Modern Statistics and Machine Learning for Global Health

Dates: February 22nd- 24th, 2024

Venue: AIMS Rwanda

Course summary

This short course will provide an introduction to modern methods at the intersection of epidemiology, statistical modelling and computer science, including probabilistic programming using Stan, computationally efficient non-parametric Bayesian inference, statistical techniques for infectious disease modelling and public health, and phylogenetics. This course will also aim to illustrate the application of these methods to epidemiological, genomic and spatial data, equipping attendees with the tools to carry out their own research projects in these areas.

The course will be delivered with a mix of lectures and computer-based practical sessions. Good knowledge of R is required to participate fully in the practical component.

Course Contents

Thursday (9.00-16.00) – February 22

Lecture 1: Introductions (9.00 - 9.30)

Bayesian Inference

Lecture 2: Introduction to Stan for applied Bayesian analyses (9.30-10.00)

Practical 1: Stan basics (10.00-11.30)

Break (*15 min*)

Lecture 3: Scalable Gaussian process regression models (11.45-12.30)

Practical 2: Scalable GP regression models (12.30-13.00)

Lunch (1 hr)

Practical 2: Scalable GP regression models (14.00-16.00)

Friday (9.00-16:00) – February 23

Infectious Disease Modelling

Lecture 4: Introduction to Infectious Disease Modelling (09.00-10:00)

Lecture 5: Introduction to phylogenetics (10:00-10:40)

Break (20min)

Lecture 6: (11.00-13:00) Four remote research talks from the Machine Learning and Global Health Network

Lunch (1 hr)

Practical 3: Running a phylogenetic pipeline (14:00-15:00)

Lecture 6: SIR models (15:00 – 16:00)

Saturday (13.00-18.00) – February 24

Practical 4: Deriving SIR type models (13:00-14:00)

Lecture 7: Introduction to Fitting an SIR model practical (14:00-14:10)

Practical 5: Fitting an SIR model in Stan (14.10-16.00)

Barbecue & cocktail: (16:00-18:00)

Pre-requisites
It is recommended that people attending are familiar with R (https://www.r-project.org/). We will provide a list of R packages to install prior to the course and instructions to set up Stan.

Maximum number of participants: 50 students

Teaching team profiles

Oliver Ratmann

I am Dr Oliver Ratmann, Reader in Statistics and Machine Learning for Public Good at Imperial College London of Science, Technology & Medicine in the UK. I am passionate about developing and applying scalable, interdisciplinary methods to tackle the grand challenges in global health research, and support underserved populations. I am co-leading the PANGEA HIV consortium, co-investigator and principal statistician on several studies of the Rakai Health Sciences Program, and co-founded the Global Reference Group for Children in Crisis and the Machine Learning and Global Health network. I am director of the MSc in Statistics at Imperial, and teach the Biostatistics course on the MSc. My group currently includes two post-docs, a research assistant, three PhD students, and several students and visitors. We are curious, diverse, and dynamic. We always look for people who want to grow with us, and are here to support young talents in Africa – get in touch.

https://mlgh.net/author/oliver-ratmann/

Juliette Unwin

Dr Juliette Unwin is a lecturer in statistical science at the University of Bristol. She is interested in developing and applying novel methods for infectious disease outbreak analysis to help inform policy makers in real time. Her current research focuses on developing spatial temporal renewal-based transmission models alongside estimating the number of children affected by COVID-19 and crises. She has previously been involved in real-time analysis of Ebola in the Democratic Republic of Congo alongside the World Health Organisation and COVID-19 in New York State with the local government.

https://research-information.bris.ac.uk/en/persons/h-juliette-t-unwin

Alexandra Blenkinsop

Dr Alexandra Blenkinsop is a Research Associate in the Department of Mathematics at Imperial College London. Her research centres around developing and applying methods to understand HIV transmission dynamics at a population level to inform policy decisions. She has also worked on projects related to COVID-19, including estimation of children affected by death of parents and caregivers during the pandemic. She has collaborated with the HIV Transmission Elimination Amsterdam Initiative, The Botswana-Harvard AIDS Institute Partnership and the United States Centers for Disease Control. She is a member of the Machine Learning and Global Health Network (https://mlgh.net/) and the PANGEA HIV consortium (https://www.pangea-hiv.org/).

Yu Chen

Yu Chen is a PhD student at CDT StatML (Modern Statistics and Statistical Machine Learning) at Imperial and Oxford. Her research focuses on modelling sexual network for HIV transmission analysis at high-resolution based on Bayesian inference. Her current research aims to understand and uncover the hidden gender-linked under-reporting issue and the reporting bias from the contemporary social survey data based on non-parametric Bayesian consistency model with Hilbert Space Gaussian Processes approximation. She is also involved in longitudinal analysis of number of children affected by primary causes-of-death in the U.S. and crises.

Google scholar profile

Shozen Dan

Shozen Dan, a PhD student at the StatML CDT, jointly hosted by Imperial College London and the University of Oxford, focuses his research on developing statistical methods with applications in public health, epidemiology, and infectious diseases. His current projects involve estimating social contact matrices with high precision, devising adjustments for reporting fatigue in longitudinal surveys, and analyzing the spatio-temporal variability of social contacts using established Bayesian statistical modeling techniques. Previously, he has collaborated with Center for Asian Research and Education (CARE) at Stanford University's School of Medicine on the analysis of health outcomes in Asian Americans. Additionally, he is an active member of the Machine Learning and Global Health Network (https://mlgh.net).

https://shozend.github.io/