THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Social Statistics group

PhD Open Day 2022

- **Lse.ac.uk/statistics/research/social-statistics**
- @LSEStatistics

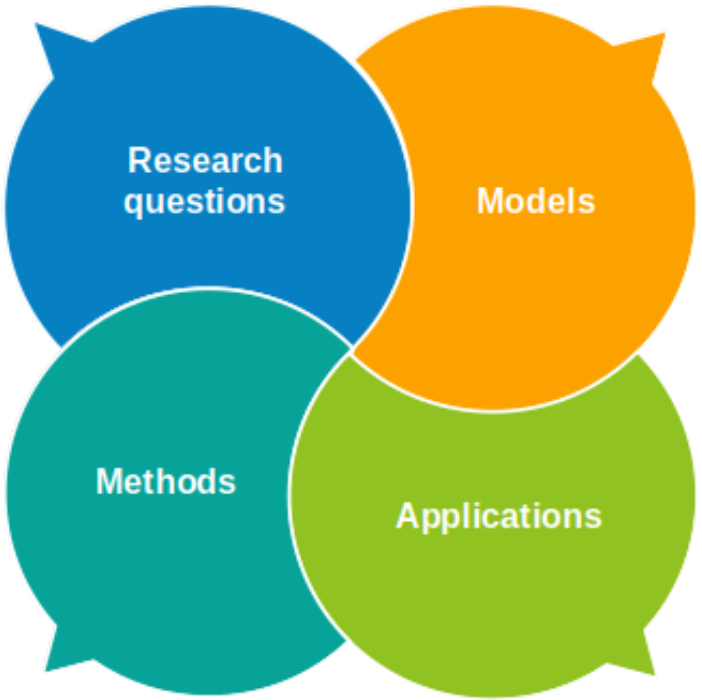
Research in the Social Statistics group

Research Questions

Social Statistics group conduct research in many areas of statistical theory and methods that are important for answering research questions in the modern social sciences

Methods

Methods to deal with data with complex structures: multivariate multilevel, clustered, survival data; large-scale item response data; data with missingness, systematic bias, high dimensionality. Methods for causal inference, categorical data and machine learning .



Models

Latent variable, mixture and random effects models; Gaussian processes; machine learning methods; marginal modelling; composite likelihood methods; models for dependence, graphical models.

Applications

Applications in criminology, demography, education, epidemiology, political science, psychology, social policy, and sociology.

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

The Social Stats group

Prof. Wicher Bergsma

Prof. Fiona Steele

Prof. Irini Moustaki

Dr. Yunxiao Chen

Dr. Sara Geneletti Inchauste

Dr. Kostas Kalogeropoulos

Prof. Jouni Kuha

- Social Statistics PhD Open Day 2022

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Professor Wicher Bergsma

- Statistical modelling and testing using reproducing kernels and (empirical) Bayes techniques.
- The I-prior methodology for parametric and nonparametric regression models
- Graphical models
- Conditional independence testing.
- Categorical data analysis, in particular marginal models which arise when there are nuisance dependencies in the data.
- The tau-star *test*, a scale invariant consistent test of independence together with Angelos Dassios from the PFI group.

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Professor Fiona Steele

- Longitudinal data analysis, multilevel modelling, survival analysis, and simultaneous equations modelling with **applications** in
  - Demography
    - residential mobility,
    - union formation and dissolution,
    - contraceptive use dynamics,
  - *Education*
    - the consequences of parental divorce for children's educational outcomes,
    - the impact of school resources on pupil attainment,
  - Family psychology
    - reciprocal influences between parents and children,
    - sibling interactions,
  - Health
    - child health,
    - mental health and employment transitions,
    - determinants and consequences of stress among nurses

- **Lse.ac.uk/statistics/research/social-statistics**
- **@LSEStatistics**

- Social Statistics PhD Open Day 2022

Professor Irini Moustaki

- Statistical methodology for analysing large and complex data sets.
  - latent variable modelling for categorical and mixed outcomes,
  - structural equation modelling,
  - estimation methods,
  - goodness-of-fit testing,
  - detection of outliers and
  - treatment of missing values and drop out in longitudinal studies.
- Areas of application include
  - Education,
  - Psychiatry
  - Health.
- Psychometrics lab at: https://psychometriclab.com/

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Dr. Yunxiao Chen

- Statistical and machine learning methods for high-dimensional data with noise
  - Analysis of large-scale item response data such as
    - Large scale educational survey analysis,
    - psychological testing, measurement and detection of aberrant behaviour,
    - numerical and stochastic optimization algorithms
  - Measurement and predictive modeling based on dynamic behavioural data
    - Development of continuous-time Gaussian processes and counting processes based algorithms to analyse:
      - Intensive longitudinal data, which are data with many measurements over time collected by smartphones, fitness trackers etc.
      - Problem-solving process data tracking individuals solving a computer-simulated task
  - Stochastic control of dynamic systems such as
    - compound decision theory for sequential hypothesis testing, stochastic control, rank aggregation, and change point detection with application to
      - online crowdsourcing,
      - item pool quality control in educational testing, and
      - technology-enhanced personalized learning.
- Psychometrics lab at: https://psychometriclab.com/

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Dr. Kostas Kalogeropoulos

- Bayesian Inference and Machine Learning Methods for
  - Factor Analysis
  - Mixture Models
  - Gaussian Process Regression
  - Sequential Learning
  - Latent Stochastic Processes
- Developing advanced computational schemes such as Markov Chain Monte Carlo and Sequential Monte Carlo.
- Target data that are potentially partial, noisy and from multiple sources.
- Applications include
  - Biomedical problems such as infectious diseases outbreaks via stochastic epidemic models
  - Financial and Econometric Time Series

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Professor Jouni Kuha

- Analysis of Social Science data with some of the following problems:
  - measurement error
  - misclassification
  - missing data
- And some of the following features
  - Categorical data
  - Survey data
- Modelling strategies including
  - Latent variable modelling
  - Structural equation modelling
- Applications include
  - The role of education in social class mobility,
  - Public attitudes to the police,
  - Safety citizenship behaviour in organizations,
  - Problem gambling
  - Intergenerational exchanges of family support.
  - Prof. Kuha has been a member of the analysis team of the broadcasters' exit poll for the four most recent UK General Elections.

Dr. Sara Geneletti Inchauste

- Developing methods for causal inference within a Bayesian framework
- Identifying and estimating effects of interventions using quasi-experiments
  - Regression discontinuity designs
  - Interrupted time series designs
  - Synthetic controls and negative outcome controls
- Adjusting for bias in causal inference due to
  - selection
  - time-varying confounding
  - informative missingness
- Applications include
  - Estimating the effects of:
    - Cholesterol lowering drugs (statins) in the general population
    - New sentencing guidelines for magistrates and judges in England and Wales
    - Austerity on mental health of minority populations in London
  - Investigating the sensitivity of the impact of ethnicity on sentencing outcomes in the presence of missing mediators and confounders

Why are you motivated to work with Social Science data?

Social science data are so varied and each dataset has its own complexities to account for in data analysis. I enjoy the challenges of working with different types of data, developing methods that mirror the structure of the data as closely as possible.

Prof. Fiona Steele

There is a chance to address research questions that might make a change to our lives for the better.

Prof. Irini Moustaki

Throughout my career I have been surrounded by social scientists. I have had many enjoyable collaborations with them, both working as a statistician on their projects and using their questions to motivate my own methodological research. Essentially all of my new ideas in statistics have been motivated by applied questions from social scientists.

Prof. Jouni Kuha

Why are you drawn to the methods you work on?

In social statistics a variety of models and methods have been developed that are useful in other statistical areas. To find such generalizations, reproducing kernel methods are extremely useful. I find these methods really exciting, as they have great mathematical beauty, computational simplicity (to a certain extent), and practical applicability.

Professor Wicher Bergsma

I mainly work on latent variable models, which are very powerful in social science research. They can model hypothetical constructs, unobserved heterogeneity, missing data, and counterfactuals, among others. These models also receive many machine-learning applications.

Dr Yunxiao Chen

The Bayesian framework makes sense to me and I find it easier to understand. It can also offer practical solutions, especially in the context social science data.

Dr Kostantinos Kalogeropoulos

Give an example of an area of research you'd like to get into

Causal inference which I found essential in the context of observational data.

Dr Kostantinos Kalogeropoulos

Latent variable models are closely related to deep neural networks. I plan to work on the interpretability of deep neural networks by bringing ideas from latent variable modelling.

Dr Yunxiao Chen

I'm working on integrating causal inference and thinking in my work.

Prof. Irini Moustaki

The "gold standard" of statistical inference is likelihood-based inference. Unfortunately, for many modern problems this approach is too computationally complex. I'm exploring quasi-likelihood methods, which sacrifice some efficiency in order to gain computational simplicity.

Professor Wicher Bergsma

What do you like about PhD supervision?

I enjoy working with young researchers trying to motivate them as well as my supervisors motivated me and also trying to help them avoid the mistakes I made.

Dr. Kostantinos Kalogeropoulos

Learning new things with the PhD researchers, exploring ideas to solve new problems.

Prof. Irini Moustaki

I enjoy the collaborative aspect, discussing ideas and approaches with the PhD researchers. It also keeps you abreast of all the new developments in the field. By the end of their PhD the researcher is the world expert in their topic!

Dr Sara Geneletti

What is good and bad about pursuing a PhD in Social Statistics?

You need to know many things because real-world problems in social science are complicated, and solving them requires many different tools. You may find that you need to learn much more than others in the beginning, but you will become very knowledgeable.

Dr Yunxiao Chen

The modern social statistician needs to develop a variety of skills in statistical theory, study design, statistical modelling, computational methods and management of complex data. It is also important to acquire some subject matter expertise in the field of application, or work with social scientists who can provide this. It is hard to juggle so many aspects, but extremely rewarding. People with these kinds of skills are in high demand and there will be lots of interesting opportunities after the PhD!

Prof. Fiona Steele

It is a long process with some phases of uncertainty but in the end it all comes together and the thesis is ready to be submitted!

Prof. Irini Moustaki

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Final Comments and Thanks

- As you can see the Social Statistics group is a vibrant and diverse group of people;
- With diverse and broad ranging research interests which develop and evolve.
- This is in part because we aim to answer questions in the Social Sciences using data produced by human beings in all their complexity and contradiction.
- Working in Social Statistics can also have real impact in the world, explaining relationships that are not clear and helping decision making and policy.
- It is also like being a detective, searching for clues hidden in the data! To me it is fun!

THANKS!