World-leading statistics education

The only talk at the Royal Statistical Society's (RSS) World Statistics Day event in London on October 20 will be given by Professor Chris Wild, below (top), Dr Maxine Pfannkuch and Matt Regan, all from the University of Auckland. The RSS says the paper "is set to transform the international landscape of statistical education". Jenny Rankine spoke with them.

The invitation to speak recognises the innovative and very visual ways they have developed for students to think about their data, and the underpinning research supporting developments in the New Zealand school statistics curriculum.

The New Zealanders will also feature strongly at the society's fourth International CensusAtSchool Workshop in Plymouth the week after, launching the RSS' ten-year statistical literacy campaign in the UK.

The recognition also reflects international excitement about the new curriculum. "The involvement of professional statisticians, researchers, teacher developers and lead teachers in developing the secondary school statistics curriculum is unique in New Zealand," says Wild.

"New Zealand is leading in curriculum scope, meeting students' future practical needs in work and life, and in how it represents data. International software developers want to work with our Ministry of Education, because they anticipate other countries drawing from the country's statistics curriculum and want their software to be part of it."

Says Pfannkuch: "Statistics in the curriculum is fairly new in many countries, and usually a lot less developed than in New Zealand. The majority of people trying to deal with statistics run into mathematical roadblocks; we can use visual methods to avoid them.

"Data imaging software can help teenagers to understand patterns in data, and allows teachers to introduce statistics concepts to younger students. We are aiming for students to make inferences about the world without taking their eyes off their data animations, so the connections between question, data and answers are immediate and obvious.

"Looking at the world using data", says Wild, "is like looking through a rippled glass window. What we see is not quite the way it really is. Statistical inference is about how to take that into account."

Pfannkuch, with Pip Arnold, leads a project called Building Students' Inferential Reasoning, developing classroom implementations of new statistics learning in Years 10 and 11 with a team of eight teachers.

Students use new hands-on activities and data animations, reinforced by physical gestures, to learn how to take sample size and variability into account when making inferences.

"Statistics was largely taught descriptively; now we're putting in the conceptual underpinnings," she says. The new curriculum focuses on fundamental thinking about questions and interpretation of the data, rather than the mechanical aspects that computers can do.

"Some of these data visualisations allow students to compare samples of 10, 100 and 1,000 to see the effect of sample size on the stability of estimates."

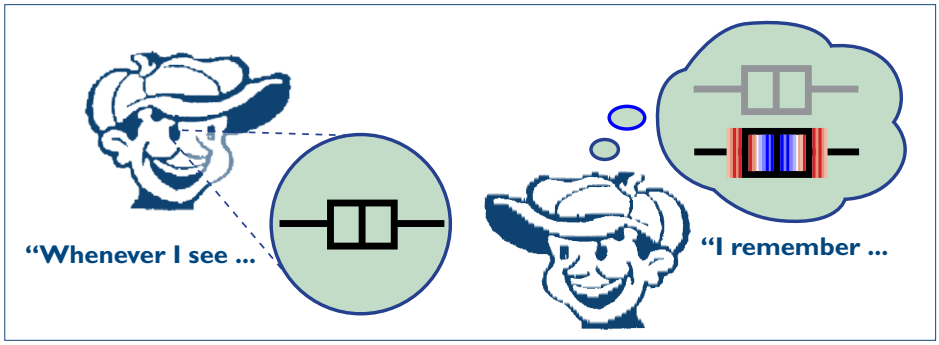
The Census At School New Zealand project, run by the University of Auckland Statistics Department, and supported by Statistics New Zealand and the Ministry of Education, enables students to collect data about themselves every two years.

"Because the data is about them they are interested and engaged," says Pfannkuch. "They understand the background, and can hypothesize about why things turn out the way they do. Also, because they contribute their own data, they know what can go wrong and they can pick up dirty data. Learning data cleaning is also new at the school level."

In a world where every sports game presents statistics, and every health article mentions health risks, where Google and other web applications are massive users of statistics, and surveys and polls monitor all kinds of activity, the New Zealand curriculum should improve young people's ability to participate in decisions and social debates about evidence. The RSS shares this goal – hence its 10-year statistics literacy campaign, getstats.

See also

www.rss.org.uk/pdf/Wild_Oct._2010.pdf (RSS talk)
www.censusatschool.org.nz/2009/informal-inference/WPRH/ (talk animations)
www.rsscse.org.uk/news/rsscse-news/315-getstats
www.getstats.org.uk/
www.gapminder.org/ (data animations illustrating the UN Millennium Development Goals)



Statistics opportunities

New Zealand statistics is strong in biosciences such as ecology, biology and genetics. This includes natural populations of animals such as birds; plants and ecologies; fisheries; crop production and animal breeding. It is also strong in graphical displays for statistical information and in earth sciences. In health New Zealand has led the world in aspects of epidemiology, the tracking of health and illness across populations. Survey sampling is strong in certain areas, such as some official statistics, and social issues such as gambling.

The field faces many opportunities. There is a world-wide shortage of survey statisticians, and computing technology is enabling new ways of seeing and animating data, such as animated population pyramids and maps of commuter flows. "The combination of mapping and dynamic graphs is taking off," says Professor Sharleen Forbes, of Victoria University. Statisticians are also challenged by huge increases in the amount of data available.

IMAges profiles a few of the many statistical developments and applications in Aotearoa.

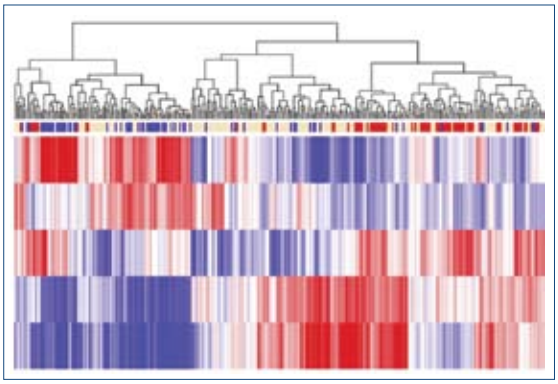
Testing for cancer

Dr Mik Black works with the Cancer Genetics Laboratory at the University of Otago, which is developing gene expression signatures – patterns of genetic activity – to predict outcomes for people with cancer. Scientists examine tumour samples after surgery, and Black uses standard statistical classification methods to predict whether the cancer is likely to come back or not.

"Three signatures indicate particularly aggressive cancer, and they have been patented by Pacific Edge Biotechnology," he says. Clinical trials are the next, very expensive, step to turn those signatures into diagnostic tests.

These take a long time; "we have to wait for five years to see if the cancer comes back. We started six years ago, patented the signatures about four years ago, but there are a few years to go to get it into hospitals as a working test."

"Genetic statistics is most rewarding if we can work closely with clinicians, as they are the people actively caring for patients. Anything we can do that they can translate into improvements for patients - that's the real reward."



Identifying blocks of genes with highly correlated activity profiles in breast cancer using the PCOT2 methodology developed by Black's former PhD student Dr Sarah Song.

Census@School

Census@School had its beginnings when Professor Sharleen Forbes convened a group of NZ Statistics Association members in 1990 to run the first children's census in New Zealand schools.

"We ran it without any money, getting government departments to pay in kind with paper and printers – I wouldn't ever attempt it again! It was totally voluntary, on top of our day jobs."

"Students wrote about 60,000 reports and six or seven of us gave up our holidays to analyse and feed it back to schools. It was picked up by the Italians and then the Royal Society in the UK developed an internet version." This became New Zealand's first Census@School in 2003.

census at school new zealand

It's about us 2010 CENSUS IN SCHOOLS

Census at School

World Statistics Day

A free international poster competition for the International Statistical Literacy Project will be announced on World Statistics Day on 20 October. The theme is 'It happens in my neighbourhood' and it offers national and international prizes. It is open to groups of up to three students born in 1995 and younger, and those born in 1992 and younger.

The one-page poster must tell a story about a set of data. Schools can register at www.stat.auckland.ac.nz/~iase/islp/competition-second; registration closes on Friday, 17 December, and a maximum of two posters per age category per school must be submitted by 18 April, 2011. National winners will be announced on 31 May.

POSTER COMPETITION ISLP

A public conference on Women in Statistics will be held on World Statistics Day at Victoria University of Wellington, with the aim of encouraging women's participation in statistics.

Statistics "has long been an area that girls and women are attracted to because it has a real-world orientation," says Professor Sharleen Forbes, of Victoria University.

One of the conference speakers, Associate Professor Megan Clark, says women make up around 60% of applied statisticians, and around 30% of theoretical and mathematical statisticians. Clark heads the School of Mathematics, Statistics and Operations Research at Victoria University.

The conference will be chaired by radio presenter Kim Hill, a member of the Minister's advisory committee on official statistics. Other speakers include Associate Professor Jennifer Brown, president of the NZ Statistics Association and head of the Department of Mathematics and Statistics at Canterbury University; Professor Natalie Jackson, director of the Population Studies Centre at Waikato University; and Rachel Milicich, Manager of National Accounts at Statistics New Zealand. Contact Lu.folau@stats.govt.nz

20.10.2010 World Statistics Day

Service - Professionalism - Integrity

Analysing interactions in natural ecosystems

Statistics is very important in ecology for analysing multiple ecological variables at once, says Professor Marti Anderson of Massey University. "You might walk a transect in a forest, or swim a certain distance underwater and count the individuals of every species you encounter. Each species is a variable and species interact with each other and the environment. One of the challenges is to understand how sets of species change together, either naturally or in response to human-induced changes."

Anderson is pictured with Associate Professor Russell Millar getting ready for a fish biodiversity survey in Northland. She has developed new computer-intensive methods of multivariate analysis for biodiversity and community data.

"The biggest problem with ecological variables is they don't behave like normal Gaussian bell-shaped curves, and almost all classical statistical methods are based on that assumption."

She has worked on communities of Antarctic plankton and bacteria, organisms living in sediments in estuaries, butterfly communities in the tropics, marine fish communities in kelp forests, microalgae in freshwater streams, insects collected in pitfall traps, forest communities, and even suites of behavioural chemicals in birds.

Her software (PERMANOVA+) enables variation in these complex systems to be partitioned, allowing effects of a disturbance to be assessed against natural variation. It is being used around the world in many ecological applications and environmental impact assessments.

How many possums?

Environmental monitoring drove the development of a simple tool to calculate the population of possums in an area.

Associate Professor Jennifer Brown, of Canterbury University, says: "We spend a lot of money trying to control possums, rats and other pests. We want to know the level at which the population impacts on the environment, and whether we have been effective in managing them. If we have reduced the size of pest populations, are we seeing a gain in conservation?"

With Pest Control Research Ltd, her team developed a wax tag that possums can bite and from which they could calculate the surrounding population. "It was revolutionary, compared with the labour-intensive traps we used to use; the tag is now used throughout the country."

She has also been involved in designing survey protocols to find rare species in their environments. "You can waste a lot of time trying to find rare species," says Brown. "The survey method seems simple but there is a lot of statistics behind it."

The method has been used to find desmans, a very rare river mole that lives in the Pyrenees in France, as well as in Southland to find invasive weeds before they start spreading.

See www.pestcontrolresearch.co.nz/research-monitoring.htm#3
www.pestcontrolresearch.co.nz/monitoring.htm
www.mathsreach.org/Videos

Photo: Malcolm Thomas, Pest Control Research.

R + L = ?

Not satisfied with a statistical package that has "revolutionised the practice of statistics", according to the Royal Society of New Zealand, Ross Ihaka, co-creator of R, is working on the next generation, with the working title L.

R is a free, open-source, extendable model with the highest hit-rate for mathematical publications in the last decade. It is available from more than 75 websites in more than 30 countries.

However, Ihaka (Ngati Kahungunu, Pakeha) says "the world is changing so fast that we desperately need something new now. Data volumes are exploding, and we have no idea as statisticians how to go about analysing petabytes [1,000 terabytes] of data."

His work on L is still theoretical – "you have to get the basics right otherwise you're constrained by your early decisions" – but shows promise of being thousands of times faster than R.

Assessing a vaccine

Roughly 200 cases of Meningococcal B were avoided by the MeNZB vaccine between 2004 to 2008, according to a statistical analysis of the vaccine's effectiveness. Eighty percent of people under 20 were vaccinated, "a quite remarkable proportion, with the highest coverage in those under five," says Dr Richard Arnold, of Victoria University.

Working with epidemiologists in the Ministry of Health, he used a Poisson regression model to compare vaccinated and unvaccinated populations. The bacterial infection is spread by airborne droplets and is associated with overcrowded households. Infection varies by age, deprivation and ethnicity, so he also controlled for those factors as well as regional, seasonal and yearly variations.

"The epidemic had peaked in 2001 and was on its way down naturally when the vaccine was introduced, but we found the vaccine was between 70% and 80% effective in avoiding the disease."

A Norway rat on Okahu Island, Bay of Islands. Photo: Stephen Cope.

Indigenous statistical power

In 2002, Te Ropu Rangahau Hauora a Eru Pomare, at the University of Otago, wrote an influential paper about the need for equal explanatory power – the production of information for Maori health and development to at least the same depth and breadth as that obtained for non-Maori.

Discussing the NZ Health Monitor surveys by Statistics NZ, Bridget Robson (Ngati Raukawa) argued that good governance "compels us to ensure that data produced by the Crown is at least as productive for Maori as it is for non-Maori".

The simplest method for equal explanatory power is to recruit equal numbers of Maori and non-Maori responders to surveys. Random surveys include approximately 15% Maori and 85% non-Maori, and "will be more likely to meet Pakeha health needs". The end result is that health surveys "may have the unintentional effect of increasing health disparities".

Robson argued that implementing equal explanatory power in surveys of health and social determinants of health, such as unemployment, "will help to break this cycle of persistent inequalities".

See www.fmhs.auckland.ac.nz/ faculty/ tkhm/tumuaki/_docs/Equal_explanatory_ power.doc

Tasting applies with the Plant and Food Research sensory science team.

Tracking rat invaders

Rodent Invasion Project member Associate Professor Rachel Fewster, of the University of Auckland, is regularly asked by Department of Conservation staff to identify the origin of rats found around the country. A few years ago she and others obtained genetic profiles for rat populations from many islands around Great Barrier and Stewart Islands, and the Bay of Islands.

"Since then they have been eradicated, but new rats have turned up. DOC or the Auckland Regional Council send us a sample and ask us where it came from." She was able to say recently about two rats from the Bay of Islands that one was almost certainly brought in by boat and the other might have been a swimmer from the mainland.

She examines 20 genes from each rat from DNA regions with a lot of variability. "We use microsatellites which don't code for anything or do any harm if they mutate. In isolated populations, rats will develop their own proportions of those genes. If I get a rat with Gene A, I think it is more likely to come from the island where Gene A is common. We take all 20 pieces of genetic information, and get a fairly clear idea of which island it came from."

Stats and the senses

Mark Wohlers one of 11 statisticians in Plant and Food Research around the country, working with scientists to ensure experiments have the statistical power to determine true treatment effects.

For example, he designs and analyses the results of blind tastings by the sensory science teams, which use panels of tasters to assess wine and fruit from New Zealand grapes and orchards. "They might be checking on length of storage or time of picking, or the effect of a different rootstock. Tasters sit in separate booths in front of a computer, ranking up to 15 variables about the taste and smell of the product.

"I determine, for example, the presentation order; they may not score the same thing similarly each time because of the tasting order. If they taste something very sweet first, the next one may be ranked lower. We might use different coloured lights to take away the effect of the colour of the fruit."

"I often use analysis of variance, sometimes multivariate analysis, and principle component analysis with bootstrapping techniques."

Census and death records

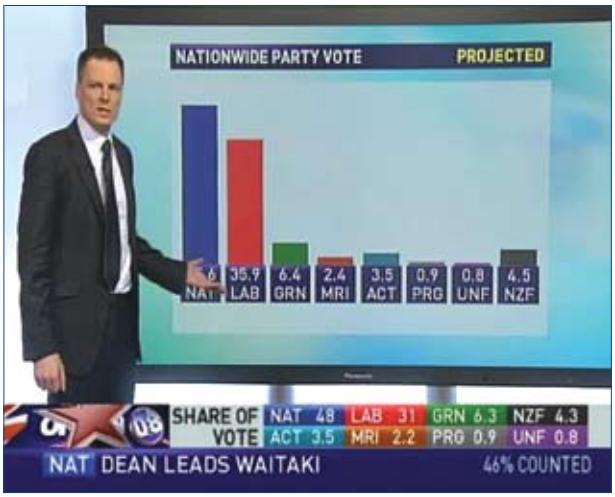
In 1998, death records were the first data set to the linked to the Census in what became the NZ Census-Mortality Study (NZCMS). "At the time," says NZCMS director Professor Tony Blakely of the University of Otago, "it was probably the biggest example of its kind in the English-speaking world."

The linkage was anonymous and probabilistic, enabling researchers to calculate death rates in the whole population for the three years after each census from 1981 to 2004.

"The NZCMS showed there was a great undercounting of Maori deaths in the 1980s and 90s. There was also little, if any, improvement in Maori mortality rates in the 80s and 90s at a time when non-Maori mortality rates dropped significantly. It's very tempting to ascribe that to the Rogernomics reforms and resulting high Maori unemployment rates."

The study also showed that relative gaps in mortality between high and low income groups widened in the 1980s and 90s.

See www.uow.otago.ac.nz/academic/dph/ research/HIRP/index.html



Statistics on the telly

Dr Richard Arnold was the face of statistics during election night coverage on TVOne in 2008, and he got his predictions "bang on". His knowledge of the country's demographics from his time at Statistics New Zealand meant he was able to develop a statistical model to forecast the result.

"There is always an early preference for National on election night because the smaller booths that finish counting first tend to be rural, and more likely to go to National." He adapted and implemented a statistical forecasting method based on matching polling places between elections, which eliminated that early bias. Making sure that the prediction has a reliable margin of error was an important part of the process, because "a prediction without a margin of error is worthless".