Mathematical Sciences

Data

Science

Methodology

14/12/2021
Jochen Einbeck

What is and isn't on these slides

What is:

A description of the current methodological activities in Statistics, Machine Learning, and Data Science by members of Statistics group of the Department of Mathematical Sciences.

What isn't:

- The Computer modelling group (different cluster)

- The Biostatistics group (different cluster)

- Activities by CS falling under this cluster (separate presentation)

- Several new hires of the group Jan 2022+

- Everyone else....

Durham
University

Everyone else who feels part of this cluster – sorry for not giving more attention!

Peter Matthews

Mohammed Alhanif

Jonathan Owen (Computer models)

Juraj Medzihorsky

Fred Worrall

Dario Domingo (Computer models)

Vincent Croset

Rui Carvalho

Ken McCaffrey

Sarah Wyer

Philip Stephens

Chunrong Feng

James Liley (Biostatistics)

Huaizhong Zhao

Rob Powell

Muhammad Hasan

Samuel Emerson

Paul Chazot

Matteo Degiacomi

Robert Lieck

John Wainwright

Toby Breckon (separate slides!)

Ulrik Beierholm:

Dan Lawrence

Marta Diaz-Guardamino

Patricia Muller

Steve Willis

Martin Cann

Tim Hawkins

Wayne Dawson

Andreanna Welch

Miguel de Lucas

Kristen Hopper

Bayesian methodology and computation

- Methodology development for...
  - Bayesian variable selection
  - Scalable Bayesian regression
  - Prior selection and elicitation

- Monte Carlo (Markov chain) methods for
  - variable dimension models
  - large data
  - intractable likelihoods

- Geometry and algorithms for Bayesian estimation
  - Invariant estimation
  - Analysis of 'graph cut' algorithms

Applied Bayesian modelling

- Uncertainty quantification and risk assessment
  - Inclusion of expert judgement (may be primary to data!)

- Application/specialization in
  - Ecotoxicology
  - Food safety
  - Banking
  - Digital commerce
  - Epidemiology
  - Energy
  - Petroleum engineering

Decision making under uncertainty

- Imprecise Probabilities
  - Foundations, algorithms
  - Applications: engineering, environment
- Nonparametric predictive inference (NPI)
  - Prediction with weak assumptions
  - Applications: medical statistics, finance
- Forecast interpretation and evaluation
  - Interpret (ensemble) forecasts
  - Probabilistic skill scores
  - Applications: weather and climate forecasts
- Energy infrastructure planning
  - Future buildings energy consumption
  - Robust decision support
  - Applications: climate change impacts on energy, reservoir modelling

Durham
University

Reliability and cryptography

- Risk and Reliability
  - Survival signatures
  - Applications: Agriculture, environment, food safety, engineering

- Cryptostats
  - Statistical methodology for homomorphic encryption
  - Privacy preserving techniques
  - Multi-party computation
  - Applications: banking, health, cybersecurity

Statistical modelling
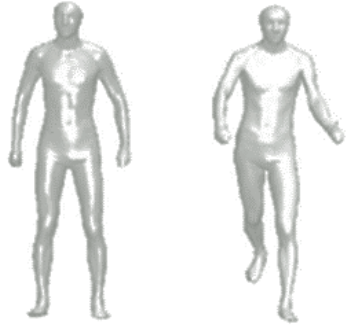
- Regression modelling
  - Generalized linear models
  - Categorical data
  - Mixed effect models
    - Diagnostics/Tests for random effects
    - Nonlinear models
    - Multilevel models

- Modelling of structure, geometry and shape
  - Random fields, statistical geometry.
  - Constrained density estimation
  - Applications: geometric data, imagery.

High-dimensional data

#MLmuse
CLAIRVOYANT
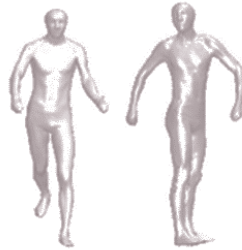
- Variable selection
  - LASSO regression and sparsity
  - Regularisation techniques

- Statistical inference
  - Valid post-selection inference

- Change point detection
  - Data splitting
  - Random projection

- Dimension reduction
  - Principal components, manifolds, etc.

Durham
University

People

Name	Theme
Louis Aslett	Cryptostats; Reliability; Biostats; Stat computing & MCMC
Camila Caiado	Bayesian modelling; Risk
Frank Coolen	Decision making under uncertainty; Risk and reliability
Tahani Coolen-Maturi	Decision making under uncertainty; Risk and reliability
Peter Craig	Bayesian modelling; Uncertainty analysis; Risk assessment
Jonathan Cumming	Bayesian modelling; Uncertainty analysis; Decision making
Reza Drikvandi	High dimensional data; Biostatistics; Statistical modelling
Hailiang Du	Machine learning; Uncertainty analysis; weather and climate
Jochen Einbeck	Statistical modelling; Biostatistics; High-dimensional data

People

Name	Theme
Michael Goldstein	Uncertainty analysis; Decision making under uncertainty
Sam Jackson	Uncertainty analysis; Statistical computation; Decision Making
Ian Jermyn	Statistical modelling of structure, geometry, and shape
Georgios Karagiannis	Bayesian methods and computation; Uncertainty quantification
Emmanuel Ogundimu	Biostatistics; Machine learning
Rachel Oughton	Uncertainty in complex systems
Dinos Perrakis	Bayesian methodology; Statistical modelling; High-dim data
Sebastian Schmon	Bayesian methodology and computation; Machine Learning
Ian Vernon	Uncertainty analysis; Decision making under uncertainty

Recent funding

Funding body, company, etc.	Theme/programme	Other info
EPSRC New Horizons	Statistical modelling of structure, geometry, and shape	U. Notts, 200k£
NERC	Hyperlocal flood warning system	ITB and CEH
EPSRC	Centre for energy systems integration	Uncertainty and decision making. Co
EPSRC/AI for Sci & Govt	Scottish Patients at Risk of Readmission and Admission	Turing & Public Health Scotland, £140k

Recent funding

Funding body, company, etc.	Theme/programme	Other info
UKRI Strategic Priorities/AI for Sci & Govt	Turing Health Programme Fellowship	Turing, £56k
HDR UK	Reproducible Machine Learning in Health Data Science	Turing
Innovate UK/UKRI	Knowledge Transfer Partnership	Carbon, £207k
Innovate UK/UKRI	Knowledge Transfer Partnership	Atom Bank, £184k
Turing	DECOVID C-19 rapid response data science taskforce	Turing secondment

Durham University

Recent funding

Funding body, company, etc.	Theme/programme	Other info
Wellcome	Uncertainty in Disease Models	£451,991
EPSRC	Decision Support Covid-19 modelling using JUNE	£440,000 (joint with Physics: PI Krauss)
European Food Safety Authority	N/A	Multiple small research contracts and departmental consultancies
German Federal Institute for Risk Assessment (BfR)	Soil organisms	Part of consortium, £40k for Durham

Thank you!