on some topics. It makes a tremendous difference if they can actually learn how to do things and see them afterwards. However it's important not to get lost in all those technology.

We teach our students to use computer packages rather than programming as in such flexible packages you can learn how to program. We don't teach them programming as a primary activity. Typically, when they come out into the world, what they need to know is how to use the computer as a tool. That is clearly the way things are progressing. Programming is done by some people but what is most important is for people to know how to develop the qualities of systematic thought and care that are required for programming.

I: Do you have any connections with the Warwick mathematics department?

K: Yes, I have a lot of friends there. In particular the Warwick probabilists are almost equally divided between mathematics and statistics. Probability is at the boundary and it is a good and interesting place to be in.