2020 Issue 02

SUTD Honours And Research Programme (SHARP) Newsletter

Innovating Research with Design

CONTENTS

02 SHARP HONOURS SESSION
AY2020/2021 HONOURS SESSION

03 DIVE INTO ‘COMPUTATIONAL STATISTICS’

05 REFLECTION: FROM SHARP HONOURS SESSION HOMEWORK TO JOURNAL PAPER

Scan to find out more about SHARP

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

SHARP HONOURS SESSIONS

SHARP students are required to attend the SHARP Honours Sessions during Freshmore Term 1 – 3. These Honours Sessions are courses that are held once a week and specially designed to equip students with deeper knowledge in Mathematics and Science, and with academic assignments which will empower them in their future research activities. The Honours Sessions are held on early weekday evenings, to allow SHARP students to reap the full benefits of the sessions while minimising potential timetable conflicts.

During each term, SHARP students have the opportunity to work with different instructors on different topics, for example:

- Term 1 is a deeper dive into Mathematics and its applications
- Term 2 in experimental aspects of Physics and Chemistry and their application in Biology and healthcare
- Term 3 will allow the students to learn about Mathematical Modelling and Statistical Physics.

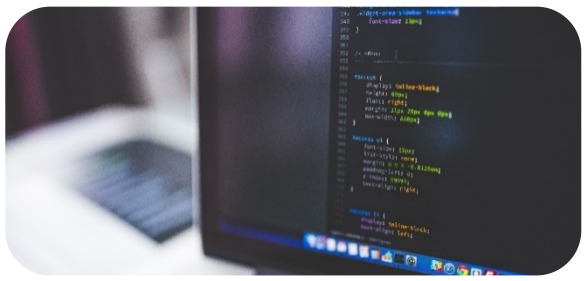
The Honours Sessions cover fundamentals that are broad enough to be of relevance to a large range of research areas, more details on the Honours Sessions can be found here.

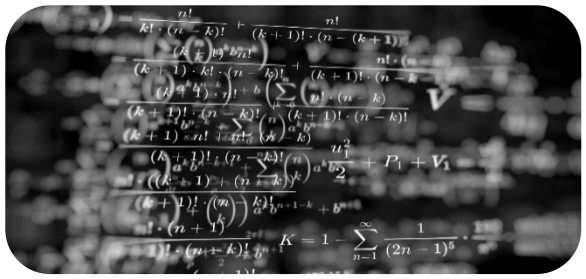
By allowing SHARP students to spend half of the term working on a hands-on Honours Session project, they benefit greatly through enhanced academic understanding and increased level of preparedness for a research environment.

In our efforts to enhance the learning experience for SHARP students, we will allow a selected group of non-SHARP Freshmore students to attend the Honours Sessions. However, the class size will be limited to twenty (20), so that the students can enjoy close supervision and interaction with the instructors and peers.

AY2020/2021 HONOURS SESSIONS

Week	Subjects	Faculty
SHARP Honours Session Term 1		
Week 1 to 3	Computational Statistics	Dr. Keegan Kang, ESD #
Week 4 to 6	Nonlinear Systems and Chaos	Prof. Ricky Ang, SMT ^
Week 8 to 13	Group Projects	
SHARP Honours Session Term 2		
Week 1 to 3	Organic Chemistry	Dr. Tan Mei Xuan, SMT ^
Week 4 to 6	Lasers	Dr. Massimiliano Colla, SMT ^
Week 8 to 13	Group Projects	
SHARP Honours Session Term 3		
Week 1 to 3	Modelling and Numerical Method	TBA
Week 4 to 6	Statistical Physics	Assoc. Prof. Dario Poletti, SMT ^
Week 8 to 13	Group Projects	





Engineering Systems and Design (ESD) Pillar
^ Science, Mathematics and Technology (SMT) Cluster

Scan to find out more about SHARP Honours Sessions

DIVE INTO 'COMPUTATIONAL STATISTICS'

Dr. Keegan Kang
Instructor for SHARP Honours Session
Lecturer, Engineering Systems and Design (ESD)

What is computational statistics?

Wikipedia states that computational statistics is the intersection of computer science and statistics, but I like to think of computational statistics as learning two languages, and translating one language in terms of the other.

One learns the language of math (linear algebra, multivariable calculus, probability), and translates this into workable code in a programming language of choice. This does not mean using a computer as a larger sized calculator, but more of **using techniques developed in the last few decades that are computer intensive and cannot be done by hand easily.**

Some people see computational statistics as machine learning (take a look at http://statweb.stanford.edu/~tibs/stat315a/glossary.pdf for a light-hearted take on this), and students would be familiar with some of these terms in the machine learning context.

However, I would like to think that computational statistics can be much more than this, and potentially covers (non-exhaustive list):

- coming up with a model for large scale data which is interpretable
- coming up with a model for large scale data which is computationally tractable
- using random numbers to solve computationally intractable problems and give close estimates rather than an exact answer

Could you share some interesting research topics related to Computational Statistics?

Sure! Here are some interesting work which I know of from researchers at two different universities.

At Cornell University, Professor Giles Hooker has done work which focusses on the interface between machine learning methods and statistical inference. He tries to understand how the process of machine learning translates into the stability of the predictions that it makes and develops ways to quantify that uncertainty, preferably without increasing computational costs. Having a notion of uncertainty then allows machine learning models to be used for statistical inference; in discovering which inputs are important, as components of a larger statistical model and in assessing the reliability of the explanations that are sometimes generated for the predictions of an automated system.

Application wise, much of Prof. Hooker's work has been motivated by collaborations with the eBird project at the Cornell Lab of Ornithology (https://ebird.org/home). Bird watchers record their observations (time, place, effort, birds seen) and these are then paired with environmental, topographical and land use features. The data is used to develop maps of bird habitat and migration, discover the environmental factors that affect bird species and monitor trends over time. Multiple complexities in the data collection process (who submits observations, where and when do they choose to go) make careful modelling challenging and important.

At Rice University, Ben Coleman has looked at approximating kernel densities for streaming data. Given a sample of data, density estimation is the task of determining the distribution that generates the data. Kernel density estimation (KDE) is a common and effective non-parametric density estimation approach, where we do not assume an underlying model. However, KDE is computationally expensive for many practical applications. The main difficulty is computing a kernel sum - a sum over the pairwise similarities between the query and each point in the dataset - because each query requires O(N) computations. Although most components in this sum are insignificant, it is hard to efficiently determine which points make large contributions. He has come up with a work "Sub-linear RACE Sketches for Approximate Kernel Density on Streaming Data," which solves the KDE problem in constant query time and limited memory. While these sketches only work for a particular set of kernels, they are sufficient for the vast majority of applications. More information can be found at https://randorithms.com/2020/09/15/RACE-KDE.html

What advice would you give to students who are interested in diving deeper into this topic?

Read more on what interests you! The field of computational statistics is rather broad, but a good grounding in linear algebra, multivariable calculus, and a bit of programming will get you started. It could be a few pages from a textbook, a blog written by data scientists or machine learning engineers, or even papers from arXiv or conferences. If the material is interesting, perhaps write some code to convince yourself that what you read is valid, and then ask yourself: "how can I use this...?“

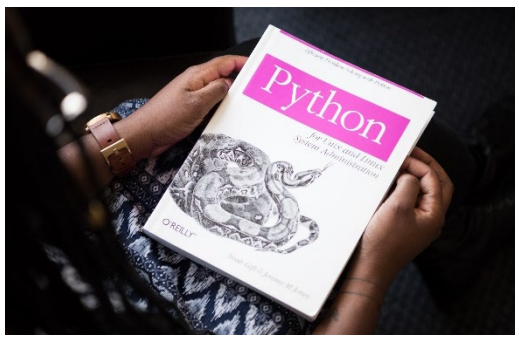


Image by Christina Morillo from Pexels

Could you also share how your portion of the course/topic relates to SUTD courses/teaching philosophy?

This is tough during the COVID-19 period, where students are supposed to sit apart in groups. But in this three weeks period, I have given two mini-projects as part of the homework assignments that allowed the whole class to work together as a group (and collaborate online to accomplish the project).

The first mini-project involved looking through several papers in computational statistics, and critiquing the plots in the experiments section. In some sense, would the plots convince a person to set aside some time to understand the paper? If not, what could be improved?

The second mini-project involved doing some basic computations for four algorithms, writing code to simulate some results, and then creating basic plots to verify their results. I have left this open ended, but encouraging the students to follow what they have done in their first mini-project (and to understand first hand if their critiques were justified or overly nit-picking).

I have also left this open up to students to decide where they want to focus on in their projects. Some students might be more comfortable in doing theoretical calculations, and other students might be more comfortable in writing code. The only requirement is that every student pull his or her own weight. And that is perfectly fine - we all have our own strengths, even as faculty members ourselves doing research.

In this way, the students not only get to delve a bit into computational statistics, they get to experience first hand what good (or bad) plots look like, and also have the opportunity to write code according to their calculations and design their plots to be comprehensible to convince others.

Image by Gerd Altmann from Pixabay

Do you enjoy teaching Computational Statistics to the SHARP students and why?

Yes! The students are very motivated, and ask a lot of questions, which in turn leads to a more productive class. There’s also more time to discuss the more interesting parts of the class, than to spend time on more basic foundations.

REFLECTION: FROM SHARP HONOURS SESSION HOMEWORK TO JOURNAL PAPER

By Jia Shuyi, Class of 2022

During my secondary and junior college studies, I was fortunate enough to be part of a few research projects in Physics and Organic Chemistry. The first-hand research experience I have acquired was invaluable and allowed me to realise that I find tremendous joy in discovering every little piece of new knowledge. Naturally, I looked for a research-intensive university after my GCE A Level examinations. I chose SUTD because of its new SUTD Honours And Research Programme (SHARP) – a premier undergraduate research programme with an emphasis on interdisciplinary curriculum and research exposure.

SHARP students are required to take advanced classes, which are also known as Honours Sessions. The very first Honours Session we had was on Parrondo's paradox – a fascinating phenomenon wherein a winning outcome is achieved by playing two individually losing games in a certain manner. I vividly recall the session during which Asst. Prof. Cheong Kang Hao (Science, Mathematics and Technology Cluster) first unravelled the paradoxical outcome. Amazement does not quite cover it. I was thrilled and could not stop thinking about the many possible applications of Parrondo's paradox. For instance, can one combine two losing trading strategies and cash out on the stock market?

Image by Lorenzo Cafaro from Pixabay

> **"This learning experience has opened my eyes to the research world out there, and I have now learnt to embrace research with an open mind."**

While the class materials for Parrondo's paradox were stimulating, **the homework given proved to be even more challenging and thought-provoking.** In one of the homework questions, Prof. Cheong encouraged us to creatively modify the rules of the paradox, as long as the paradoxical effect (obtaining a win from two losing strategies) is retained.

He also gave us the time and space to explore the handful of open questions that he set in the homework. Interestingly, a particular open question reminded me of my previous research experience in a Physics project. In that project, unexpected results were observed when the constant strength of an electric field was periodically varied. Then it dawned upon me that integration of Parrondo's paradox and periodic functions might be worth exploring. While having a lot of fun remodelling the paradox the way I wanted, I was entirely unaware that my naïve attempt of incorporating periodic functions into the paradox would morph into a proper research statement and eventually end up in a journal publication in *Physica A* [1].

Even though I had some experience in scientific writing, writing a proper journal paper was an entirely different "beast". I spent hours battling the blank page and getting stuck in typesetting. Making the story flow while giving an outlook on what is already known, explaining the motivation behind the work and highlighting the novelty was a real struggle that taught me to scrutinize my research results more critically, think outside the box and always link back to the bigger picture. Furthermore, the peer review process was probably more stressful than the actual research itself. It was only with the patience and everlasting support from Asst. Prof. Cheong and his research team that I pushed through the difficulties and honed my research skills. **This learning experience has opened my eyes to the research world out there, and I have now learnt to embrace research with an open mind.**

Looking back, I realise how far I have come from merely attempting a homework question. In my view, this is exactly what SHARP Honours Sessions are about: **exposing students to open-ended problems with proper support, while concurrently equipping them with the right fundamentals.**

[1] Shuyi Jia, Joel Weijia Lai, Jin Ming Koh, Neng Gang Xie, and Kang Hao Cheong (2020), "Parrondo effect: Exploring the nature-inspired framework on periodic functions". Physica A, 556, 124714. IF: 2.92.

DOI: https://doi.org/10.1016/j.physa.2020.124714

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

SHARP Newsletter: An SUTD publication from the SHARP office

Singapore University of Technology and Design

8 Somapah Road, Singapore 487372

SUTD Website: https://sutd.edu.sg/

SHARP Website: https://sutd.edu.sg/SHARP

Contact: sharp@sutd.edu.sg