

**Australasian Region**

**IBS-AR**

**Newsletter**



**Issue 21  
August 2017**

## President's Corner



What connects 1,388,757,042 and 29,276,272? They are the populations of China, and of Australia and New Zealand combined, sourced from <http://www.worldometers.info> just before noon on 1 August 2017. Yes, I needed an opening for this column!

Because those numbers are vastly different, this suggests that there are huge differences between the two regions. However, there are surprising similarities. For example, the China region of IBS has approximately the same number of members as our region. I learned this from Andrew Zhou whom I met on 1 August while attending the 2017 JSM in Baltimore (see photo right). Andrew is an active and visionary President of IBS-China. With others, Andrew was instrumental in establishing a new journal: *Biostatistics & Epidemiology*, the official publication of the International Biometric Society – Chinese Region. We discussed various opportunities to strengthen each other's region and to work towards establishing stronger links. I am positive that good initiatives will eventuate and bear fruit.

Hopefully a contingent of IBS-China members will attend our November [Biometrics by the Border](#) conference. It will be a memorable one and I am looking forward to our six female keynote speakers, as well as having both the incumbent IBS President, Elizabeth Thompson, and the incoming IBS President, Louise Ryan, attending and sharing their thoughts on topics relevant to our community and their research in biometrics.



*Zheyu Wang (Johns Hopkins), Andrew Zhou (President of IBS-China) and Samuel Mueller.*

### In this Issue

President's Corner.....	1
Australasian Regional Council .....	2
Member Profiles.....	3
Software Corner .....	5
Recent Papers .....	8
Student Scholarships.....	9
Membership News.....	10
Biometrics by the Border .....	11
Special Features .....	12

How to share ideas effectively, intriguingly and entertainingly seems to be a rare skill – this is me reflecting on the last three JSM conference days. Based on these vivid memories, here are **six** personal **S's for better presentations** at Kingscliff:

- **Smile!**
- **Open Strongly!** If someone walked out 60 seconds into your talk, they should remember something that reflects positively on your work.
- **Keep it Simple!** Everyone knows you are smart when they understood what you said.
- **Keep it Short!** All but one talk I've attended were or felt too long.
- **Stress the Novel!** What is new and why does it matter I kept wondering.
- **Summarize the Important!** If someone entered 60 seconds before the end of your talk, they should remember something that reflects positively on your work.

This issue is again packed with information and I hope that you keep enjoying the new format. I particularly appreciate the continued contributions from young statisticians; just have a look at the [Software Corner](#) written by two PhD students Kevin Y. X. Wang (U Sydney) and Charles T. Gray (La Trobe) and the two [Member Profiles](#) by Emi Tanaka (U Sydney) and Daniel Fernández (U Albany).



Samuel Mueller ([samuel.mueller@sydney.edu.au](mailto:samuel.mueller@sydney.edu.au))

## Australasian Regional Council

<b>President</b>	Samuel Mueller (University of Sydney)
<b>Past-President</b>	Ross Darnell (CSIRO)
<b>Secretary</b>	David Baird (VSN NZ)
<b>Treasurer</b>	Warren Müller (CSIRO)
<b>News Correspondent</b>	Vanessa Cave (AgResearch, VSNi)
<b>Regional Website Manager</b>	Hans Hockey (Biometrics Matters)
<b>IBS Incoming President</b>	Louise Ryan (University of Technology Sydney)
<b>IBS Executive Board</b>	Brian Cullis (University of Wollongong)
<b>IBS Representative Council</b>	Ross Darnell (CSIRO) Alison Kelly (Queensland Department of Agriculture and Fisheries) Kevin Murray (University of Western Australia)

[www.BiometricSociety.org.au](http://www.BiometricSociety.org.au)



Facebook group: [IBS-AR](#)

## Member Profiles

Thank you to Emi Tanaka and Daniel Fernández for kindly agreeing to provide profiles. We invite our mid-career members (i.e., members approximately 5 to 15 years into their professional career) to provide a profile for inclusion in a future newsletter. Please email the editor ([Vanessa.Cave@agresearch.co.nz](mailto:Vanessa.Cave@agresearch.co.nz)) for more information.

### Emi Tanaka

*The University of Sydney, School of Mathematics and Statistics*

*When and why did you join IBS?* I joined in 2015. This was because I started to work in the biometrics area and community involvement to me is important. IBS had a good reputation so I thought it was a good avenue to join to get to know fellow local biometricians. I wasn't wrong!



*What biometrics related work do you do?* My PhD research was in statistical methodology for bioinformatics (specifically, analysis of DNA motifs) but I've switched areas and now my application is in agriculture (generally related to plant breeding). The bulk of my research is related to statistical modelling now but I'm hoping to go back to a bit of bioinformatics and marry my two disparate areas of knowledge together. It's in the courting stage now so time will tell if it'll work out.

*What other research interests do you have?* Separate to my agriculture and bioinformatics research, I like to keep tabs on the so-called data science and posts related to data visualisation. The data science phenomena have been interesting to me and I go regularly to data science meet ups and occasionally play some kaggle competitions. It's probably more of a hobby than a research interest. As a bonus, I learnt a fair bit about machine learning, drawing pretty plots and interactive data analysis. It served well when I was writing up some teaching materials recently.

*What is your most memorable conference moment?* Biometrics by the Harbour (2015)! It was my first regional biometrics conference. It was well organised. What stood out for me was really the people. We would have fun strolling along the harbour, hanging out every night and enjoying the company of others. I'm certainly going to the 2017 one!

*Where do you see IBS-AR as a professional organisation in 10 years from now?* I hope it continues to grow and only get better. I would hope that the organisation would still be inclusive and supportive of its members. In particular, I hope that it improves in the recognition of the diverse work of the members and also its reach to the younger members. I think it is already heading that way and good to see the social media presence is improving.

## Daniel Fernández

*State University of New York: University at Albany.*



*When and why did you join IBS?* I joined in 2012 when I was a PhD student at Victoria University of Wellington. The first motivation to join IBS-AR was because, as a newcomer to the area and to New Zealand, I wanted to know more closely the professionals of the AR region and their research directions. IBS is a well-known international society, which is a leading organization for applied statistics. The IBS conference has been always an enriching experience and I have met colleagues, who have become not only my collaborators but also my friends.

*What biometrics related work do you do?* It has changed through the years. My main research has been the development of model-based clustering approaches for ordinal and mixed-type data. I have been working more on methodological research, but I love applied statistics. While I was doing my PhD, I was working more on ecological statistics. After moving to the United States, I am working more frequently with health data sets as a biostatistician.

*What other research interests do you have?* The development of goodness-of-fit techniques in different settings such as GEE, multilevel/hierarchical, and mixture models. Another field of interest to me is model-based classification approaches for non-continuous responses. I'm also very interested in the application of visualization tools for explaining fuzziness in clustering structures. Additionally, I am always interested in the application of statistical models (or the development of new ones) in problems arising in different fields such as ecology, linguistics, and public health.

*What is your most memorable conference moment?* There are a few. I have very fond memories of my first IBS-AR conference in Mandurah, back in 2013: "Biometrics by the Canals". The atmosphere there was very friendly and enriching at the same time. I met wonderful people and incredible professionals. It was special because I convinced myself that I wanted to work in academia. My long conversations with Shirley Pledger talking about it were great. Additionally, I remember the social dinner and how a bunch of us went to sing in a karaoke bar after that, and ended up in the swimming pool of the hotel.

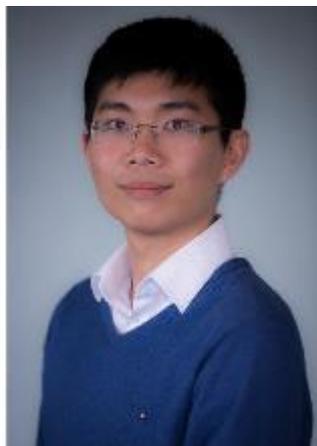
*Where do you see IBS-AR as a professional organisation in 10 years from now?* I think (and hope) IBS-AR will grow a lot, with more members (hopefully with a more balanced gender representation) engaged in all activities. During the last years, I have been lucky enough to attend other regional IBS conferences and I would say that IBS-AR is the strongest in ecological, agricultural, and environmental statistics. However, I would like to see the application of statistics in a more diverse range of fields presented at IBS-AR conferences.

## Software Corner

### A quick and dirty guide to getting started with purrr

#### Introduction

The **R** meta-package **tidyverse** is well known for its data-wrangling packages [dplyr](#) and [tidyr](#), as well as its visualisation capabilities through [ggplot2](#). It is less widely known that the **tidyverse** also offers a way of performing functional programming in **R** through the **purrr** package. The aim of this article is to highlight a few key techniques we wish we had known when we first started with **purrr**.



*Kevin Wang and Charles Gray*

We will make extensive use of the core **tidyverse** packages and functions. In particular, [dplyr](#), [tidyr](#), [tibble](#) and [piping](#). You might like to take a look at the series of [cheatsheets](#) published by **RStudio**. To get started, we install the **tidyverse** meta-package which is a shorthand way of installing [all the packages in the tidyverse](#).

```
# install.packages("tidyverse")  
library(tidyverse)
```

#### From lapply to map

The **purrr** package extends on the base **R** function `lapply`, allowing us to apply functions coordinate-wise along a list object. The `lapply` -analogous function within **purrr** is the `map` function. The syntax of `map(.x, .f, ...)` is similar to `lapply` where the first argument `.x` specifies a list, `.f` specifies a function, and `...` refers to optional arguments for `.f`. The `lapply` and `map` lines of code below both calculate the mean for each vector in a list and give identical output (a list of means).

```
list.of.vectors = list(c(1, 2, 3), c(2, 4), c(1, 3, 6))  
lapply(list.of.vectors, mean)  
map(list.of.vectors, mean)
```

A very useful feature of the `map` family of functions is their ability to control the output type. For example, `map_dbl` returns results as a numeric (double) vector:

```
map_dbl(list.of.vectors, mean)  
## [1] 2.000000 3.000000 3.333333
```

There are analogous map functions for logical `map_lgl` and character `map_chr` outputs, and more. The `map` function can be extended to two variables `map2`, or any number of variables with `pmap`.

## Tibbles

Where **purrr** becomes really powerful is when it can be used in conjunction with the data-type [tibble](#). A tibble is essentially a fancy version of a data frame that allows for list-type variables.

For example, take the classical `iris` data. Suppose we want to perform simple linear regression on the variable `Sepal.Length` against `Sepal.Width`, `Petal.Length`, and `Petal.Width` in the `iris` dataset, separately for each `Species`.

To perform this task, we would first need to partition the data using the column `Species` and then create three new `lm` objects (one for each species of `iris`) separately. Using **purrr** we can execute these steps by modifying a tibble of the `iris` dataset.

To partition the data, we can use the function `tidyr::nest` to break up the `iris` data by the `Species` column.

```
irisNest = iris %>% group_by(Species) %>% nest()
irisNest

## # A tibble: 3 x 2
##   Species      data
##   <fctr>      <list>
## 1 setosa <tibble [50 x 4]>
## 2 versicolor <tibble [50 x 4]>
## 3 virginica <tibble [50 x 4]>
```

Note that the `irisNest` data is now a tibble and the data-type of the `data` column is a list of tibbles. Each tibble has 50 rows, since there are 50 observations per `Species` in the `iris` dataset.

An excellent way to take a closer look at nested data is through the **listviewer** package by running `listviewer::jsonedit(irisNest)`. This produces an interactive widget that eases exploring the different levels of your nested dataset.

## Mapping tibble variables

Continuing with our `irisNest` data, a natural next step is to fit a linear model to each species. The model fitting is repetitive in the sense that we are fitting the same model formula to different subsets of the data. An easy way to achieve this is to use our new friend, the `map` function.

Specifically, the `map` function is used to create a list of `lm` objects. Using the `dplyr::mutate` function ensures that this list of `lm` objects will be returned as a new column in our tibble.

```
irisModel = irisNest %>%
  mutate(model = purrr::map(
    .x = data,
    .f = ~ lm(formula = Sepal.Length ~ Sepal.Width + Petal.Length,
              data = .x)
  ))
irisModel

## # A tibble: 3 x 3
##   Species      data      model
##   <fctr>      <list>   <list>
## 1 setosa <tibble [50 x 4]> <S3: lm>
## 2 versicolor <tibble [50 x 4]> <S3: lm>
## 3 virginica <tibble [50 x 4]> <S3: lm>
```

Manipulating data in this way allows us to perform modelling while retaining a copy of the data which the modelling was performed on. This is especially important for reproducibility as every data and model is properly aligned in a tabular form.

As a final touch, we will run `broom::tidy` on the `lm` objects. This will give us a list of data frames, each containing summary statistics of the corresponding linear model and then we can use `unnest` to simplify the output.

```
irisOut = irisModel %>%
  unnest(map(model, broom::tidy)) %>%
  mutate_if(is.numeric, round, 3)
knitr::kable(irisOut)
```

Species	term	estimate	std.error	statistic	p.value
setosa	(Intercept)	2.304	0.385	5.979	0.000
setosa	Sepal.Width	0.667	0.090	7.387	0.000
setosa	Petal.Length	0.283	0.197	1.437	0.157
versicolor	(Intercept)	2.116	0.494	4.282	0.000
versicolor	Sepal.Width	0.248	0.187	1.325	0.191
versicolor	Petal.Length	0.736	0.125	5.896	0.000
virginica	(Intercept)	0.625	0.525	1.190	0.240
virginica	Sepal.Width	0.260	0.153	1.695	0.097
virginica	Petal.Length	0.935	0.090	10.433	0.000

This way of mutating columns is particularly useful when dealing with grouped datasets for which a collection of statistics and diagnostic plots need consistent naming and organisation. Notice how all of our data and modelling objects can be stored compactly in one tibble object.

## Where to next?

We have only scratched the surface of **purrr**'s capabilities here. There are a couple of tutorials in development on GitHub that provide further information on **purrr**:

- Charlotte Wickham's **purrr** [tutorial](#);
- Jenny Bryan's excellent [tutorial](#) in development; and
- another place to keep an eye out for **purrr** vignettes is on the [R-bloggers website](#), where new **R** tutorials are posted daily.

Kevin Y. X. Wang (University of Sydney) and Charles T. Gray (La Trobe University)

*Thank you to Kevin Wang and Charles Gray for generously agreeing to write an article for our Software Corner. Should you wish to provide a software article for the next newsletter, please email the editor ([Vanessa.Cave@agresearch.co.nz](mailto:Vanessa.Cave@agresearch.co.nz)).*

## Recent Papers

The International Biometrics Society publishes two journals, *Biometrics*, and, jointly with the American Statistical Association, the *Journal of Agricultural, Biological, and Environmental Statistics*. Papers recently published by our members in these journals include:



## Journal of Agricultural, Biological, and Environmental Statistics

### March 2017, Volume 22, Issue 1

Joanne De Faveri, [Arūnas Verbyla](#), [Brian Cullis](#), Wayne Pitchford and Robin Thompson. Residual Variance–Covariance Modelling in Analysis of Multivariate Data from Variety Selection Trials.

To submit a manuscript visit: <http://www.editorialmanager.com/jabe/>

## Biometrics

To submit a manuscript visit: <http://www.biometrics.tibs.org/>

*If your paper has been inadvertently omitted from this list, please email the editor ([Vanessa.Cave@agresearch.co.nz](mailto:Vanessa.Cave@agresearch.co.nz)) so that it can be included in the next edition.*

## Student Scholarships

To help attract enthusiastic and talented students to career paths in biometrics, our region offers scholarships for students undertaking a fourth or honours year of study, or a coursework Masters, in statistics, mathematical statistics, biostatistics, bioinformatics or biometrics. The winners of this year's highly contested award are Yu Yang and Evelyn Tay. Congratulations!

### Yu Yang



Yu Yang is an honours year student majoring in statistics at the University of Otago, New Zealand. After earning a BSc degree in pharmacy from Shanghai Jiaotong University, she worked as a clinical research associate for 3 years. The main job task was to check and verify subjects' trial data in clinical studies. Attracted by study design and data analysis, she came to Dunedin to study statistics. Undergraduate study equipped her with basic knowledge in statistics. After graduating in 2016 as one of the top students, she was awarded the summer studentship, a 10-week research project. Her project focused on recovery of covariance matrix by probabilistic PCA model.

Encouraged by her supervisors, Dr Matthew Schofield and Prof Richard Barker, Yu is currently working on quantifying climate change. The project is about prediction of missing temperature data from the 14th century until current time based on tree ring data. Its aim is to develop a method of getting predictions from incomplete high-dimensional data.

### Evelyn Tay



During her previous career as a General Practitioner, Evelyn observed the growing need for specialists who are able to bridge the gap between medicine and mathematics. She wishes to utilise her experience in medicine, aptitude in mathematics and interest in statistics to contribute to public health practice and policy through research in statistical methods.

She is currently completing a BSc (Honours) in the field of statistics at the University of Western Australia. Her honours project will investigate the sensitivity and timeliness of commonly used Bayesian signal detection methods to identify batch dependent adverse events following immunisation. The timely detection of any abnormal rise in the incidence of adverse effects due to a batch of immunisations can lead to prompt removal of this batch from circulation. It is also important in maintaining public confidence in immunisation. Her primary supervisor is Dr Julie Marsh who is a lecturer in statistics at the University of Western Australia and a senior research fellow at the Telethon Kids Institute.

## Membership News

Welcome to our new members: Niranjan Adhikari (Student, RMIT University), Joshua Joannes Bon (Student, University of Western Australia), Lauren Borg (University of Wollongong), Petra Kuhnert (CSIRO), Alastair Lamont (Student, University of Otago), Wendy (Yiwen) Li (University of Adelaide), Esther Meenken (AgResearch), Kenyon Ng (Student, University of Western Australia), Nishika Ranathunga (Student, La Trobe University), Alethea Mary Rea (University of Western Australia), Evelyn Tay (Student, University of Western Australia) and Yu Yang (Student, University of Otago)

Membership as of 1 July 2017:

	Australia	New Zealand	Other*	
Life Member	2			2
Regular Member	136	39	7	182
Senior Retiree	13	3	1	17
Student Member	16	7		23
Supporting Member	1		7	8
	<b>168</b>	<b>49</b>	<b>15</b>	<b>232</b>

\* China, Germany, Thailand, United Kingdom, USA, Vietnam

### *Congratulations to Shirley Pledger*

Shirley was an Honour speaker at the [EURING Analytical Meeting](#) in Barcelona July 2017. Shirley's presentation on "Wrong Models - and when to use them" was an overview of her published capture-recapture models, mostly using finite mixtures to model heterogeneity of capture and survival probabilities.



### *Alastair Scott (1939-2017)*

It is with sadness that we learned of the death of Alastair Scott in May. Alastair, an Emeritus Professor of Statistics at the University of Auckland, was one of the finest statisticians New Zealand has produced. His research was characterised by deep insight and he made pioneering contributions across a wide range of statistical fields. Alastair was acknowledged, in particular, as a world leader in survey sampling theory and the development of methods to efficiently obtain and analyse data from medical studies.

Alastair joined the Australasian Region of the International Biometric Society in 1992.

A tribute to Alastair highlighting some of his many achievements can be found on the [Auckland University website](#).



## Biometrics by the Border

Australasian Regional Conference: 26th - 30th November 2017

The next regional conference of the Australasian Region titled "Biometrics by the Border" is to be held from 26-30 November at the Mantra on Salt Beach, Kingscliff, which is located just south of the Gold Coast in Australia. Registration and abstract submission is now open: <http://www.biometric2017.org>

The all-female line-up of keynote speakers celebrates achievements by women in the field of biometrics. We are excited that Professors Elisabetta Carfagna (University of Bologna, Italy), Di Cook (Monash University), Rachel Fewster (University of Auckland), Sonja Greven (LM University Munich, Germany), Louise Ryan (University of Technology Sydney, incoming IBS President), and Jean Yang (University of Sydney) will be sharing their latest insights in their respective fields. We will also be joined by the IBS President, Elizabeth Thompson.

Three pre-conference workshops are being offered: "Spatio-Temporal Statistics with R" presented by Chris Wikle and Petra Kuhnert (25-26 Nov), "The use of geospatial technology for agriculture & agri-environmental statistics" presented by Elisabetta Carfagna (26 Nov), and "Exploring data and models visually" presented by Di Cook (26 Nov).

The venue provides a range of accommodation types and budgets. It is adjacent to a great surf beach which is patrolled during that period. The local organising and program committees are working hard to make this a conference you will appreciate attending from both a social and science perspective.

Visit the conference website <http://www.biometric2017.org> for more information.



*Elisabetta Carfagna*



*Di Cook*



*Rachel Fewster*



*Sonja Greven*



*Louise Ryan*



*Jean Yang*



## Special Features

### SEEM conference

The 2017 Statistical Ecology and Environmental Monitoring (SEEM) conference is being held in Queenstown on Dec 6-8. There is a great line-up of speakers, including Alan Gelfand, Jennifer Hoeting, George Seber, Marti Anderson and Len Thomas. More details are available on the website: [www.stats.otago.ac.nz/conferences/seem2017](http://www.stats.otago.ac.nz/conferences/seem2017).

### 50th anniversary of the Cormack-Jolly-Seber model

Just over 50 years ago, three papers appeared which independently described the fundamental approach for analysing capture-recapture data. It is now called the Cormack-Jolly-Seber model. This anniversary is celebrated in the second issue of Statistical Science for 2016, guest edited by Steve Buckland and Byron Morgan. It features transcribed interviews with George Seber and Richard Cormack. In addition there are eight research papers that demonstrate how the capture-recapture area is still developing, with applications to genetics, social and medical areas, as well as ecology.

Shown in the photographs are Steve presenting a copy of the issue to Richard at St Andrews University; Rachel Fewster, a co-author of two of the papers in the issue, presenting a copy to George at the University of Auckland; and Byron presenting two copies to George Jolly's two daughters Heather Hannah and Fiona Davies. A third copy goes to their brother David Jolly, who lives in Saudi Arabia.



[www.BiometricSociety.org.au](http://www.BiometricSociety.org.au)



Facebook group: [IBS-AR](https://www.facebook.com/IBS-AR)