UNITY
2011

Silverstream, Wellington
21 – 24 November 2011

Later mathematicians will regard set theory as a disease from which one has recovered.

Henri Poincaré

ROOM ALLOCATIONS

Timm Treskatis and Peter Jaksons

Agate Ponder-Sutton and Caludia Seibold

Josh Collins and Michael Snook

Peter Green and Daniel Turek

Edoardo Persichetti and Adam Ward

Ben Clark and Kyle Tate

James Dent and Nick Brettell

Sabariah Saharan and Valentina Baccetti

Mohammed Daher and Imran Khaliq

Michael Welsh and Melissa Welsh

Faezeh Frouzesh and Howida AlFran

Andreana Holowatyj and Penny Bilton

Time is the most undefinable yet paradoxical of things; the past is gone, the future is not come, and the present becomes the past even while we attempt to define it, and, like the flash of lightning, at once exists and expires.

Charles Caleb Colton

TIMETABLE

Tuesday 22 November

0850	Opening Rant
Session α: Chair -- Michael Snook	
0900	Peter Green
0925	Josh Collins
10 minute break	
1000	Ben Clark
1025	Mohammed Daher
1050	Morning Tea
Invited Talk -- Pure	
1115	Peter Donelan
1215	Conference Photo
1220	Lunch
Session β: Chair -- Lisa Woods	
1305	Michael Snook
1330	Adam Ward
10 minute break	
1405	Peter Jaksons
1430	Imran Khaliq
1455	Afternoon Tea
Session γ: Chair -- Josh Collins	
1520	Kyle Tate
1545	Daniel Turek
10 minute break	
1620	Andreana Holowatyj
1645	Edoardo Persichetti

Wednesday 23 November

Session δ: Chair -- Edoardo Persichetti	
0900	Timm Treskatis
0925	Michael Welsh
10 minute break	
1000	Lisa Woods
1025	Penny Bilton
1050	Morning Tea
Invited Talk -- Statistics	
1115	Richard Penny
1215	Lunch
Invited Talk -- Weta	
1305	Weta Workshops
Session ϵ: Chair -- Peter Jaksons	
1415	Sabariah Saharan
1440	Howida AlFran
1505	Afternoon Tea
Session ς: Chair -- Daniel Turek	
1530	Valentina Baccetti
1555	Nick Brettell
10 minute break	
1630	Claudia Seibold
1655	Agate Ponder-Sutton

Thursday 24 November

Session ϑ: Chair -- Agate Ponder-Sutton	
0900	Faezeh Frouzesh
0925	Melissa Welsh
10 minute break	
1000	James Dent
1025	Morning Tea
Invited Talk -- Applied	
1100	Michael Plank
1200	Prize Giving
1220	Lunch
1250	Departure

A mathematician is a scientist who can figure out anything except such simple things as squaring the circle and trisecting an angle.

Even Esar

INVITED TALKS

PETER DONELAN — PURE — VICTORIA UNIVERSITY

Singularities and Geometry in our Favourite Spaces

CURVES AND SURFACES in the plane and space generally have points where their geometry changes: inflection points, points of zero torsion, umbilic points. These can be characterised by their degree of contact with special sets: lines, planes, circles, and spheres. One way to make this precise to via height and distance functions whose level sets are these model sets. The key ingredient is the classification of singularities of such functions. There is no need to restrict to three dimensions nor to Euclidean space. Recent results tell us about the local and global geometry of surfaces in four dimensions and in spaces with indefinite metrics.

RICHARD PENNY — STATISTICS — STATS NZ

Statistics Includes Data, and Data is more than Numbers

THE WORD STATISTICIAN was first used at the beginning of the 19th century and was originally "one versed in or engaged in collecting and tabulating statistics". Statistics in this case meant data. Since then the data has appeared to become less and less a part of the profession of statistics. This is unfortunate as data is still the bedrock upon which statistics exists and many statistical analyses are wrecked by an under appreciation of this fact. As an unapologetic dataphile I want to talk about data. I will start by showing data is not just a set of numbers or answers in some computer file. From this starting point it can be seen it is important to understand how data is accumulated, what can go wrong and how to improve your chances of getting your data right. As data is often shared you need to be aware how you can prevent the creation and spread of your own Statistically Toxic Data, and how to protect yourself from another's STD.

MICHAEL PLANK APPLIED UNIVERSITY OF CANTERBURY

Size-structured Models of Marine Ecosystems and Fishing

THE FIRST RULE of the ocean is "never eat anything bigger than your head". Apart from that, most fish aren't too fussy about what they eat and a predator's diet is determined primarily by the size of the prey, rather than by its species. Furthermore, predators move steadily up the food chain as they grow larger, meaning that the traditional food web description of the ecosystem is not appropriate. In this talk, I will give an introduction to size-structured models of marine ecosystems based on the McKendrick-von Foerster equation. I will show that a model of this type predicts the remarkable "invariance of biomass" with respect to body size that is observed in real systems. I will also show some results on stability and the effects of different size-based fishing protocols.

Anyone who cannot cope with mathematics is not fully human. At best he is a tolerable subhuman who has learned to wear shoes, bathe, and not make messes in the house.

Robert Heinlein

APPLIED MATHEMATICS TALKS

VALENTINA BACCETTI

Matter Degrees of Freedom for a Quantum Theory of 3D Gravity

IN THIS TALK we will illustrate a possible way for coupling matter degrees of freedom to a "Spin Foam Model". Spin foams are a particular non-perturbative approach formulated to tackle the problem of a quantum theory of gravity. Specifically, we will use a supersymmetric spin foam model and show how to identify the fermionic degrees of freedom embedded within it.

JOSH COLLINS

Constructing Phylogenetic Networks

THE PROBLEM OF constructing evolutionary trees given some appropriate data is a problem with a long history in phylogenetics. However, it has become apparent that in many situations a tree structure is not appropriate to accurately model the evolutionary history of a set of organisms and thus attention is increasingly turning towards the problem of constructing networks for the same types of data. This talk will contain a comparison of such methods and an overview of the approach I am taking.

ANDREANA HOLOWATYJ

Characterising Moves in Hexagonal Regions of Hextile Knot Mosaics

A HEXAGONAL KNOT MOSAIC is a knot diagram that lies in a hexagonal grid in a particular way such that the intersection of the knot diagram and the

interior of one of the hexagons is one of 27 basic types. Within the knot mosaic there exist simply connected regions of finitely many hextiles known as hexagonal regions. We enumerate planar isotopy moves in n-hextile regions and explore methods used to condense the theorised infinite number of isotopy moves on a hexagonal grid. We then explain parameters that limit and determine the range of planar isotopy moves possible in cases of specific hexagonal regions.

Imran Khaliq

Extracting Winning Strategies in a Subclass of McNaughton Games

We investigate algorithms for extracting winning strategies in two-player games played on finite graphs. We focus on a special class of McNaughton games called update games. We present a procedure for extracting winning strategies in update games by constructing strategies explicitly. This is based on an algorithm that solves update games in quadratic time. We also show that solving update games with a bounded number of nondeterministic nodes takes linear time.

Agate Ponder-Sutton

Beetle-licious: Building a Better Weed Control Mode for Tradescantia Fluminensis Vell. (Commelinaceae)

Tradescantia fluminensis Vell. (Commelinaceae) is a concerning invasive weed within New Zealand. Branching process models; a stochastic method; are used to simulate multiple cases of individual, independent growth of T. fluminensis. A system of linear ordinary differential equations is used to check the validity of the simulations. The models are utilised to explore possible management strategies and ask "What can be known about the probabilities of controlling the growth of this plant, given the plausible ranges of branching rates and death rates?" The simulations show show that it will take a large proportion of tip death to control T. fluminensis. This may recommend other forms of control which may be more applicable.

In this talk, branching process analysis will be presented in conjunction with the simulation algorithm; simulation output will be compared to an ODE approximation.

Sabariah Saharan

Elitist Non-Dominated Sorting Genetic Algorithm (NGSA-II): An Application to Large Imbalanced Traffic Road Accident Data Set

Genetic algorithms (GAs) have been extensively studied by numerous researchers from various fields of study over the last decade. Genetic algorithms for clustering analysis and pattern recognition have been shown

to be very successful by many researchers. GAs only require minimal information of the problem and use robust search and optimisation techniques which make these algorithms more efficient than classical methods. For this research, we have implemented Non-Dominated Sorting Genetic Algorithm (NSGA-II) to a traffic accidents data set. We use multi-objective optimisation to find one or more feasible solutions. Our data set is in binary form which have highly imbalanced classes. The data relates to 26440 traffic accidents with information on their level of severity and accident-cause. The clustering task is to group the causes of accidents to K groups where there are no given group labels. We compare our results to Generic Clustering for Unknown K (GCUK).

Claudia Seibold

Fractals and Faults

Not all faults are visible – most of them are below the surface. Identifying fractal pattern can help determining the location of faults. Therefore it may have the potential to support the prediction of natural hazards such as earthquakes because of the relationship between earthquakes and faults.

A fractal has been defined as a rough or fragmented geometric shape that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole.

In the last decades the fractal view of faults has been supported by long-range correlation of various measurements in geophysical boreholes and by geometrical analysis of fault traces. But fractal patterns at faults seem to have breaks, which divide the spatial distribution into more than one band.

This study will examine how the fault system in the South Island of New Zealand can be described in fractal terms. Data will be compared with findings in Japan to determine differences and getting insights to improve the characterisation of faults.

Kyle Tate

Simplicial Approaches to Quantum Gravity

In this talk I will introduce how the construction of a simplex can be used to study both classical and quantum General Relativity. I will describe recent work in which the significant differences between simplices in Euclidean and Lorentzian signature have been utilised to formulate a new model of Lorentzian Quantum Gravity.

Timm Treskatis

Optimal Control of Burgers' Equation with Adaptive Moving Mesh Methods

The spatially one-dimensional Burgers' equation often serves as a touchstone for numerical methods, which have been developed to solve problems in computational fluid dynamics. As optimal control problems for the Navier-Stokes equations are becoming more and more affordable in terms of computational resources, much research has already been conducted on the subject of initial and boundary value control of the viscous Burgers' equation. However, these previous works avoid the numerically demanding solution of the PDEs involved, when the viscosity ν assumes realistically small values.

By implementing a suitably constructed adaptive moving mesh solver for Burgers' equation we extend the admissible range for the viscosity parameter. It can be shown that even for $\nu = 10^{-6}$, the algorithms converge within a reasonable amount of time while the solutions exhibit neither spurious oscillations nor extensive numerical diffusion.

Adam Ward

Spectral Theory, Scattering Theory and the Smilansky Conjecture

New Zealand is under attack! A nuclear missile has been launched and is targeted at Wellington. Modern nuclear missiles are designed to ascend into the atmosphere and then break up into several smaller warheads which then fall back down to earth on the target. Some of these individual warheads will contain nuclear material and others will be decoys. Obviously, if you were working the NZ military you would be interested to know which of these warheads contained nuclear material so that you could then fire your own missiles at them to destroy them before they detonated. The question is – how do we determine which of the warheads contain nuclear material?

Spectral theory is a branch of quantum mechanics and functional analysis that looks to describe and characterise the spectrum of self adjoin operators in Hilbert space (typically in $L_2(\mathbb{R}^n)$). If we make the assumption that the warhead is a smooth, compact impenetrable shell, and that each contains a single particle (which is either nuclear or non-nuclear) then the operator under investigation is the Dirichlet Laplacian i.e. if Ω_{int} represents the interior of the warhead and $\partial\Omega$ its boundary then

$$-\Delta_{\text{int}} = -\sum_{j=1}^{3} \frac{\partial^2}{\partial x_j^2}$$

$$D(-\Delta_{\text{int}}) = \left\{ f \in W_2^2(\Omega_{\text{int}}) \mid f|_{\partial\Omega} = 0 \right\}$$

Spectral analysis of this operator implies that the energy levels that the particles can take are discrete, i.e. the energy levels correspond to the eigenvalues of the Dirichlet Laplacian. The Smilansky conjecture connects these eigenvalues to the fix points of the scattering operator, which (in theory) would enable us to determine which of the warheads contain nuclear material.

Melissa Welsh

Rheumatic Fever Epidemiology and Modelling

Acute Rheumatic fever is a disease of particular interest in New Zealand, especially for individuals of Maori or Pacific Island ethnicity. Mathematical modelling can help use to better understand how the disease might behave and spread in the population. These models can also help us to predict how certain medical interventions and treatments may affect the epidemiology of Rheumatic fever in the population.

We will look at some simple models, using differential equations and stochastic variations, and how we can apply these models specifically to Rheumatic fever.

Lisa Woods

Mapping Multiple Quantitative Traits Using Structural Equation Models

Quantitative traits are continuous physical properties displayed by an organism, such as yield, which are influenced by regions of the genome known as quantitative trait loci (QTL). The identification and mapping of QTL is of interest to geneticists and breeders looking to select for particular traits.

Multiple trait mapping is useful as it examines the correlations between traits such that, unlike single trait analysis, it allows the user to fit more complex and accurate biological models. Many multiple trait mapping methods have been developed, for example: Seemingly Unrelated Regression, Structural Equation Modelling (SEM), composite interval mapping for multiple traits.

SEM has an advantage over other multiple trait methods in that it allows incorporation of the casual structure. Other methods only work to map QTL and test for pleiotropy, which incorporating correlations between traits. By incorporating the casual structure, direct and indirect QTL effects on each trait can be estimated, thus allowing more accurate inferences of the genetic architecture to be made.

In most analyses that use SEM, the casual structure is assumed to be known a priori. We will investigate the use of a Bayesian mixture SE to infer the casual structure, effects and location of QTL that influence multiple traits. We use simulated data to explore the accuracy of our method under various scenarios, including the effect of potentially confounding environmental factors.

If you ask mathematicians what they do, you always get the same answer. They think. They think about difficult and unusual problems. They do not think about ordinary problems: they just write down the answers.

M. Egrafov

PURE MATHEMATICS TALKS

Howida AlFran

Distinguishing Generalised Knot Groups by Counting Homomorphisms to Finite Groups

Kelly (1991) and Wada (1992) independently identified and defined the generalised knot groups (G_n). The square (SK) and granny (GK) knots are two of the most well-known distinct knots with isomorphic knot groups. Tuffley (2007) confirmed Lin and Nelson's (2006) conjecture that $G_n(SK)$ and $G_n(SK)$ were non-isomorphic by showing that they have different numbers of homomorphisms to suitably chosen finite groups. He concluded that more information about K is carried by generalised knot groups than the fundamental knot groups. Soon after, Nelson and Neumann (2008) showed that the 2-generalised knot group distinguishes knots up to reflection. The goal of this study is to show that the difference can be detected by counting homomorphisms into suitable finite groups. This study extends Tuffley's work to analogues of the square (SK) and granny (GK) knots formed from connect sums of other torus knots. It gives further information about generalised knot groups of connect sums of torus knots which differ only in their orientation.

Nick Brettell

Removing Elements from a Matroid Relative to a Minor and a Fixed Basis

A standard matrix representation of a matroid M represents M relative to a fixed basis B, where contracting elements of B and deleting elements of $E(M) - B$ corresponds to removing rows and columns, respectively. If M is 3-connected, it is often desirable to perform such a removal while maintaing 3-connectivity.

Whittle and Williams (2011) showed that provided M doesn't contain any 4-element fans, there are at least four distinct elements k_i, for $i \in \{1,2,3,4\}$, such that $\mathrm{si}(M/k_i)$ is 3-connected when $k_i \in B$, and $\mathrm{co}(M\backslash K_i)$ is 3-connected when $k_i \in E(M) - B$.

We show that, subject to a mild essential restriction, either there are at least two distinct elements that can be removed in this way and also retain a copy of a specified N-minor of M, or M has a 4-element fan with a specific labelling of basis elements.

We also investigate the structure of M when there are precisely two such elements.

Ben Clark

$U_{2,5}$-fragile matroids

The fundamental problem in matroid theory is to characterise the classes of matroids that arise naturally from matrices. The most commonly sought way to characterise these classes is via a list of excluded minors for the class, that is, a list of the smallest matroids that are not in the class. In this talk we discuss the existing proof techniques for finding excluded-minor characterisations, including the crucial concept of fragility. We then describe the structure of a class of matroids that are fragile with respect to the matroid $U_{2,5}$, and outline how this can be used as a step towards finding new excluded-minor characterisations.

Mohammed Daher

Dual Numbers and Invariant Theory in the Euclidean Group

The Euclidean group is a mathematical model of rigid body motion in space and so is used in applications in Physics and Engineering such as robotics. The dualisation of the special orthogonal group $SO(3)$ is isomorphic to the special Euclidean group $SE(3)$. In this talk I will talk about invariant theory of the special Euclidean group by using the principle of transference.

The principle of transference proposes that true statements about $SO(3)$ dualise to true statements about $SE(3)$.

James Dent

An Introduction to Constructive Reverse Mathematics

Constructive reverse mathematics (CRM) is a project in which non- and semi-constructive principles are classified in accordance with which other principles they imply or are implied by, relative to the system of Bishop-style constructive mathematics. Of particular interest are the various

versions of Brouwer's fan theorem and several antitheses of Specker's theorem (a characteristic result of Russian recursive mathematics): principles that belong in the hierarchy of fan theorems are semi-constructive in the sense that they are intuitionistically valid, and a result of Berger and Bridges shows that this is also true of the so-called anti-Specker properties.

I give an exposition of the current state of the CRM programme, showing how these principles are related to one another, and framing the scope for further investigation — particularly of certain principles at the weaker end of the spectrum.

Edoardo Persichetti

Cryptography: Mathematics of Crypts?

Despite the analogy that could seem strange, the two concepts are not so far away: Cryptography is in fact the science which studies the hiding the information. This talk focuses on a "new wave" branch in the discipline, an approach called Public-Key Cryptography, and showcases the efforts made towards important security and practicality results, with perspective applications in an eventual Post-Quantum scenario.

All welcome!

Michael Snook

The Trouble with Counting Bases

It is #P-complete to count the number of bases of a representable matroid over any fixed field. This result was claimed by Vertigan in 1991. However, no publication was produced and this result has somewhat fallen into Myth. We will discuss similar results and our progress towards remedying this situation.

Michael Welsh

Maximum-sized Golden-mean Matroids

A matroid is an axiomatic combinatorial structure that arises out of notions of dependence, such as linear dependence in matrices. Using linear independent sets of columns in a matrix, one can form the independent sets of a matroid. The rank of a matroid is the maximum size of all such independent sets. A matroid that can be represented over both $GF(4)$ and $GF(5)$ is a golden-mean matroid. A maximum-sized matroid is a matroid that holds the maximum size with respect to a class of matroids for a given rank. In this talk, we will look at the maximum-sized matroids in the class of golden-mean matroids. No knowledge of matroid theory will be required.

Statistics may be defined as "a body of methods for making wise decisions in the face of uncertainty."

W.A. Wallis

STATISTICS TALKS

Penny Bilton

A Classification Model for Poverty Mapping in Nepal

Poverty and malnutrition are serious problems in many Third World Countries. Elimination of extreme poverty by 2016 is the fist of the United Nation's Millennium Development Goals. To achieve this objective the World Bank, through the World Food Programme, expends billions of dollars of aid resources into the poorest countries in the world.

To ensure optimal distribution of aid resources a statistical model is employed to proved estimates, at low geographical levels, of poverty measures which are then incorporated into a poverty map. The current standard methodology for poverty mapping is the ELL method, named after its proposers, Elbers, Lanjouw and Lanjouw (2002). The technique utilises a linear regression model to predict household per capita expenditure, the variable of interest, from which are derived various poverty measures.

The purpose of this thesis topic is to investigate the application of classification techniques as an alternative methodology for poverty mapping. The challenges inherent in the proposed research include the need to incorporate complex survey design elements, weighting, stratification, and clustering. Another issue to be addressed is the aspect of small area estimation, since modelling at low levels introduces the problem of small sample sizes and imprecise estimates.

Many benefits could arise from a classification model for poverty mapping; direct prediction of a household as poor or non-poor; a reduction in the number of predictor variables required for modelling; simplicity, and the ability to represent interactions between predictor variables. The ultimate goal of the research is to find a more efficient modelling process, to facilitate better allocation of the billions of dollars currently spent on aid funding.

Faezeh Frouzesh

Evaluation of Starting Points for Optimisation in Mixture Models

While mixture models are frequently used to classify data, complications arise in the analysis of very large data sets. The use of mixture models in statistical models for data sets is increasing with heterogeneity and/or redundancy in the data. These models are based on likelihood. Maximum likelihood estimates of parameters are obtained using the EM algorithm. If poorly chosen, starting points for the optimisation through the likelihood surface may lead to only a local maximum, not a global maximum.

Shirley Pledger and Richard Arnold have a research programme using finite mixtures to analyse an $n \times p$ data matrix, which may have either binary data (0/1) or count data $(0, 1, 2, \ldots)$. They have a range of models, to do the analysis, and they propose to turn this into a R package. Thus, having the automated starting points is important in the R package.

In this research, we will evaluate and compare different methods of choosing starting points on simulated data. I have tried to find a procedure which will make intelligent choices of possible starting points, and fast evaluations of their usefulness.

Peter Green

Sequential Analysis and Computational Ecology

Computer simulations provide a "general solution" for analysing ecological models. In principle we should be able to answer questions about any well specified model by running computational experiments. In practise we are constrained by finite computational resources, especially if a model takes a long time to converge or if we are interested in generalising a model's behaviour over a multidimensional parameter space.

If we can emulate a simulation with an analytically tractable stochastic model, then we can overcome some of the limitations of computational experiments. In particular, we can calculate the distribution of outcomes for a process with a long running time. We can also evaluate parameter sensitivity without requiring extensive numerical coverage of the parameter space.

The Moran process from theoretical biology is a birth-death process used to model the spread of mutant genes in a population. This process can be used to calculate the probability that a beneficial mutation will establish itself throughout a population. The Moran process is the cumulative sum of random changes in the population state, and is therefore amenable to sequential analysis. The central result in sequential analysis, Wald's Fundamental Identity, can be used to calculate the probability that a random walk arrives at a given state, for example at the extinction or fixation of a beneficial mutation.

Our simple ecological model, implemented in NetLogo, can be emulated by a generalised Moran process. Sequential analysis provides a unified toolkit for analysing

that emulation. Our emulation allows us to demonstrate, under certain conditions, a sharp tipping point where very small advantages due to beneficial mutations are almost certain to become established in a population. We plan to extend this analysis to study possible explanations for the development of nervous systems during the Cambrian explosion.

Peter Jaksons

Ecological Sampling: Random, More Random, Totally Random

Despite the increasing need and demand for well designed long term sampling strategies of natural resources, research in this field of sampling theory is rather limited. An appropriately designed sampling strategy can be highly precise and accurate and will often require a smaller sample size compared to other sampling strategies, saving time and money.

There are three key characteristics a long term sampling strategy should possess. Most important the sampling strategy should have a well spread, spatially balanced design. Second, the sampling strategy should allow for unequal probability sampling. Finally, the design should be adaptive over time without losing its spatially balanced distribution. Only this last characteristic is a special requirement for long term sampling compared to cross sectional sampling. So far the gold standard for ecological sampling is the Generalised Randomised Tessellation Sampling methodology (GRTS). A new methodology, the Balanced Acceptance Sampling (BAP) based on the Halton sequence, meets all three of the key characteristics for long term sampling and is believed to perform better than GRTS.

In this talk I will start with an introduction to simple random sampling (SRS), GRTS, and BAP. Second, a comparison will be made of these three sampling strategies to test how will they perform for each of the three key characteristics a long term sampling model should have. Finally, the methodology and applications of the BAP design will be explained in further detail.

Daniel Turek

Model-Averaged Wald Confidence Intervals

In the frequentist setting, a model-averaged estimate of a parameter is calculated as the weighted sum of single-model estimates, using weights derived from an information criterion. A standard method for calculating a model-averaged confidence interval is to use a Wald interval centred around the model-averaged estimate. I propose a new method for construction of a model-averaged Wald confidence interval, based on the idea of model-averaging tail areas of the sampling distributions of the single-model estimates. Simulation study is used to compare performance of the new method and existing methods, in terms of coverage rate and interval width. The new method consistently outperforms existing model-averaged methods in terms of coverage, often requiring

little increase in the interval width. I also consider choice of model weights, and find that AIC weights are preferable to either AICc or BIC weights in terms of coverage.