After Dinner Talk to Statistical Society Members

Don McNeil, Sydney, 28 November 2001

I'd like to reminisce on 45 years learning and teaching Statistics at universities. First I'll summarise the ups and downs of what has been on the whole a very lucky career, and describe some of the people who've influenced me. I'll finish by viewing how computers have developed and make a prediction, and hopefully convince you that some very exciting times are ahead for all of us.

When I was 16 I sat for the Entrance Exam to the University of Tasmania and did extremely well, gaining the top mark in the State. This was a proud moment for my parents and teachers, who concluded that I must have been some kind of child prodigy. In fact, the reason for my success was quite simple. The State Library kept copies of past examination papers, going back for many years. Perusing them, I discovered that questions repeated themselves. So I worked out the answers to all the past questions and learnt them off by heart. Also, one of the librarians was an extremely attractive girl, so I went to the library every day in the hope of seeing her.

When I went to the University of Tasmania, it was much more difficult, and I nearly failed two of my first-year subjects. My grades improved slightly in second-year, but deteriorated in third year. Having no idea of what to do with a pass degree in Maths, I applied for 4th year honours, and despite my poor record, was accepted.

The Professor of Mathematics was Edwin Pitman. He could teach on any aspect of pure or applied mathematics or statistics. First, we would listen to him while he explained an elegant solution to a problem. Then he wrote the argument on the blackboard, and we transcribed it into our notebooks. I can still remember the lectures where he explained asymptotic relative efficiency. And I'll never forget his motto that a mathematician's main objective is to avoid work.

Graduating at the end of 1960, aged 20, with a BSc honours second class, school-teaching was all I seemed qualified for. But the Tasmanian Education Department required teachers to be aged 21 to get adult pay. Annoyed at this, I applied for a job at the Gordon Institute of Technology in Victoria, where no such age discrimination existed. However, my teaching at this institution was neither enjoyable nor successful, and when LS Goddard, the newly appointed head of Maths at the University of Tasmania offered me a fellowship to teach half-time and study for a research master's degree, I seized the opportunity.

For my thesis, Professor Goddard gave me some papers to read on spectral representations of birth and death processes, published in the *Proceedings of the London Mathematical Society*. Pure mathematics was not my strong point – I never got past rings and fields – so progress was nonexistent, and he told me that I had no hope of success in the academic world and should not think of attempting a PhD. Depressed at this diagnosis, I joined the University dramatic society and descended into the world of drink.

Fortunately, I was saved by two lucky events. First, a new professor of applied mathematics, David Elliott, replaced Edwin Pitman, who had retired a year earlier. **Although Edwin Pitman was a great scholar, he had never had any graduate students in Hobart, and he never encouraged his undergraduates to do further study. In this respect David Elliott was the opposite. He wanted to build up a graduate**

program, with his students in adjoining rooms so he could interact with them. He agreed to supervise me, even though he knew nothing about my topic.

My second, and more important, stroke of luck was meeting a girl and instantly falling in love. She supported me and eventually reformed me. Abandoning my wicked ways for a higher calling, my thesis progressed too, after David Elliott suggested using traditional complex variable calculus rather than spectral theory to tackle my problem. He arranged for the thesis to be examined by Pat Moran in the Statistics Dept at the ANU, where I was offered a PhD scholarship. Lyn and I honeymooned in Sydney and moved to Canberra, where I commenced working under Pat Moran's supervision in January 1965.

Although I had now tasted success both in love and work, my years in Canberra were plagued by fears that my original supervisor was right to think that I would never make a scholar, and that sooner or later I would be exposed as a fraud. **I was lucky to get the opportunity to succeed and throw off these self-doubts, but I think that supervisors should never be so discouraging, Having now supervised dozens of masters and PhD students myself, I can say quite definitely that the few failures I've had were the result if insufficient encouragement on my part.**

Pat Moran was a man of few words. Like David Elliott, he would sit in his office writing papers, with the door open so his students could visit him whenever they needed help. Because I was worried that he would discover my lack of ability, I rarely consulted him. I just imitated him and sat in my room and wrote chapters of my thesis. He had given me a good topic, and there were plenty of supportive people around who could help me, like David Vere-Jones, Warren Ewens and Ted Hannan, so it wasn't too difficult. After two years he asked to look at what I had written, and with trembling hands I showed him my chapters. He just nodded his head and said "you've got enough", dismissed me, and continued his own writing.

There was another department of Statistics at the ANU, headed by Ted Hannan, and containing interesting people like Warren Ewens, Eugene Seneta, Des Nicholls, and Chip Heathcote. When I learned that Warren had accepted a Chair at LaTrobe University, I asked Ted if he would consider me for a temporary lectureship while seeking Warren's replacement. Ted happily agreed, and I thus spent an interesting year, learning all sorts of things from Paul Winer. **And I learnt from Ted how to run a department, simply by focusing energies on research and supervision of graduate students, and delegating administrative matters to trusted staff. Two of Ted's students, Dean Tyrrell and Des Nicholls, went on to become Vice Chancellor and Dean of the Faculty, respectively.**

My next move was to the Statistics Department at Johns Hopkins University, where Pat had arranged a job for me with Geof Watson. **Geof was the opposite of the doubting supervisor. If he saw a skerrick of talent in any of his students or junior colleagues, he magnified and extolled it, and I was no exception.** Filled with the confidence Geof gave me, my research thrived at Johns Hopkins. And when Geof moved to Princeton in 1969 to set up a new department of Statistics separate from Mathematics, I was invited to chair the department at Johns Hopkins.

But Baltimore in 1969 was not an attractive place to raise a young family, and a Readership in Mathematical Statistics was going in Perth, with the prospect of a Chair in a year or so. Full of excitement at the opportunity to develop a department of Statistics in my home country, I moved to the University of Western Australia.

Alas, my hopes and dreams were soon shattered. First, **I realized that I was in a Mathematics department where deductive logic was king and applications, data and computing were paupers.** I was a fish out of water. Second, a mining boom in Perth had caused land prices to skyrocket. With no assets other than a hilly block of land in Hobart, a reader's salary was hardly enough to build a home in Perth. My relations with the Head of department plunged so much that he replaced me as head of the Statistics section by a visiting combinatorial analyst. In despair I resigned and returned to Hobart, intending to have another go at school-teaching, and to build a house overlooking the Derwent.

But luck again intervened. Geof Watson hadn't been able to attract the people he wanted to Princeton, and in desperation he turned to me. I took my family back to the US, and slowly gained self-confidence. Princeton had a big IBM computer, the most powerful in the world, and to simplify accounting, the first 2 seconds of

CPU time were free. With a bit of ingenuity almost any statistical problem could be solved in 2 seconds on this computer, so our computing was mostly free. I had fun developing a crude GIS program that had all the boundaries of European provinces stored and could print character-based contour maps of demographic data in the 2 seconds allowed.

The department was small. **Apart from Geof and me, the only other permanent member at first was John Tukey. Peter Bloomfield joined us the next year. John was writing a book on exploratory data analysis and was teaching this material to first-year students. The trouble was that the students couldn't understand him, and the course was in danger of collapsing through lack of enrolments. So Geof and I decided to give John a rest, and to hire a visitor to teach the material. Needless to say, we couldn't find any visitor brave or foolhardy enough to teach John's holy writ under his nose while he was still developing it, so of course I ended up doing it.**

Mastering and teaching exploratory data analysis transformed my life. I soon realized that most of John's material could be made more attractive to students by computerizing it. For example, he had a simple robust method for two-way anova called *median polish*. This is extremely tedious to do by hand, but is easily programmed in a computer. Box plots demanded similar treatment. Princeton had a new IBM 370, which ran an interactive language called APL. When I taught the course with the students using terminals interactively, enrolment increased dramatically, and the whole Princeton football team, mostly burly black guys, enrolled at one stage. **I wrote a book called *Interactive Data Analysis*. People called it the first English translation of Tukey!**

It was hard to escape Tukey's shadow. At a local meeting of the statistical society I was asked to give a talk on "Tukeyisms", and a huge crowd turned up. I knew that John was going to be out of town that day, so decided to have some fun with the audience. I pretended to have misunderstood my assignment, assuming that "Tukeyisms" meant stories about John Tukey, rather than explaining things he had invented like stem-and-leaf plots and hanging rootograms. The audience appreciated my stories, particularly the one about the time when I was walking in a blizzard and he offered me a lift in his car. You see, the maths professors at Princeton had a competition to see who could drive the oldest car. It was clear that Tukey was aiming to win this competition, because when entering his car I found that the floor on the passenger's side had rusted away and my feet were touching the snow-covered ground! The audience was clearly enjoying my stories, which became increasingly outrageous, and I'm afraid, for effect, I even embroidered them a little. So you can imagine my shock when, in the darkened room, I actually saw John Tukey himself sitting in the front row. Hearing that such a large crowd was coming to hear the talk, he had decided to cancel his trip to Chicago. I quickly sobered up and started explaining box plots and median polish.

In 1973 we decided that our department needed its own computer, so we bought a PDP 11/40 running a new operating system called Unix, and put it in a room we shared with the maths department. This was before computers had been used to solve the four-colour problem, and the Princeton mathematicians didn't take them very seriously. In fact they regarded computers as the devil's invention. Putting one in their area infuriated them, particularly when students like Dave Donoho switched their major from maths to stats to use our computer.

Princeton was tremendously exciting, but looking back, a great opportunity was missed. My first two years were largely devoted to research in stochastic models, and Feller Volume 2 was my bible. Robert Macarthur was Professor of Biology, and had written a book on mathematical ecology. Sadly, Robert was dying of cancer, and he persuaded Bob May, an Australian Physics Professor, to continue his work. Bob told me about Ito processes, which had received only a brief mention in Feller, so were quite new to me. At about the same time the economists asked me to give a course of lectures to them on time series analysis. If I'd taken Ito processes more seriously, I would have told the economists about them, and then learned about their applications to options pricing, and perhaps been able to get involved in the exciting development of the Nobel prize-winning Black-Scholes formula.

But I could see some things going wrong at Princeton. Our department was rich with its endowments and research grants and had a healthy number of good PhD students, but its undergraduate students, the university's life blood who provided for its future through their endowments, were being neglected. I missed Australia, and wanted to get back. I applied for a Chair in Computer Science in Hobart, but not having a

PhD in Computing, ran second to a more qualified candidate. When the Chair in Statistics at Macquarie became available, I didn't hesitate.

The difficulty was that the fledgling department of computing was firmly in the grip of the mathematicians, and statistics was firmly in the grip of the economics school, and neither department was the slightest bit interested in statistical computing. But I was full of confidence so I accepted the position.

Statistics, being associated with the data-starved discipline of Economics, was not thought to need funds for computers. This made it extremely difficult to develop computing labs for our students. Having seen what computers could do, I was like a carpenter being asked to build a house after his tools had been taken away. During my first 12 years at Macquarie, I devoted all my energies to establishing these computer labs. Progress was slow and exhausting, and often we went backwards. Computers cost money, and there was never enough. I set up a consulting lab and used its profits to buy computers, and we developed a statistical package that we sold commercially. But when the money started to flow in, the university objected to me running a business on campus, and put a stop to it. So the problem was not only lack of money, but lack of support from the university administration.

During my early years at Macquarie the number of students majoring in Statistics grew to respectable levels, and we established a Master's program. This strengthened our case for a second Chair, and **in 1990 Victor Solo was appointed. He immediately improved our research profile, something that I had largely ignored in my preoccupation with computing. And a stream of excellent PhD students started arriving.**

But under attack from the administration with my statistical lab destroyed, another low point had been reached. I began my love affair with South-East Asia.

I'd been there before. In 1982 I went to Pakistan to carry out a nationwide sample survey of household health expenditure. I remember an official in Islamabad telling me about group of Thai demographers who had recently visited his office. He took a dim view of them. "All they did was laugh and drink beer!" he said. In 1985 I visited Hasanuddin University in South Sulawesi as a member of an IDP team, and returned to various islands in Indonesia on several occasions, becoming increasingly drawn to the land and its people. I was like a small boy who had escaped through a hole in a garden wall to find an exotic wonderland on the other side.

Our relations with Indonesia were better in those days than they are now, but ignorance fed prejudice on both sides. In 1985 three academics accompanied me to Ujang Pandang. We stayed at an expensive hotel in the city where the food was atrocious and the service snail-paced. After a tasteless western-style breakfast washed down by stale instant coffee in a land that grew the best coffee in the world, we were chauffeured to the university, about 15 km outside the city, arriving at about 10 am. There we met various officials and talked vaguely about curricula, and had lunch at 12 o'clock. After lunch the campus emptied, and there was nothing much to do but go back to our hotel in the city and write reports on our lack of progress. We complained about the difficulties of helping people who worked for such short hours, and joked about Indonesian "rubber time".

I learned the truth when I went back by myself a few weeks later. I persuaded the IDP bureaucrats that my Indonesian counterparts needed a computer. The IBM PC had not yet been invented, so I took an 8-bit machine called a Datamax. I chose this computer because we were using them to teach first-year Statistics at Macquarie. A small company in Manly manufactured them, and I also wanted to show the Indonesians that they didn't need to get all their technology from America.

The machine was heavy and well exceeded the weight limit for air travel, but I got it there in one piece and installed it in a room on campus. I offered to teach the locals how to use it for word-processing, database management, and statistical analysis. The response was good: 12 persons registered to come in pairs for two-hour sessions, from 6 am to 7 pm, every day for two weeks. Since I had to get to campus before 5 am to prepare my classes, I stayed in a small guest-house outside the city, and travelled to campus in a bicycle-taxi called a *bejak*. **I thus discovered that accommodation in Indonesia was excellent and cheap, that the**

food was delicious, and that people actually worked very hard. I found that when I arrived on campus at 5 am, students and teachers were already there, and that the teachers who disappeared after lunch were actually going to teach at another institution because their salaries were so low that they needed a second job to survive. They returned in the late afternoon to teach in the evening as well.

My Hasanuddin experience had a sad ending. After I returned to Australia, the dean took the Datamax to his office, where it sat, covered with a dust-jacket like a shrouded statue in a long-vacated church, until advancing technology made it obsolete. My students never got to use it again. I know this because one of them wrote to me and told me of her frustration. I learnt an important lesson from this experience. **You can't just go to a developing country, drop some pearls of wisdom, and leave never to return. You've got to keep going back.**

In 1987 IDP invited me to run a short course in Applied Statistics at Prince of Songkla University in Southern Thailand. Remembering the words of the Islamabad official, I readily accepted this assignment. I helped the medical Faculty to set up a Master's program in Biostatistics and Epidemiology, and then established my own MSc program in the department of Mathematics and Computer Science. In breaks between semesters and the odd study leave, I was actually doing more teaching and research in Thailand than in Australia, and enjoying it more. I accepted an appointment as external examiner for the department of Applied Statistics at the University of Malaya in Kuala Lumpur, and PhD students came to Macquarie from both places. I now supervise PhD students in Biostatistics for the Medical Faculty, and next year we plan to start a PhD program in the Faculty of Science and Technology. I've made about 30 trips to Prince of Songkla University in 14 years. In contrast to my sporadic and relatively ineffectual visits to the Indonesian universities, I'm confident that the Thai experience will have a lasting effect. In the five years since we started the MSc program in 1986, 21 students have graduated, not a lot less than the number to have graduated in the same time with our Master of Applied Statistics from Macquarie. The course content is much the same and their theses are of similar quality to those of their Australian counterparts. If anything, the Thai students have been more successful in getting their theses published.

Finally, after 20 years at Macquarie and only four years before my retirement, our Statistics department got a lucky break. The university decided to establish a Master of International Business program, and invited departments to contribute units. It was clear that these students weren't going to be attracted to Statistics *per se*, so we offered a unit called "Computer Applications in Business". It was largely just showing the students how to use Microsoft Office and HTML, but we gave it some academic backbone, using the software to illustrate statistics and database concepts. Pretty soon we were getting nearly 100 students every semester. We then offered another unit called "Information Management", comprising more advanced database and data analysis emphasizing statistical graphics, which I'd been teaching for years to a handful of our own master's students. Its enrolment jumped to 50 within two years. Flushed with success, we invented units in "Decision Support Systems" and "Web-Database Engineering", thus providing all the core units in an IT major.

These postgraduate enrolments have allowed us to broaden our enrolment base, and have finally given us enough money to set up our own computer lab. The irony is that we don't need money any more for computers. The price of computers is now so low that anybody can just do this! Just consider what has happened with computers.

At the University of Tasmania in 1963, we used an Elliott 503 with paper tape to read stored data and programs into its tiny memory. The computer was so expensive that the university needed to share its cost with the State government. The 60s were the decade of bulky, expensive, batch processing mainframes, like Princeton's IMB 360. Individual departments couldn't afford their own computers.

The 70s saw the advent of time-shared computers such as the IBM 370s and the DEC PDP 11s, for which Unix was invented as an affordable and powerful alternative to the IBM 370 Operating System. The PDP 11/40 in our department at Princeton cost about twice an academic's salary.

In the 80s, Vaxes continued the reign of time-shared computers for a few years, but were supplanted by personal computers, first Apples and a variety of 8-bit processors and then IBM-compatibles, running DOS, and networked using Novell software. These cost about 10% of an academic's salary. But more important

than cost, they steadily increased their power during the decade, doubling in speed, memory, and disk storage every two years.

This technology continued to develop through the 90s, when Microsoft's Windows operating system became dominant. And their hardware improved even faster. A computer with 1000 MHz speed, 256 Mb RAM, and 40 GB disk storage, much faster than anything we've seen before, now costs about 1% of a professor's salary. Allowing for peripherals and software, this means that a department can set up a lab with 50 workstations for what it pays just one professor in one year.

The decade ahead has already signaled its computing profile. It's the decade of the Internet and the World Wide Web. The limitation is no longer the power of the computer on one's desk, because even massive number crunching and Monte Carlo simulations – those that would take hours on expensive mainframes in the past – can now be done in seconds on a desktop PC. The limitation is now the bandwidth of the Internet connection.

The Web offers a golden opportunity for statisticians. After struggling for 25 years, at last my department has its own independent computer system, just as it did at Princeton when I left. As I had feared, the department at Princeton didn't survive, but like Rip van Winkle, it's really wonderful to be able to wake up and continue those exciting times. And it's so much more interesting now. In contrast to the crude character-based maps I had fun creating using the IBM mainframe, we can now produce them in beautiful colour using standard GIS software. My students can create a Visual Basic program to graph data stored in a spreadsheet, simply by recording mouse-clicks as a macro, tidying up the source code, and attaching it to a customized toolbar. Now that spreadsheets have solver tools it's so easy to do maximum likelihood estimation. You can even show second-year undergraduates who hate mathematics how to do this, and they enjoy it. And they love using HTML and ASPs to create their own web pages. While lots of people are teaching data mining and data warehousing, these fashionable topics are central to Statistics.

And at last, the time might have finally come when distance teaching is not only possible, but cost-effective on the Web.

Of course it's easy to say this. Doing it is more difficult. We must learn to harness the technology. One thing we must give up is our dependence on so-called computer experts who think they know what's best for us. The academic world, particularly in Australia, is littered with the debris from botched attempts to set up all-embracing computer systems for academic and administrative use. They want to use the web to wrap us up so tightly that everything is slow and the system is regularly down for repair, justifying their actions as necessary for security.

Don't believe this. Nowadays, teaching how to set up a web-site does not require a network or even an Internet connection. In fact, building a web-site while connected to the Internet is a lot like trying to build a boat in the middle of an ocean. Yet that is how many academic departments are forced to do things.

With the new information and communications technology that's enveloping us, there are so many exciting things to do now that I really don't want to retire gracefully. Anyway, thank you for your indulgence in bearing with me.