



# The Statistical Society of Australia NEWS



## AND THE WINNER IS... AUSTRALIA!

The votes are in for the International Statistical Literacy Project, and our fledgling National Secondary Schools Poster Competition has produced a winner on the world stage in just its first year!

The Junior Division Winners from the 2014 Australian National Secondary Schools Poster Competition (Year 10 students from Lisarow High School, Central Coast, NSW) were announced at the International Statistical Institute's 60th World Statistics Congress in Brazil as the **Winners of the International Statistical Literacy Project – Junior Division**. The winners receive 450 Euros and a certificate. There were more than 20 countries competing, see <http://iase-web.org/islp/PosterCompetition2014-2015.php?p=Prizewinners>.

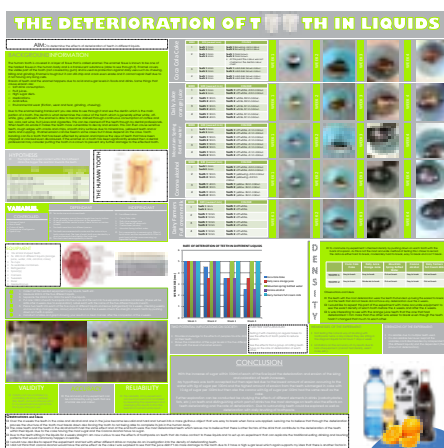
On 7th August, Peter attended a special assembly at Lisarow High School to award the students. Local media covered this event.

### In this issue

Editorial	2
Events	3
Meta-Analysis Workshop	6
National Secondary Schools Statistical Literacy Poster Competition	8
Canberra branch	11
NSW Branch	17
SA Branch	21
From the SSAI Office	24



The three winners with Peter Howley



Winning poster

To be recognised on the international stage is a phenomenal achievement. And in just our first year – the future looks bright! This is an outstanding result for Australia and the Central Coast!

Meanwhile, the national competition continues to grow. The pilot in the Hunter Region in 2014 saw 85 students engage in teams to complete 32 posters, this year there are **over 350 students engaged in over 140 posters from schools across NSW, ACT, Victoria, WA and Tasmania**.

If you would like to be involved please contact [peter.howley@newcastle.edu.au](mailto:peter.howley@newcastle.edu.au) or call 02 49215518.

**Peter Howley and Michael Martin**  
Chairs of Statistics Education

September 2015  
Issue 152

#### SSAI

PO Box 213, Belconnen ACT 2616  
We are located on the ground floor of  
ABS House, room GN 311.  
Phone 02 6251 3647  
Email [eo@statsoc.org.au](mailto:eo@statsoc.org.au)  
Website [www.statsoc.org.au](http://www.statsoc.org.au)

#### Editor

**Sonia Langford**

#### Correspondence

Please direct all editorial  
correspondence to Sonia Langford  
Email [eo@statsoc.org.au](mailto:eo@statsoc.org.au)

#### Disclaimer

The views of contributors to this  
Newsletter should not be attributed to  
the Statistical Society of Australia, Inc.

#### Subscriptions

The Newsletter of the Statistical  
Society of Australia is supplied free to  
all members of the society. Any others  
wishing to subscribe to the newsletter  
may do so at an annual cost of  
A\$30.00 for an issue of four numbers.

#### Advertising

Advertising will be carried in the  
Newsletter on any matters which  
the Editors feel are of interest to the  
members of the Society.  
For details of advertising rates, etc.  
contact the SSAI Executive Officer at  
[eo@statsoc.org.au](mailto:eo@statsoc.org.au)

**DEADLINE FOR NEXT NEWSLETTER**  
**10 November 2015**

## EDITORIAL

How exciting! We have a fantastic membership adding value in a variety of ways to our community. In this newsletter we share some of our members' activities and achievements. For instance, you have been taken along on the journey of development of the Secondary Schools Statistical Literacy Poster Competition, and the success of this young contest has been reaffirmed by recognition on the international stage. In addition, we celebrate the award of an Australian Research Council Laureate Fellowship to a former SSAI President Professor Kerrie Mengersen. At a branch level the selection includes learning about collaborations aimed at mining the "Internet of Everything" to useful ends, spatial modelling as a tool to identify geographical clusters of disease risk, and discovering how statisticians are contributing to improvements in Aboriginal health in South Australia. I hope you agree as confirmed by the articles in this newsletter, that statisticians are truly useful contributors to our society.

Just a brief update on a couple of general matters – the SSAI Executive will be searching for a second editor, and the newsletter survey will be in circulation before the December edition.

Once again, if you have any questions or constructive comments regarding the newsletter, feel free to contact us via [eo@statsoc.org.au](mailto:eo@statsoc.org.au).



With warm regards from,

**Sonia Langford**

### ACSPRI 2015 SPRING PROGRAM WILL BE HELD AT THE UNIVERSITY OF TECHNOLOGY, SYDNEY FROM SEPT 28 TO OCT 2

ACSPRI courses cater not only for researchers in the social and political sciences, but also in areas such as behavioural, health and medical sciences, policy research, education, economics, epidemiology, law, management, marketing, public relations and human resource management. Courses at each of our programs range from those that offer a basic grounding in qualitative and quantitative research methods to state of the art techniques for experienced researchers.

For more details about the courses on offer, visit [ACSPRI's Spring program](#) course page on our website or email: [info@acspr.org.au](mailto:info@acspr.org.au)

## EVENTS

### 2015 RSS CONFERENCE

4-10 September 2015, Exeter UK

### BIG DATA 2015

20-21 October 2015, Sydney

### FIRST RUSSIAN STATISTICAL CONGRESS

20-22 October 2015, Novosibirsk City, Russia

### 23RD CONFERENCE PROBABILITY AND STATISTICS IN THE ATMOSPHERIC SCIENCES

10-14 January 2016, New Orleans, USA

### XXVIII<sup>TH</sup> INTERNATIONAL BIOMETRIC CONFERENCE (IBC 2016)

10-15 July 2016, Victoria, BC Canada

### ECO-STATS '15: TECHNOLOGICAL ADVANCES BETWEEN ECOLOGY AND STATISTICS

8-10 December 2015, Sydney NSW

### 12TH GERMAN PROBABILITY AND STATISTICS DAYS 2016 – BOCHUMER STOCHASTIK-TAGE

1-4 March 2016, Bochum, Germany

### AUSTRALIAN STATISTICAL CONFERENCE 2016 (WEBSITE NOT YET AVAILABLE)

5-9 December 2016, Canberra

#### CALL FOR PAPER: INTERNATIONAL CONFERENCE FOR ESTABLISHMENT SURVEYS, 20-23 JUNE 2016

The Fifth International Conference on Establishment Surveys (ICES-V) will be held June 20-23, 2016 in Geneva, Switzerland. The Program Committee invites you to submit a proposal for an invited paper from June 1st, 2015 to September 30, 2015.

For more information on invited sessions, including session formats, suggestions for topics, criteria, instructions and a template for submitting proposals, please consult <http://www.portal-stat.admin.ch/ices5/invited-sessions/>.

## SECTION CHAIRS

### Bayesian Statistics

Chair: Scott Sisson  
[Scott.Sisson@unsw.edu.au](mailto:Scott.Sisson@unsw.edu.au)  
Assistant Chair: Jannah Baker  
[Jannah.Baker@qut.edu.au](mailto:Jannah.Baker@qut.edu.au)  
<http://www.statsoc.org.au/bayesian-statistics.htm>

### Environmental Statistics

Co-Chair: David Warton  
[dwarton@maths.unsw.edu.au](mailto:dwarton@maths.unsw.edu.au)  
Co-Chair: Jakub Stoklosa  
[j.stoklosa@unsw.edu.au](mailto:j.stoklosa@unsw.edu.au)  
Assistant Chair: Mayukh Samanta  
[Mayukh.Samanta@qut.edu.au](mailto:Mayukh.Samanta@qut.edu.au)  
<http://www.statsoc.org.au/environmental-statistics.htm>

### Social Sciences

Chair: Michele Haynes  
[M.Haynes@uq.edu.au](mailto:M.Haynes@uq.edu.au)  
<http://www.statsoc.org.au/social-statistics.htm>

### Statistical Education

Co-Chair: Michael Martin  
[Michael.Martin@anu.edu.au](mailto:Michael.Martin@anu.edu.au)  
Co-Chair: Peter Howley  
[Peter.Howley@newcastle.edu.au](mailto:Peter.Howley@newcastle.edu.au)  
<http://www.statsoc.org.au/statistical-education.htm>

### Surveys and Management

Chair: Stephen Horn  
[srthorn@effect.net.au](mailto:srthorn@effect.net.au)  
<http://www.statsoc.org.au/surveys-and-management.htm>

### Biostatistics

Co-Chair: Jake Olivier  
[j.olivier@unsw.edu.au](mailto:j.olivier@unsw.edu.au)  
Co-Chair: Kylie-Ann Mallitt  
[kmallitt@kirby.unsw.edu.au](mailto:kmallitt@kirby.unsw.edu.au)  
Assistant Chair: Nicholas Tierney  
[Nicholas.Tierney@gmail.com](mailto:Nicholas.Tierney@gmail.com)  
<http://www.statsoc.org.au/medical-statistics>

### Section for International Engagement

Mark Griffin  
[m.griffin@adasis-oz.com](mailto:m.griffin@adasis-oz.com)  
<http://www.statsoc.org.au/IntEngagementSection>

### Young Statisticians' Network

Rory Tarnow-Mordi  
[rtmordi@gmail.com](mailto:rtmordi@gmail.com)  
<http://www.statsoc.org.au/about-young-stats.htm>

Further contact details for Society Secretaries and Section Chairs can be obtained by contacting the Society on (02) 6251 3647

## SSAI CENTRAL COUNCIL

### Executive Committee

President: John Henstridge  
Secretary: Doug Shaw  
[secretary@statsoc.org.au](mailto:secretary@statsoc.org.au)

### Branch Presidents and Branch Secretaries

#### Canberra

President: Robert Clark  
Secretary: Warren Müller  
[secretary.canberrabranch@statsoc.org.au](mailto:secretary.canberrabranch@statsoc.org.au)

#### New South Wales

President: Louise Ryan  
Secretary: Ryan Defina  
[secretary.nswbranch@statsoc.org.au](mailto:secretary.nswbranch@statsoc.org.au)

#### Queensland

President: Peter Baker  
Secretary: Dimitrios Vagenas  
[dimitrios.vagenas@qut.edu.au](mailto:dimitrios.vagenas@qut.edu.au)

#### South Australia

President: Lisa Yelland  
Secretary: Paul Sutcliffe  
[sutters@bigpond.net.au](mailto:sutters@bigpond.net.au)

#### Victoria

President: Luke Prendergast  
Secretary: Sandy Clarke  
[vic.branch@statsoc.org.au](mailto:vic.branch@statsoc.org.au)

#### Western Australia

President: Alethea Rea  
Secretary: Ryan Admiraal  
[R.Admiraal@murdoch.edu.au](mailto:R.Admiraal@murdoch.edu.au)

## ARC'S AUSTRALIAN LAUREATE FELLOWSHIPS ANNOUNCED

The Minister for Education and Training, the Hon. Christopher Pyne MP, recently announced 15 outstanding recipients of new Australian Laureate Fellowships, funded through the Australian Research Council (ARC).

The Fellows will receive a total of \$42 million over the next five years and will commence research programs exploring fields including: harnessing intellectual property to build food security; translating 'big data' to meet challenges in industry, environment and health; and exploring a new 'Pharming' industry that uses plants to deliver medicine.

The SSAI is proud to advise that one of the Australian Laureate Fellows is former SSAI President Professor Kerrie Mengersen. Kerrie's project title is "Bayesian learning for decision making in the big data era" and her fellowship project aims to develop new techniques in evidence-based learning and decision-making. Big data has arrived, and with it a huge global demand for statistical knowledge and skills to analyse these data for improved learning and decision-making. Kerrie's project will seek to address this need by creating a step-change in knowledge in Bayesian statistics, and translating this knowledge to real-world challenges in industry, environment and health. The new, big data, statistical analysts trained through the project could also create much needed capacity at national and international levels.

A full list of new Australian Laureate Fellows and their project details are [available here](#) and further details are also available from the [ARC's Funding Announcements web page](#).

For more information about the Australian Laureate Fellowships scheme please visit the [ARC website](#).

### SAVE THE DATE!

#### AUSTRALIAN STATISTICAL CONFERENCE 2016

**in conjunction with the Australasian Data Mining Conference (AusDM)  
and the 19th Australian Conference on Teaching Statistics (OZCOTS)**

5th – 9th December 2016  
Hotel Realm, Canberra

#### **Big Data: Mining, Analyzing and Teaching**

Join delegates from all areas of statistics, data mining and teaching to discuss, network and learn. Develop and share knowledge and expertise with world class Australian and International colleagues. The Conference will provide an excellent opportunity to be involved with presentations on a wide range of topics recognising the role that statistics and data mining play in all aspects of the modern life.

Watch this space for more information:

<http://www.statsoc.org.au/events/ssai-events/australian-statistical-conference-2016/>

## META-ANALYSIS WORKSHOP

The University of Southern Queensland's (USQ) School of Agricultural, Computational, and Environmental Sciences and the School's Division of Research and Innovation, hosted a workshop on 'Statistical Meta-Analysis with Applications' in collaboration with the SSAI.

This workshop was held at the newly acquired Ipswich Campus of USQ from 16-17 June 2015. The two invited presenters were Professor Bimal Sinha from the University of Maryland, Baltimore County, USA and A/Professor Suhail Doi from the Australian National University, Canberra, Australia.

The organiser of the workshop, Professor Shahjahan Khan of USQ, opened the inaugural session with a brief introduction on the role of meta-analysis within the systematic review and evidence based decision making process and the use of statistical methods in synthesising data from independent studies. He also highlighted applications of meta-analysis in many fields of medicine, agriculture, education and business, and discussed some the issues related to methods of allocation of weights under various models in the estimation of the common effect size of meta-analysis.

Professor Bimal K Sinha started with some motivating real life examples of data leading to the definition of measures of various effect sizes for continuous and binary outcome variables. He covered all commonly used estimators of common effect size and discussed their variance estimators and confidence intervals. He also discussed inference about the common mean of univariate normal distribution, publication bias, vote counting procedures, and utilising the random effects (RE) model and meta-regression to deal with the heterogeneity issue.

Professor Suhail Doi highlighted the main purpose of meta-analysis and focused on some of the problems inherent with conventional statistical meta-analysis, especially the issue of unfair redistribution of more weights to smaller studies under the random effects model. Under the title of '*Recent advances in the methodology of statistical meta-analysis*', he presented the inverse variance heterogeneity (IVhet) estimator as an alternative to the RE model estimator, and introduced the quality effect (QE) model estimator as discussed in his recent publications. Through extensive simulation examples he demonstrated the advantages and appropriateness of the new estimators.

The twelve workshop participants attended from government departments, industry and academia, from Queensland and other parts of Australia. They were very happy with the presentations and management of the event, and thanked USQ and SSAI for organising this valuable workshop.

### Professor Shahjahan Khan

University of Southern Queensland

Participants and presenters of the meta-analysis Workshop at USQ, Australia





## XXVIII<sup>th</sup> International Biometric Conference VICTORIA CONVENTION CENTRE, JULY 10 – 15, 2016



*“I warmly invite you to the XVIII<sup>th</sup> International Biometrics Conference in Victoria, Canada, traditional territory of the Lekwungen First Nations, and capital city of British Columbia. Victoria is located on the southern tip of Vancouver Island, off Canada’s Pacific Coast. The conference will be held at the Victoria Conference Centre, adjacent to the famous Empress Hotel. Victoria is named the City of Gardens and has easy access to recreational activities such as kayaking, whale watching, hiking, and much more. The temperate climate and relaxed island lifestyle should make for a memorable 28<sup>th</sup> IBC.”*

– LAURA COWEN, UNIVERSITY OF VICTORIA, LOCAL ORGANIZING CHAIR

### SCIENTIFIC PROGRAMME

- Opening Ceremony & Presidential Address
- Invited Oral Sessions
- Contributed Oral and Poster Sessions
- Young Statisticians Showcase Session
- Biometrics and JABES Showcase Session
- Short courses

### SOCIAL PROGRAMME

SUNDAY JULY, 10  
MONDAY JULY, 11  
TUESDAY JULY, 12  
WEDNESDAY JULY, 13  
THURSDAY JULY, 14

**Welcome Reception**  
**Young Statisticians Mixer**  
**Regional Officers’ Reception**  
**Range of social excursions**  
**Gala Cultural Event**



For constantly updated information,  
please see the website [www.biometricsociety.org](http://www.biometricsociety.org)



# COME AND JOIN IN THIS NATIONAL COMPETITION

## National Secondary Schools Statistical Literacy Poster Competition

In previous SSAI Newsletters we wrote about the successful pilot of the National Secondary Schools Poster competition in the Hunter Region in 2014. In brief, this competition encourages teams of 2 to 5 secondary school students, to undertake data based investigations on a practical research question that interests them, then creatively report their results in a poster format. The project's aims and the support available to engage in this activity are described on the website ([www.ssaipostercomp.info/](http://www.ssaipostercomp.info/)). Last year, we had 85 students produce 32 posters. Examples of their many and varied entries are available at [www.ssaipostercomp.info/winners.html](http://www.ssaipostercomp.info/winners.html). Two examples are provided below.

### How to Make 2 Minute Noodles in 3 Minutes!!!


**Hypothesis:** Adding salt to water will cause the water to take longer to boil  
**Aim:** To find out how salt affects the boiling time of water  
**Independent Variable:** The amount of salt that we put in the water  
**Dependent Variable:** The time it takes for the water to boil  
**Controlled Variables:**

- The amount of water
- The temperature of the stove
- The same equipment each time
- The amount of salt needs to be accurate

**Equipment:**

- Timer
- Salt
- Measuring equipment
- Stove
- Saucepan
- Gloves
- Safety Glasses
- Water

**How does salt affect the boiling time of water?**

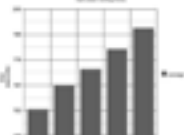


**Was my results valid?**  
 Yes our results were valid because we controlled the variables. We used the same amount of water each time. We put the stove on the same heat each time. The same saucepan was used. There was only one independent variable, the amount of salt. These factors make our results valid.

**Conclusion:**  
 It was found that adding salt to water increases the boiling time of water. The more salt you add, the higher the boiling temperature becomes, therefore the solution takes a longer period of time to boil.

**Was our hypothesis supported by this experiment?**  
 Our hypothesis was supported by the experiment. We predicted that adding salt to the water will increase the boil time. Considering the experiment proved that our hypothesis was correct. You can easily figure this out by reading our results table and graph.

Salt (teaspoons)	Time - test 1 (minutes)	Time - test 2 (minutes)	Average
0	2:35	2:30	2:32.5
0.5	2:48	2:45	2:46
1	2:52	2:51	2:51.5
1.5	2:58	3:03	3:00.5
2	3:10	3:12	3:11

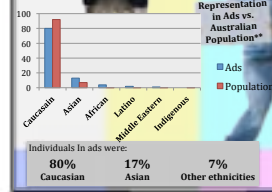


### What are the representations of Ethnicities in Primetime TV Advertisements?

The objectives were to answer these questions:  
**What ethnicities are represented and how frequently?**  
**How do these representations compare to the national population?**  
 The answers to these questions would prove or disprove the Hypothesis that ads would have an over representation of Caucasian individuals with other ethnic minorities underrepresented

**Methodology:** Channels Nine, Seven and Ten were reviewed between 6-9pm over the course of 14 days. Five ads were assessed each night based on randomly selected minutes of the 180 minute sample size. This primary data was analysed.

**Representation in Ads vs. Australian Population\*\***



**Individuals in ads were:**  
 80% Caucasian, 17% Asian, 7% Other ethnicities

**Research disproved the hypothesis with Caucasians underrepresented compared to the Australian population and ethnicities including Asian, African and Latino overrepresented**

**17% Non-Caucasians had speaking roles in primetime ads**

**NEXT TIME:** Larger sample size, manage time constraints

\*Primetime defined between the hours of 6pm and 9pm. TV Channels defined as commercial stations including Nine, Seven, and Ten  
 \*\* Population according to the ABS and the Australian Census 2011

The great news is that we have expanded nationally this year and attracted sponsorship from SAS and the Teachers' Mutual Bank. **We presently have over 350 students engaging in over 140 projects from schools across NSW, ACT, Victoria, WA and Tasmania.** Considering the competition began in its current form this time last year, we are delighted with such a positive start and incredible growth.

We wish to establish multiple 'coordinator sites' around Australia, each replicating the initial Hunter Region experience, including a Poster Display and Awards Night. The 2014 pilot helped establish processes and materials that we would like to see utilised elsewhere as we expand the competition.

It would be nice to hear that there are others who think this competition is a good idea and/or would like to see how they may get involved. Having a local friendly and interested person as a point of contact for schools will be invaluable to the expansion of the competition. So the questions for you are:

- Do you have some interest in helping create the next generation of Statisticians, working with Schools and Teachers, developing interest in the types of activities you enjoy, and helping to arrest the concerns of declining interest in mathematics, statistics and the sciences?

> Continued on next page



- Would you like to know more about what is involved in coordinating locally as part of the national competition?

Please contact [peter.howley@newcastle.edu.au](mailto:peter.howley@newcastle.edu.au) or phone 02 49 215518 to discuss.

I am pleased to advise that the CSIRO's Scientists-and-Mathematicians-In-Schools Coordinator has offered their program's involvement in the competition and will provide personnel to act as project facilitators within schools where possible. We are also being promoted through the Mathematical Associations of NSW, Tasmania and WA.

In 2016, we will be expanding the competition to Primary schools. In the meantime, primary schools who wish to be involved may do so in the lowest year division (Years 7-8).

So please contact [peter.howley@newcastle.edu.au](mailto:peter.howley@newcastle.edu.au) or phone 02 49 215518 to discuss.

**Peter Howley** and **Michael Martin**  
Chairs of Statistics Education

## 35% SSAI Member Discount Promotion with Wiley

**WILEY**

**SSAI members receive a special discount of 35% on online purchases with Wiley or Wiley-Blackwell (<http://www.wiley.com/statistics>). A discount of 25% applies to textbooks. This offer excludes school books.**

To activate this benefit, please go to <http://au.wiley.com/WileyCDA/Section/id-410891.html>. Register on the SSAI/Wiley Landing Page and from then on you can purchase Wiley books at the 35% discount *without use of a Promotional Code*. For those members outside of Australia and New Zealand, please use the Promo Code SDP92 at checkout to activate the discount. Members who registered on the Wiley site before the SSAI link was available may encounter problems when trying to get the discount. To get the new discount these members unfortunately need to register using another email address.

Toll free phone (from within Australia only) 1800 777 474  
Toll free phone (from New Zealand only) 0800 448 200  
Other overseas phone + 61 7 3354 8455  
Email [custservice@johnwiley.com.au](mailto:custservice@johnwiley.com.au)

# GPSD

**12th German Probability and Statistics Days**  
*Stochastiktage Bochum [March 1 – 4, 2016]*

## Sections and Invited Speakers

- 1 Stochastic Analysis**  
Davar Khoshnevisan (USA)
- 2 Spatial Stochastics**  
Jean-Francois Coeurjolly (France)
- 3 Limit Theorems, Large Deviations, Extremes**  
Erwin Bolthausen (Switzerland)
- 4 Finance, Insurance, Risk: Modeling**  
Thomas Møller (Denmark)
- 5 Stochastics in Physics and Biology**  
Christophe Garban (France)
- 6 Stochastic Processes**  
Ben Hambly (UK)
- 7 Time Series**  
Piotr Kokoszka (USA)
- 8 Data Analysis and Computational Statistics**  
Finn Lindgren (UK)
- 9 Nonparametric and Asymptotic Statistics**  
Alessandro Rinaldo (USA)
- 10 Statistics of Stochastic Processes**  
Emmanuel Gobet (France)
- 11 High Dimensional Inference**  
Valen Johnson (USA)
- 12 Finance, Insurance, Risk: Statistics**  
Jose E. Figueroa-Lopez (USA)

## Plenary Speakers

- Sandrine Dudoit**  
Berkeley School of Public Health (USA)
- Laszlo Erdős**  
Institute of Science and Technology Austria (Austria)
- Martin Hairer**  
The University of Warwick (UK)
- Iain Johnstone**  
Stanford University (USA)
- Walter Schachermayer**  
University of Vienna (Austria)

## Congress Coordinator

**Prof. Dr. Herold Dehling**  
Fakultät für Mathematik  
Ruhr-Universität Bochum

## Congress Office

**event lab. GmbH**  
Dufourstraße 15  
D - 04107 Leipzig  
Phone: +49 (0) 341 24 05 96 75  
Fax: +49 (0) 341 24 05 96 51

[www.gpsd-2016.de](http://www.gpsd-2016.de)

## CANBERRA BRANCH

### Making inferences and predictions based on data subject to selection bias

Dr Ray Lindsay gave the presidential address at the 2015 March meeting of the Canberra Branch, marking the end of his successful two year tenure as Branch President.

He highlighted that "big data" is often collected under uncontrolled conditions or is otherwise subject to selection bias. So, his presentation examined circumstances where such sample data can and cannot be used to make inferences or build models for the full population. Additionally, he noted that some models are built with the intention to make predictions for populations that are different in some key aspects, essentially a form of extrapolation. Ray based his comments partly on some past and current work at the Australian Tax Office, whilst noting that all interpretations were his alone. He used as examples during his talk some moderately big datasets i.e. of the order of 10 million observations and 1000 variables.

Ray observed that much of the research on big data analysis is not in the traditional statistical literature, with key papers from the International Conference of Machine Learning (ICML), and artificial learning and econometric journals. He then reviewed a range of alternatives for dealing with selection bias, including:

- ridge regression, applied to surveys in a seminal 1984 paper by Bardsley and Chambers;
- Heckman's approaches for self-selection problems, for which he won the 2000 Nobel Prize in Economics;
- Zadrozny's (of IBM) work on learning and evaluating classifiers under sample selection bias, published in ICML in 2004;
- Gretton's 2007 correction of selection bias by distribution matching; and
- Pearl's work, which was couched in the language of graph theory.

He concluded that real world problems of this kind are challenging. Data is typically not very clean, and traditional (and even modern) algorithms often do not properly deal with missing values.

The talk provoked a lively discussion.

**Robert Clark**

## Mining the Internet of Everything

The June meeting of the Canberra Branch was held in conjunction with the Institute of Analytics Professionals Australia (IAPA). Dr Ric Clarke, of the Australian Bureau of Statistics (ABS), spoke on 'Mining the Internet of Everything (IoE) – Official Statistics in the Information Age'.

He described the Internet of Everything as a unified web of information on people (the social web), places (the geospatial web) and things (the sensor web), with 'Big Data' as the collective residue of interactions in the IoE. He outlined how at the ABS a multidisciplinary team of specialists in mathematics, econometrics, computer science and information management are working to extract value from the IoE. For instance, they are developing a set of prototypes to represent data as a network of entities and relationships, to describe the semantics of data in a machine-interpretable form, to enable machines to reason on the data to derive new insights, to extract and transform the content of multi-structured data sets and to embed advanced visualisation in information systems. The ABS is exploring the application of these methods to:

- satellite sensor data to categorise land use and type of crop,
- mobile device location data to produce day time as well as usual resident population counts,
- Point-of-Sale payment data for improved coverage and efficiency of the Consumer Price Index,
- Smart meter data to understand the dynamics of household energy consumption), and
- administrative data from multiple sources e.g. on businesses and their employees.

There was a healthy attendance of both IAPA and SSAI members, and plenty of discussion before and after the talk.

Both groups are positive about holding other joint events in the future.

### Bill Gross



Dr Ric Clarke  
presenting at the  
June Canberra  
branch meeting

# BAYES ON THE BEACH

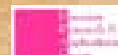
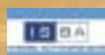
December 7<sup>th</sup> – 9<sup>th</sup> 2015

Surfers Paradise, Gold Coast

[botb2015.wordpress.com](http://botb2015.wordpress.com)

Bayes on the Beach is an intentionally small international forum for discussing and exploring developments in Bayesian Statistics and its applications. This three day conference will take place in the world-famous Surfers Paradise on the Gold Coast and is designed to have a diverse format including presentations, contributed sessions, workshops, a poster session and tutorials.

This conference is supported by the Bayesian Statistics Section of the Statistical Society of Australia, Inc. (SSAI), the CSIRO; the Australasian chapter of the International Society for Bayesian Analysis (ISBA); Queensland University of Technology (QUT) and the ARC Centre of Excellence for Mathematical & Statistical Frontiers (ACEMS).



## Identification and Definition of Lexically Ambiguous Words in Statistics

Is 'significant' a significant word in statistics? What does it really mean, and in what context? Such issues are being addressed by Alice Richardson of the University of Canberra (UC), in collaboration with Peter Dunn and Rene Hutchins of the University of the Sunshine Coast (USC). Alice presented their work at the May meeting of the Canberra Branch.

Teaching statistics used to be concentrated on presenting relevant mathematics and hoping the students knew what to do with it. But now it is just as important to make students aware of the language and concepts of statistics, so they can understand and correctly use them. Inevitably there are many words that can mean different things in a statistical context to everyday usage, or even in different applications in statistics. These are the lexically ambiguous words.

It isn't only 'significant' that is so misused. How about 'random', or 'correlation' or 'distribution'? There are many such words, in fact Kaplan and colleagues have identified 36 words they thought students may find lexically ambiguous. So to research the relative levels of difficulty various words would provide for students, Alice and her colleagues ran an experiment. In 2013, first year statistics students at both USC (924 students; 10 tutors) and UC (256 students; 8 tutors), were provided in week 1 with an extract of a journal article highlighting certain words, and asked to define them in context. For example:

"Results of the multiple regression analysis revealed a negative association between average speed in training with race time. There was a *significant* positive association between 100km race time and personal best time in a marathon."

Different words were selected in different tutorials, with each student seeing two extracts with up to five words to define. All this was embedded in a complex experimental design. Student definitions were allocated to one of six categories: 1) statistical and correct; 2) statistical, but not correct; 3) ambiguous; 4) non-statistically correct; 5) non-statistical and not correct; 6) no response. The activity was repeated at the end of the semester to look for improvements.

Some words were well-handled, for example 'mean' was defined statistical and correct 83% at the beginning and 94% at the end, and 'correlation' was respectively 60% and 73%. However, 'significant' was the real outlier going from 1.7% to only 5.4%. Students knew that 'significant' had a special meaning in statistics, but could not articulate exactly what it was. At least the numbers who provided a statistical (but incorrect) definition jumped from 2.9% at the beginning to 20.5% at the end.

Whilst tutors generally do identify words that students find lexically ambiguous, and students do improve after instruction, the need for some specific vocabulary-learning activities was recognised. Alice described four such activities that have been used and which were generally well-received:

1. Word find – finding words in rectangular arrays of letters,
2. Crosswords – inserting the correct words into a crossword based on clues,
2. Gap-fill – inserting the correct word into a gap in a sentence, and
4. Ordering – contests where students are presented with strips containing statistical activities, for example 'define the parameter', 'calculate test

> Continued on next page

statistic' etc. and then have to place them in the correct order in a grid with categories 'State', 'Plan', 'Solve' and 'Conclude'. The first to correctly fill the grid is the winner.

Whilst Word find was an awareness-raising activity, summaries of student assessments showed that Crosswords and Gap-fill were generally seen as fun and helpful. The Ordering activity was enjoyed greatly by students.

So as a practitioner who has to grapple with statistical words, use them in scientific work, and explain them to non-statistician collaborators, I was pleased to hear about innovative approaches to recognising difficult words and concepts, and teaching the next generation of statisticians the language, not just the mathematics. Alice and her colleagues found that both context and tutors mattered. Using these teaching techniques is having a significant impact on teaching!

**Warren Müller**

#### References

Richardson, A.M., Dunn, P.K. & Hutchins, R. (2013). Identification and definition of lexically ambiguous words in statistics by tutors and students. *International Journal of Mathematical Education in Science and Technology*.

Richardson, A.M., Dunn, P.K. & Hutchins, R. (2013). The impact of tutor, extract and word on the correct definition of lexically ambiguous words in statistics. *Australian Conference on Science and Mathematics Education*.

Kaplan et al. (2009). Lexical ambiguity in statistics: what do students know about the words association, average, confidence, random and spread? *Journal of Statistics Education*.

## STATISTICAL JOB BOARD NOW AVAILABLE!

Looking for a change? [Create a Job Seeker account](#), upload your resume or simply create a job alert and let the job find you.

Not looking for a job but offering one? Simply [create an Employer account](#) and see where it will lead you!


Best of luck finding top talent...or that dream job!

[http://careers.statsoc.org.au/home/index.cfm?site\\_id=18859](http://careers.statsoc.org.au/home/index.cfm?site_id=18859) (Job Board)



# Bio INFO Summer 15

**THE UNIVERSITY OF SYDNEY**  
7 - 11 DECEMBER 2015



**REGISTER NOW:**  
[WWW.AMSI.ORG.AU/BIS](http://WWW.AMSI.ORG.AU/BIS)

Research  
Australian Mathematical Sciences Institute




## “

Andrea Baratta,  
Managing Director at Epsilon Security

*[The intern] also gained valuable industry experience – she was able to solve problems and produce tangible results that have an actual business context.*

**AMSIIntern**  
[www.amsiintern.org.au](http://www.amsiintern.org.au)

## STATISTICALLY SAFER

Before stepping down as Chief Commissioner of Victoria Police, Ken Lay urged a move away from traditional policing to more sophisticated approaches. As Victorian lifestyles change and demographics shift we have seen rapid growth in high-density dwellings. Sixty-three per cent of offences recorded by Victoria Police in 2014 were property-related. So, what should home buyers and renters be looking for when house hunting?

*“Security systems have become a very important and integral part of both residential and commercial living,”* says Andrea Baratta, Managing Director at Epsilon Security.

*“We have identified a gap in how the security of these buildings is managed,”* Andrea says. Andrea proposes a global risk measure be given to all new buildings. He believes this as a good starting point for residents to understand how their property compares with other properties and will allow them to identify key risks and implement a solution to improve their safety.

Statisticians from the University of Melbourne are using Victorian crime data along with other measurable variables to evaluate the safety of your next apartment.

*“The AMSI Intern program enabled us to collaborate with Dr Davide Ferrari and his postgrad student Puxue Qiao to devise a statistically sound model for this project. Puxue also gained valuable industry experience – she was able to solve problems and produce tangible results that have an actual business context,”* says Andrea. *“Something not nearly enough Australian students have the opportunity to do.”*



## NSW BRANCH

### NSW Branch May Meeting: Craig Anderson

The NSW Branch had their May monthly meeting on May 26, 2015. We were delighted to hear from Dr Craig Anderson, who gave a talk on his main research area, disease mapping in spatial epidemiology. Craig is a recent graduate from the University of Glasgow, whose main research focus is on spatial and spatio-temporal modelling of disease risk. He is currently a postdoctoral research fellow working with Professor Louise Ryan at the University of Technology, Sydney NSW.

Craig began his talk by presenting a historically important example i.e. the most intense outbreak of cholera in the Soho district of London, in 1854. Physician John Snow identified the source of the outbreak as the public water pump on Broad Street, and discovered that cholera was spread by contaminated water, rather than polluted air, which was the then dominant "miasma theory". He produced one of the first ever "disease maps" and this case is now known as a classic study in epidemiology which uses extensive spatial analysis. For those of us who had no background in spatial modelling, this was a fascinating example for us to see its importance in the context of disease mapping.

Craig then proceeded to present a case study of respiratory hospital admissions in Greater Glasgow and Clyde Health Board, for which various spatial modelling methods were applied. To help the audience ease into the technical aspects, the basic model for disease risk was first discussed. It was a Poisson GLM incorporating a random effect modelled using a conditional autoregressive prior, which led to the Leroux model. Problems with the model were then identified, namely constant spatial smoothness across the study region. Thus, the objective was to find a method which allowed for more flexible smoothing, which was the core of Craig's talk.

He then proposed two approaches which use clustering techniques to allow for discontinuities in the spatial structure, namely the fixed effect model and the random effect model. A rather technical issue was fitting the fixed effect model using Integrated Nested Laplace Approximation as a computationally efficient tool for approximate Bayesian inference, which Craig briefly touched on. Comparison between the two proposed models was made in terms of their advantages, disadvantages and different aims. The most interesting aspect of the talk was to see the similarities and differences between the models, in terms of their ability to pick out geographical clusters of disease risk, when displayed on the disease map of the Glasgow case study. Craig concluded his talk by suggesting some potential future work.

Overall, Craig delivered a very interesting and informative presentation on the fascinating topic of spatial modelling, with just the right amount of technicality, real application and jokes to keep the audience engaged and entertained.

**Joanna Wang**



Dr Craig  
Anderson

## NSW Branch June Meeting: Dr Patricia Menéndez

The guest speaker at the June NSW SSAI meeting was Dr Patricia Menéndez from the NSW Bureau of Crime Statistics and Research. Dr Menéndez presented the methodology and early results from a recent study on the effect of the 2014 lockout laws in Sydney and within NSW.

In January 2014, after two alcohol related deaths in Kings Cross and amid growing pressure from the electorate and the media, the NSW government passed into law in a special sitting the wide ranging Sydney lockout laws. These laws did face some opposition, for instance at the time a Greens MP stated: "We don't believe there's evidence to justify what they're doing".

This ostensibly politico-social experiment is easily cast as a statistical one. "Treatments" in the legislation included: a 1:30 am lockout, 3:00 am last drinks, a freeze on liquor licenses, banning orders on trouble makers, a new risk based licensing fee, a suspension of on-line training for security staff, a state-wide ban on take-away liquor sales after 10:00 pm, as well as increased fines and minimum sentencing. Over one year after the law's initiation the following hypotheses were of interest:

H1 - Have the laws lowered the incidence of assault in Kings Cross and the CBD?

H2 - Have the laws increased the incidence of assault in neighbouring areas?

H3 - Is there evidence of 'displacement of crime' to other areas or 'diffusion of benefits'?

For those not familiar with policing and crime statistics, when a police intervention targets a problem in a specific area, but not its root causes, crime can simply relocate to another, untargeted area in a process known as displacement. The counterpart of displacement is diffusion of benefits, whereby other areas outside of the targeted intervention also share the benefits.

The available data were geo-coded monthly counts of assaults in the 7 months following the laws' introduction. As might be expected, analysing these data demanded more than a simple Poisson or negative binomial regression. The data show in some cases strong seasonal features along with temporal auto-correlations, as well as the obligatory instances of low cell counts. Time series structural models for count data were used to decompose the time series into seasonal variation and the underlying trend, as well as to estimate the effect of the lockout laws while accounting for the temporal dependencies. AIC (Akaike Information Criterion) based model selection was used to rank 3 competing models encapsulating different potential responses:

M1 - An immediate but transitory effect;

M2 - An immediate and permanent change;

M3 - A slow but steady change.

The results displayed a downward trend in violence in Kings Cross before the lockout and a subsequent dramatic drop immediately afterwards. The same prior downward trend was observed in the CBD, with a further decrease post lockouts. A 9% drop in assaults was noted across the entire state. There was no evidence of a displacement of alcohol related violence to proximal and distal areas in the 7month post lockout data.

> Continued on next page

These initial results indicate the lockout has served its political objective. However, to statisticians the results invite further conjecture: How did the lockout work? Were there fewer assaults simply because there were fewer hours to enact an assault or fewer people going out? Are complex knock-on effects now in play such as drinking at home and more unobserved domestic (as opposed to public) assaults? Are there more profound definitions of success to be analysed? The Bureau of Crime Statistics and Research plans to gather more information and utilise richer sources of data to attack these more complex issues.

This subject is particularly topical in Sydney at the moment and the lecture was extremely well attended. We would like to thank Dr Menéndez for an entertaining presentation and a fascinating introduction to crime statistics for most of the audience.

**Richard Walton**



Dr Patricia Menéndez presenting to the NSW branch



NSW branch discussing Dr Menéndez's presentation



# AMSI 16 Summer School IN THE MATHEMATICAL SCIENCES

**RMIT University**  
4 - 29 JANUARY 2016



**REGISTER NOW:**

[WWW.AMSI.ORG.AU/SS](http://WWW.AMSI.ORG.AU/SS)



IN YOUR 3RD YEAR OF MATHS AND STATS?

Find out what it's like to be a researcher!

# 2015/16

AMSI VACATION RESEARCH SCHOLARSHIPS

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Australian Government  
Department of Education and Training

**APPLY TODAY:**

[WWW.AMSI.ORG.AU/VRS](http://WWW.AMSI.ORG.AU/VRS)



## SA BRANCH

### The Aboriginal Health Landscape Project: Estimating resident populations for intercensal years



Craig Hansen



John Gray

John Gray and Craig Hansen presented their research about “Producing South Australian small area<sup>(1)</sup> estimated resident population (ERP) counts for 5 year age/sex and Indigenous status for intercensal years” to the SSAI South Australia Branch meeting on July 2015. John is the South Australian Health and Medical Research Institute (SAHMRI) Wardliparingga Aboriginal Research Unit’s Manager, Data and Analysis Support, as well as Analyst on the Landscape project. Craig recently joined SAHMRI in the position of Senior Epidemiologist within the same unit, with particular responsibility for management and analysis of data for the Aboriginal Diabetes Study (PROPHECY), a cohort study of Type II Diabetes amongst Aboriginal South Australians.

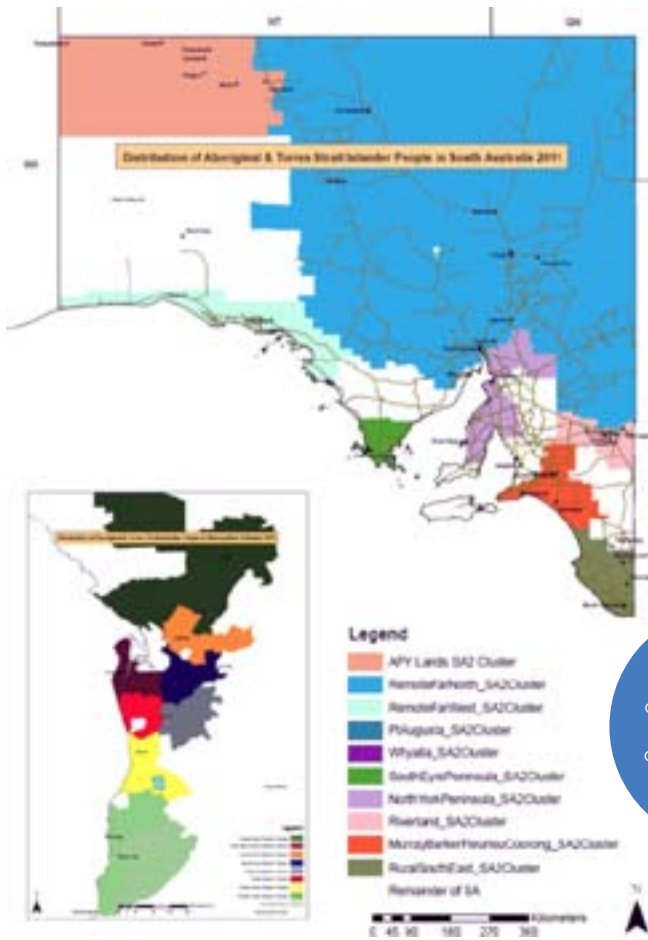
John and Craig’s current work informs the “Aboriginal Health Landscape Project” a flagship project of the Wardliparingga Aboriginal Research Unit. This project is intended to provide data and analysis that will be available to assist Aboriginal communities and service providers with:

- setting targets for better health outcomes,
- determining funding priorities for a fairer distribution of health programs and potentially with improved efficiency,
- prioritising Aboriginal health and health-related research,
- advocating for health improvements, and
- fostering informed debate on the work needed to achieve Aboriginal health equity.

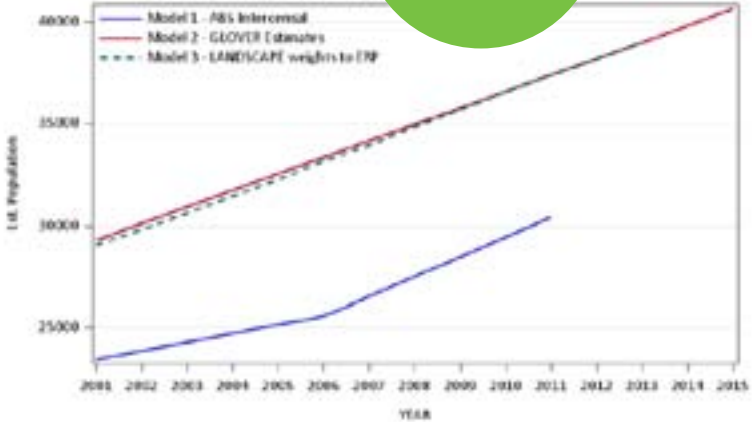
The Landscape project will report on contemporary population level health and health related data at the small area level, including population rates for selected measures such as hospital separations, emergency department admissions, immunisation rates, as well as on perinatal statistics, community mental health service activity and cause of death data. In particular, the project will describe the method used for creating meaningful geographic units of analysis (Landscapes) within South Australia, and then generating estimated resident population counts by Indigenous status, 5 year age groups and sex at the Landscape level for the years 2001 to 2013 inclusive.

John and Craig’s talk presented their attempts to estimate population counts for the Aboriginal and Torres Strait Islander population in South Australia in intercensal years. To generate these estimates they considered three models. Model 1 describes a linear progression between census figures for 2001, 2006 and 2011 by SA2<sup>(2)</sup>, Indigenous status, age and sex. Model 2 uses 2011 to 2015 population estimates provided by John Glover (at the Public Health Information Development Unit - PHIDU) as a base series, then backcasts<sup>(3)</sup> from 2011 to 2001 by SA2, Indigenous status, age and sex. Model 3 applies the LANDSCAPE distribution (% e.g. weights) of population for census years 2001, 2006, and 2011 to the ERP estimates (2001 to 2013) by Indigenous status, 5 year age and sex.

> Continued on next page



Comparison of 3 models across years (M1 = ABS, M2 = GLOVER, M3 = LANDSCAPE)



Typical map created as part of the Landscapes project (based on SA2 boundaries from the 2011 Census)

Some of the main challenges in estimating population counts of this type are uncertain quality of attribution of Indigenous status in baseline demographic datasets such as births and deaths, inconsistency in Indigenous identification over time and apparent under-counts of Indigenous people in census data. The presentation highlighted that generally, ABS population counts for Indigenous people between intercensal years are only available by age and sex for large geographic units such as "all of state" and remoteness level, or at a small area level for the total population.

Following the presentation, the speakers and branch members continued their discussion over dinner at a nearby Indian restaurant.

**Shahid Ullah**  
Flinders University

**Notes:**

- (1) Small geographical area.
- (2) For those unfamiliar with SA2s, please see the following definitions:
  - a) Statistical Areas Level 2 (SA2s): An area defined in the ASGS (Australian Statistical Geography Standard) which consists of one or more whole SA1s. SA2s are based on officially gazetted State suburbs and localities. In urban areas SA2s largely conform to whole suburbs and combinations of whole suburbs, while in rural areas they define functional zones of social and economic links.
  - b) SA1s: The second smallest geographic area defined by the ASGS. The SA1 has been designed for use in the Census of Population and Housing as the smallest unit for the processing and the release of Census data.
  - c) Mesh Blocks: The smallest geographic unit defined by the ASGS. Mesh blocks can be aggregated to represent many different geographical regions. There are approximately 347,500 mesh blocks in all of Australia.
- (3) Forecasting in reverse time.

## SSAI GOLDEN JUBILEE TRAVEL GRANT

**It provides overseas travel funds to SSAI student members, who can prove consecutive SSAI membership for a minimum of two years and who wish to attend overseas conferences at which they present a paper or poster.**

### **A complete application will consist of**

- Information on the conference and its importance to student's work (2-3 lines)
- Details of the paper/s/poster student wants to present at the conference
- A list of other funds sought or promised, including student's home institution
- Student's out of pocket expenses expected
- Any other supporting material student feels is necessary
- A letter of support SIGNED by one of student's supervisors AND student's Departmental Head
- Student's CV

**The application deadline is 31 March 2016.**



A maximum of \$1000 is available per application, limited to a single trip during the course of the student's studies. Students will not be supported in their first year of study and will have had to be members of the Society for at least 2 years prior to the application deadline. Applications are required to be lodged in advance of travelling. In exceptional circumstances an application can be for post-conference support, but the application will then have to be made within 1 month of returning and the 2 year mandatory membership period prior to departure must still be met. Exceptional circumstances are limited to unforeseeable student out of pocket expenses arising from other funding sources not fulfilling their obligation or changes to the trip that could not have been avoided.

If successful the student member is required to produce original receipts for amounts of equal or greater value than the grant. These receipts will be returned to the student marked with how much has been reimbursed. The student will therefore still be able to use the receipts for proof of attendance or to claim any funding shortfall from other organisations. The student member will also need to supply a report of his or her involvement in the conference to be published in the SSAI newsletter. This report should confirm the actual travel details and papers presented.

Recipients of the grant are asked to acknowledge the SSAI's support in the presentations and in any published version of the paper.

One travel grant is available per year. Assuming that more than one application will be received per year, either the Executive Committee or a special committee would help with the selection process.

For more information or to apply, contact the SSAI Office ([eo@statsoc.org](mailto:eo@statsoc.org)).

With this travel grant program the SSAI seeks to underline its objective to further the study, application and good practice of statistical theory and methods in all branches of learning and enterprise. It has been implemented to confirm to members that the SSAI is willing to support student statisticians and their budding careers.

## FROM THE OFFICE

Recently you would have seen a few emails asking for volunteers for our various committees. While the call for volunteers for the Accreditation Committee was very successful, unfortunately the call for a new member for the CPD Committee to represent the WA Branch was not. Obviously when events such as workshops and symposia are planned it is important that the CPD Committee is in a position to consider the interests of every SSAI branch and this can only happen if the committee represents a cross-section of the society.

SSAI would not be the society that it is without its volunteers. I wonder how many of our members are aware that SSAI actually only has two paid employees and both only work part-time? SSAI's Executive Officer works 20 hours a week for SSAI and our Membership Officer comes in for three hours a week. They manage memberships, the SSAI office and support SSAI's various committees, events and volunteers.

Everything outside the SSAI office is managed by volunteers. The articles you read in this newsletter have been put together by volunteers. These articles are generally about events, which have been organised by volunteers. The incredibly successful poster competition run by our very own Statistical Education section was put together by volunteers. The biennial Australian Statistical Conference is planned by volunteers with the help of a professional conference organiser. SSAI's website was initially planned by volunteers. SSAI's Accreditation process is managed by volunteers. There are many, many more examples that I could list.

I'm mentioning this today because I sometimes receive disgruntled messages from members who are not entirely satisfied with certain aspects of the society. What they are actually unhappy about is the way other members of the society have managed something – may it be an event, the website, the wording of an email, the structure of a poll, the time it took to get their accreditation through, or perhaps something to do with the newsletter. Are these critics aware though that the matters they are complaining about were handled by people with busy jobs, plenty of family commitments and on top of that perhaps some studying on the side? Yet these busy volunteers still have the stamina to give more of their time to their professional society.

I know none of the volunteers do their work because they expect accolades. However, today, the paid staff members of SSAI want to say a big "thank-you" to all the volunteers. Thank you to all you wonderful members who give so much. Your contribution to SSAI does not go unnoticed. SSAI needs you and more people like you.

**Marie-Louise Rankin and Irene Kiely**



Marie-Louise Rankin