Statistics are working


The World Food Programme in Nepal

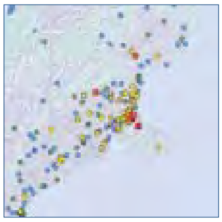
IMAges continues with illustrations of some of the many statistical developments and applications in Aotearoa.

Modelling earthquakes

An automated statistical model of the stresses in the earth's crust may eventually be incorporated into Geonet, the national array of seismometers monitoring earthquake activity in Aotearoa.

With his colleagues in Earth Sciences at Victoria University, Dr Richard Arnold has developed new and more robust methods of estimating the properties of the ruptures that cause earthquakes.

"An earthquake occurs when two blocks slide against each other in a particular plane, sometimes with an area of only a few square metres. We need to understand the three-dimensional orientation of the fault plane and the movement in that plane to characterise the tectonic stresses that drive the earthquake.



These stresses are anisotropic – the crust is compressed more strongly in one direction than another.

"A single earthquake gives only a very limited view of what's going on in the crust, but multiple earthquakes accumulate statistical evidence to characterise the stresses: Smaller earthquakes on lesser faults help us understand the stresses the large faults are experiencing. Changes in stress can indicate impending seismic events, including volcanic eruptions." He is automating the model so that it can monitor stress changes routinely across the country.

Targeting aid

Countries including Nepal, Cambodia, Timor-Leste, Bangladesh, the Philippines and Bhutan have benefited from New Zealand statistical expertise in estimating poverty for efficient aid allocation, in work funded by the UN World Food Programme.

Professor Stephen Haslett of Massey University in Palmerston North has worked with teams using generalised linear models combined with the country's own census data to produce detailed small-area maps of poverty levels. "If you just look at national sample surveys, they're not accurate enough. What we're doing is affecting the allocation of $100m of aid a year in Nepal alone; the money may not be spent anywhere near as well without it. A lot of sample survey results get used to form policy."

He has also designed new water and sanitation surveys of small and medium enterprises in countries such as Azerbaijan and Vietnam for the World Bank, and national surveys monitoring entrepreneurship in Tonga and Uganda.

Survey design

The reliability of population estimates in ecology is always a concern for statisticians. Standard techniques using randomly-placed transect lines in forests or oceans to measure population numbers were known to produce estimates with a high degree of uncertainty, says Associate Professor Rachel Fewster, of the University of Auckland.

Systematic survey designs, where transects start from a certain point and evenly cover the area, were known to be more reliable, how much more was unknown so they were assigned the same poor reliability as random lines. "With systematic designs, the first line determines everything, so we can't use standard statistical theory," she said. Fewster's variance estimation for systematic surveys enabled the improvement to be estimated, and showed from repeated simulations that the systematic estimates can be much more reliable than random estimates.

The result "made a big difference to a Canadian survey of threatened dolphins in fiords", and to surveys of hyenas in the Serengeti. When correctly estimated, the surveys were up to twice as reliable as previously thought. The method was included in the software package Programme Distance, distributed by the University of St Andrews in Scotland.

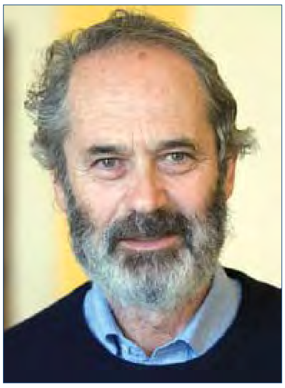
Finding gene copies

Dr Mik Black at the University of Otago is involved in next generation gene sequencing, "which is driving the Thousand Genome Project, an international effort to sequence fully the genomes of 2,500 individuals throughout the world. The amount of data is phenomenal – we're all upskilling ourselves on how to handle it."

One aspect of the project is looking at particular gene copy numbers. "We have two copies of each gene, one from mum and one from dad, but we can also have more copies of any gene from ancient variations. Sometimes this has become an increased risk for particular diseases. The analysis is not particularly hard but the volume of data is huge, so finding the multiple copies is difficult."

Abstract beauty

Professor Jan Saxl of the University of Cambridge was fascinated by the abstraction of group theory when he first studied it at university, and has been working in the field ever since.



Groups are algebraic structures made of a set and an operation that combines any two of its elements to form a third element. The presence of groups in many areas of mathematics and other disciplines makes them a central organising principle.

"By solving an abstract problem you can solve different questions in different subjects all at once. Abstraction allows you to do things you can't do otherwise."

"Octagonal coins have symmetry group D_8. Theoretical chemists can explore what shapes molecules may have, using groups of symmetries. The graphs in which I am interested have a large group of symmetries, and are best studied using this group."

With others, Saxl has worked to classify the maximal subgroups of almost simple groups. "Simple groups are the building blocks of finite groups. They're similar to primes in that all the finite groups decompose into chains of simple groups."

He is studying maximal subgroups: "my groups invariably act on some set, whether the vertices of the coin or atoms of molecules. Subgroups are substructures; maximal are the largest possible substructures."

He is also working on a big project about distance transitive graphs. "These are highly symmetric graphs, which have a very large symmetry group acting on them." During his visit he was working with Professor Eamonn O'Brien at the University of Auckland "on an enormous group acting nicely on a very large graph. A graph for me is really a group acting on that graph. This graph has about 2^{28} vertices, so it is impossible to draw. Eamonn can work with a group of size 4^{30} on a computer."

"The group is much larger, but more manageable. The Hoffman-Singleton graph (right) is complicated, but some aspects are much clearer within its group of symmetries, which is manageable despite its size."

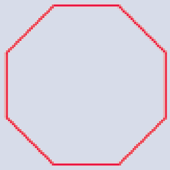
"Abstraction excites me; it is beautiful. Lots of people can enjoy the beauty of music, but not so many people can understand the beauty of abstract mathematics. It is a great pleasure seeing clever students learning to see its beauty."

Saxl was in New Zealand in 2011 as a visiting Maclaurin Fellow. ***JR***

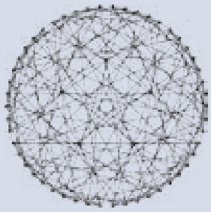
> **The mathematician's patterns, like the painter's or poet's, must be beautiful. The ideas, like the colours or the words, must fit together in a harmonious way. Beauty is the first test: There is no permanent place in the world for ugly mathematics.**
>
> *G. H. Hardy, 1877 – 1947*

Levels of abstraction

1 Shapes



Octagon



Hoffman-Singleton graph

2 Symmetries

Octagons have 16 different symmetries - eight rotational (first row) and eight reflectional symmetries (second row).

The **Hoffman-Singleton graph** on 50 nodes and 175 edges is the only regular graph of vertex degree 7, diameter 2, and girth 5.

3 Groups of symmetries

The set of the **octagon** movements, with composition as operation, forms the algebraic structure of a finite dihedral group D_8. Below is a group table for D_8, and other tables for some of its subgroups.

●	1	6	7	16	17	22	23
1	●	7	6	22	23	16	17
6	7	●	1	17	16	23	22
7	6	1	●	23	22	17	16
16	17	22	23	●	1	6	7
17	16	23	22	6	7	●	1
22	23	16	17	1	●	7	6
23	22	17	16	7	6	1	●

●	1	6	7
1	●	7	6
6	7	●	1
7	6	1	●

●	7	16	23
7	●	23	16
16	23	●	7
23	16	7	●

●	7	17	22
7	●	22	17
17	22	7	●
22	17	●	7

The symmetry group of the **Hoffman-Singleton graph** is of order 252,000.