Raf Bocklandt
Korteweg-de Vries Institute for Mathematics
University of Amsterdam
r.r.j.bocklandt@uva.nl

Nicos Starreveld
Korteweg-de Vries Institute for Mathematics
University of Amsterdam
n.j.starreveld@uva.nl

Interview David Aldous, winner Brouwer Medal 2020

Probability and the real world

On 6 April, during the NMC 2021 Scientific Days there was a postponed ceremony for David Aldous, the winner of the Brouwer Medal 2020. To get a better picture of the man behind the medal, Nicos Starreveld en Raf Bocklandt, editors of NAW, conducted a Zoom-interview with the laureate, during which they talked about life, work, teaching, and the interplay between probability and the real world.

Cambridge

David Aldous was born in 1952 in Exeter, a city in the southwestern tip of England. In 1970 he went to Cambridge University to study mathematics.

"At that time in England there was a sudden transition: in high school you are a child, in college you're an adult, and I think that is much less the case nowadays. You had a lot of elective classes and you were not required to attend the lectures. Basically, you only had to do the exams at the end of the year. The unique feature of Cambridge and Oxford wasn't the lectures but the fact that you had one-on-one or two-on-one tutorials with a graduate student or a faculty member. They assigned you problems for homework, but it was not graded. You just showed up for the tutorials, your tutor would ask how you did on the questions, and you would discuss the problems. In this way the supervisor would find out much quicker who was doing well and who needed help. This system is I think more personal than anywhere else in the world."

At first his passion was mainly analysis, the interest in probability theory started growing in the second year of his studies, when he followed a course on Markov Chains given by the probabilist and queueing theorist David Kendall.

"A funny story that I remember regards the final exams in the last year, my strongest subjects were general topology and probability, and Cambridge had this rule that the exams were set by someone who was not the instructor of the course. We had very good instructors for both courses but the final exam in topology was made by John Conway and in probability by Béla Bollobás. They posed very creative questions which had almost nothing to do with the content of the courses, so I didn't do so well.

After graduating, I went for one year to Yale University as part of an exchange program between the two universities. I still remember the lectures on weak convergence given by queueing theorist Ward Whitt, who was then a young assistant professor. Weak convergence was a topic that directly drew my attention as it combines in a very elegant manner analysis and probability theory. After the year in Yale, I returned to Cambridge. The university took me on as a graduate student without having completed all the official requirements. Things were all a bit more informal back then."

The Brouwer Medal for David Aldous

Exchangeability

David Aldous obtained his PhD with a dissertation with the intriguing title *Two Topics in Probability Theory*. One of these topics concerned exchangeability and was inspired by work of John Kingman, who was a professor at the university of Oxford.

"At that time, John Kingman was investigating the subsequence principle. This was a very deep idea concerning limiting arguments of sequences of random variables. I remember taking the bus to Oxford, which although only 80 miles apart took about three hours to get to, to listen to Kingman's lecture about this topic. At the end of the lecture, he formulated a very strong conjecture, which seemed natural but turned out to be wrong. I had already thought about this conjecture and had come up with a counterexample, but I wanted to be very cautious because he was the most famous probabilist in England. So back home, I wrote to a friend in Oxford to check it before giving it to Kingman."

Traces of the subsequence principle are already present in the strong law of large numbers which states that, for any sequence of independent and identically distributed random variables $X_1, X_2, \dots$ the sample average $\frac{1}{n}\sum_{i=1}^{n} X_i$ converges almost surely to the mean $\mu = \mathbb{E}X_1$. This holds not only for the average of the whole sequence, but also for the average of any of its subsequences.

In 1967 János Komlós proved the astonishing result that, for any sequence of random variables with bounded expectations $Y_1, Y_2, \dots$ (but not necessarily independent or identically distributed) there exists a subsequence $(Y_{n_i})_i$ and a random variable Y, such that its sample average, and that of all its subsequences, converge to Y. This discovery ignited a discussion among mathematicians on the universal character of such asymptotic results.

"Chatterji had formulated the following heuristic principle. Given a limit theorem for independently and identically distributed random variables under certain moment conditions, there exists an analogous theorem such that an arbitrarily-dependent sequence (under the same moment conditions) always contains a subsequence satisfying this analogous theorem. The limit theorem above is the precise form of this principle for the strong law of large numbers. Instead of formalizing a general framework Chatterji had proven some special cases [1].

Kingman was working on this topic and was certain exchangeability played a crucial role. His endeavour was to find an appropriate framework which would allow him to make the principle rigorous and to include all the special cases that had been proven by Chatterji in one general theorem. His approach relied on asymptotic exchangeability and resulted in Kingman's conjecture (mentioned earlier) for which I had a counter-example."

David Aldous

Photo: Wikimedia / George M. Bergman

But this was not the end of the story. Aldous devised a more refined and mathematically precise version of the subsequence principle and showed that it is true both for any almost sure limit theorem and any weak limit theorem, subject only to some mild technical conditions.

The idea of Aldous was the following: from an arbitrary sequence of random variables $Y_1, Y_2, \dots$ with tight probability distributions (a natural condition which assures that no probability mass evades to infinity) he managed to extract a subsequence $(Y_{n_i})_i$ and relate it to an exchangeable sequence $(Z_i)_i$ with similar characteristics. Exchangeable sequences are sequences of random variables whose joint probability distribution remains unchanged after permutations of the variables. This implies that every subsequence has the same distribution, and Aldous used this to show that certain types of properties of $(Z_i)_i$ are shared by subsequences of $(Y_{n_i})_i$.

"Exchangeability was back then an obscure topic in probability theory and not so many people cared too much about it. No one since 1981 has ever cared about this specific subsequence work. But another aspect of exchangeability, the analog of de Finetti's theorem for arrays known loosely as Hoover–Aldous–Kallenberg theory, did emerge from obscurity twenty years later in the context of graphons. For me exchangeability is just there. What are the key ideas in mathematics? Symmetry is one of them if you think of group theory. If you ask how symmetry appears in probability theory then the answer is exchangeability."

Berkeley

Not long after his PhD, Aldous got a job at the University of California in Berkeley, so in 1979 he moved across the Atlantic once more. He would remain there for the rest of his academic career because for someone working in probability theory Berkeley was an excellent choice.

"Berkeley was well known because of the Berkeley Symposium on Mathematical Statistics and Probability, which was held about every five years from 1945 to 1972. In that era you didn't have many big international meetings in particular subjects. Somewhere in your library you can probably still find the proceedings of the Berkeley symposium and they contain a lot of interesting papers.

Everyone in the field knew that Berkeley was a center for probability and statistics even though they might not know who was actually working there. When I applied for my job in Berkeley, I remember walking down the corridor and being introduced to a kind old man. This turned out to be Jerzey Neyman, one of the founders of modern statistical testing. Just as you don't expect to walk into Euler or Gauss, you kind of assume that all the big textbook names are dead, so I had to restrain myself from shouting 'my god, you're still alive'. Now forty years later I'm waiting for someone younger to have that same reaction to me. I haven't heard it yet, but maybe someday ...”

Unlike in many other universities, the statistics group in Berkeley is not part of Mathematics but has its own separate department.

“When I arrived in 1979, its focus was mainly on mathematical statistics. It had the reputation of being one of the few places in the world where they were doing theoretical statistics. We had about 15 to 20 faculty members and we were teaching half of all the Berkeley students some introductory course in statistics. This was part of the reason of our existence. Most of the staff back then were hired for their research but we all taught some of these service courses.

Nowadays that has changed, and these courses are taught by specialist teachers, which I think is good for the standard of teaching because some of my colleagues and I were frankly not very good at teaching elementary statistics. In the first half of my career, I taught an introductory course in statistics to students who did not know any calculus or algebra. I actually enjoyed it because, contrary to an introductory course on analysis, you could draw a lot of your data and examples from newspapers. On the other hand, teaching a statistics course is like being a dentist, people don't really want to be there.

Because the statistics department and the mathematics department were separate, math students could graduate without ever having followed a probability course. The math curriculum at Berkeley was very much focused on pure mathematics and probability was optional. In fact, some of the staff at the mathematics department did not consider probability to be part of mathematics at all. The two departments did not mingle that much; there were some older people who taught fifty-fifty at both departments but that had almost stopped by the time I arrived there. By the end of the nineties the connections were renewed and people like me were given a formal appointment at both schools. They finally agreed that probability and statistics were part of mathematics.”

Microsoft

Although Aldous often jokingly says that he only held one job in his life, this is not entirely true. In 2009 he took a sabbatical year and moved to Seattle to work at Microsoft Research.

“My wife was born and raised in the San Francisco Bay Area and had lived there for whole of her life. She had some interest in moving sometime somewhere else and Seattle came to mind. Yuval Peres, a colleague of mine at Berkeley, worked at Microsoft Research and had some funding to invite me over for a sabbatical.

Microsoft Research is an interesting place. It acts a bit like an interface with the academic world and they employ a lot of

The Brouwer Medal

On the 6th of April David Aldous received the Brouwer Medal 2020 for his contributions to both pure and applied probability theory. The prize was awarded during the Scientific Days of the NMC 2021, which were held online due to COVID-restrictions.

In his laudation, Frank den Hollander, chair of the nomination committee, praised the great originality of Aldous's work combined with depth, breadth and beauty. He also indicated that Aldous has introduced a wealth of new concepts, which have had an enormous impact not only on probability itself, but also on other scientific fields. He has initiated flourishing new areas of research and has inspired researchers worldwide. In addition, he has contributed to the popularization of probability through numerous channels, such as his website 'Probability and the Real World'.

In particular, the committee highlighted the following scientific contributions:

- The ALOHA protocol, which is used in communication networks and was modified after Aldous proved that the original protocol was flawed.
- The probabilistic analysis of performance of algorithms in combinatorial optimization.
- The Aldous tightness criterion, which provides a connection between stopping times and tightness and links general theory of stochastic processes and weak convergence.
- The cut-off phenomenon for mixing times, a new concept for Markov processes pioneered with Persi Diaconis.
- His Landmark series of papers on limits of tree-like structures which lead to a better understanding of universality.
- His work on exchangeability that led to extensions of de Finetti's theorem and became a major tool in different areas of probability theory, including random graph theory.
- Local weak convergence, a key concept for convergence in random graphs.
- The Poisson clumping heuristic, a guiding notion for description of rare events with a wide range of applications
- His description of critical clusters in the Erdős–Rényi model for random graphs.
- Miscellaneous work on scale-invariant random spatial networks, discrete spatial networks, flows through random networks, recursive distributional equations, and stochastic models for phylogenetic trees.

The Brouwer medal is only one of the prizes and fellowships awarded to Aldous. He also received the Rollo Davidson Prize and the Loève International Prize in Probability, and became a Fellow of the Royal Society, the American Academy of Arts and Sciences, the American Mathematical Society and many more.

post-docs temporarily to know what is going on academically. Microsoft is incredibly careful about who they hire permanently. You can't get a job at Microsoft unless you worked for them as a kind of consultant for several years. You'll be well paid, so it is not like the gig economy, it is a bit like the tenure system in academia.

The research lab has about forty groups, each of which had about ten permanent people. They were supposed to be looking ahead at what Microsoft might be doing ten years down the line. Thirty-nine of these groups were doing something recognizable that might be relevant. Some were working on speech recognition or simultaneous spoken language translation, which were not at all doable at the time. Others were working on the Kinect-system, a motion controller that allowed you to play games by moving around.

I remember helping one of the groups that was modeling the parallel computations for their search engines. They were writing a paper about it in seven weeks to present at a conference. I was supposed to help them out with the math, but they kept changing the model every week. In mathematics we are a lot concerned with the complexity of algorithms and their running times, but it turned out that most of their problems were related to moving data around such that they were accessible by the different processors, which led to some interesting queueing problems."

Working at Microsoft also came with another perk for someone interested in statistics and probability: data.

"The fortieth group was the probability group, which was only there for obscure historical reasons. I mainly worked with them, and I didn't do much that was directly relevant for Microsoft. Around the same time their search engine Bing had just come into operation, so they had access to a huge amount of new data. I got hold of a file with all the search queries that contained the word probability or chance. This gave me some real-world data to play with and I obtained some surprising insights.

It turned out that about half the chance queries had to do with health and medicine and about half of those with birth control and pregnancy. Another big part of the medical questions concerned cancer: if you are diagnosed with a certain cancer, what are the chances of surviving? Most of the nonmedical questions are still related to concrete personal events like accidents or career opportunities. Only a few people asked questions of a more philosophical nature, like the probability of life and evolution, or the vague 'probability of chance of something going wrong'."

An in-depth discussion on these search queries can be find on Aldous' website: https://www.stat.berkeley.edu/~aldous/RealWorld/bing_chance_topics.html.

Probability and the real world

Although many of Aldous' contributions to probability theory approach the subject from an abstract mathematical level, his ideas are often inspired by real life situations. An example of this is a toy model for waiting lines at a gate before boarding an airplane.

"Imagine you are waiting back in line at a passport control checkpoint. As people reach the front of the line they are being processed steadily, say it takes ten seconds to have a passport checked. Hence every ten seconds there is a departure from the queue. But not everyone in the queue moves forward constantly every ten seconds. When a person leaves the checkpoint, the next person moves up to the checkpoint, the next person moves up and stops behind the now-first person, and so on, but this 'wave' of motion often does not move through the entire long line; instead, some person will move only a short distance, and the person behind will decide not to move at all.

I decided to make a model of this which is rather simple to understand quantitively. This spatial queueing model is very interesting because it relates to something that has been very well studied, which is coalescing Brownian motion on the real line, but it relates bizarrely to it in the sense that time and space are interchanged. This was a funny discovery because I could just draw a picture of coalescing Brownian motion and then say that our model has exactly the same behavior with the only difference that time and space are interchanged, which unfortunately messes up the Markov property, and at a precise level this makes everything difficult. It was a fun problem to work on, in the end, we managed to explain an observation, but it didn't lead to something that can help you control or design queues."

Aldous has also contributed to the popularization of probability. In recent years he has taken an interest in various aspects of the real world that involve chance. On his website visitors can find a section 'Probability and the Real World', where they can read a plethora of interesting and stimulating ideas about everyday facts that relate to probability. So, if you want to know more about human biases, making predictions, (mis)using toy models, probability paradoxes or other mathematical curiosities, we strongly recommend you have a look and take a chance...

Reference

1 D.J. Aldous, Subsequences of sequences of random variables, *Bulletin of the American Mathematical Society* 83(1) (1977), 121–123.