ACTIVITY REPORT 17/18

DATA SCIENCE FORUM

This report summarises Data Science Forum's activities for the 2017/2018 academic year.

```
tic
datatrans = log(10.^(-c)+data.^2);
mufi=0;sigmafi=10; %%% prior for phi
phi=truncatednormal(1,mufi,sigmafi,-1,1);

locsigma=2.5;scalesigma=0.01*locsigma;%%% prior for
sigmasqv
sigmasqv=1./(gamrnd(locsigma,1,1)./(scalesigma));
n = length(datatrans);
sigmasq = 1;%%% initial value for sigmasq
M = 3;
Maccept = 0;
Mcount = 0;
mu0 = 0; %%% initial value for mu0
emean = 0; %%% readjusted mean

s = zeros(1,n); %%% indicator variables

W = dirichlet([0.1 0.9]); %%% weights

for j=1:n, s(j) = discrete(W);end
n1=sum(s==1); n2=sum(s~=1);
mu = [log(10^(-c)) 0];
numb = [sum(s==1) sum(s==2)];

holdphi = zeros(1, numbofits);
holdsigmasqv = zeros(1, numbofits);
holdsigmasq = zeros(1, numbofits);
holdlogsigmasq = zeros(numbofits,n);
holdM=zeros(1,numbofits);
holdvar=zeros(1,numbofits);
holdmean=zeros(1,numbofits);
holdk=zeros(1,numbofits);
holdW=zeros(1,numbofits);
x = linspace(-20, 5, 1000); %%% predictive
predstar = zeros(1, length(x));

probparam=[0.5 0.5];
for it = 1:(burnin+every*numbofits)
```

118

Total registered members

63%

graduate-level students (MSc, Phd)

59

members from Department of Economics, Mathematics and Statistics

23

members from Department of Computer Science and Information Systems

89%

of respondents found the events 'useful' or 'very useful'

About our student society

Birkbeck's Data Science Forum is dedicated to the dissemination of methodological insight and applications in the areas of computational statistics and data analysis. We also take a keen interest in every major AI area, including machine learning, particularly when applied to financial markets. We aim to provide a bridge between research and industry to apply emerging technology, science, and entrepreneurship.

The Data Science Forum is a not for profit educational initiative with a simple but important mission. We help people gain understanding, acquire knowledge, and develop skills in: computational methods, and their applications in statistics and mathematics.

To fulfill that mission, the Forum provides both education experiences and career development opportunities for students of Birkbeck.

The Forum aims at bringing together students, researchers, and practitioners to discuss recent developments in computational methods, methodology for data analysis, and applications in statistics and mathematical finance. All topics within the broad interface of Computing, Applied Mathematics & Statistics will be considered for presentation.

The Data Science Forum will introduce the skills necessary for you to be both clients and designers of numerical methods for computer science applications. Here is what you should expect as an attendee of the workshops:

• Active participation. You will be actively engaged as a full participant throughout the workshops. Unlike typical events where most attendees are passive listeners, your active participation genuinely matters to the overall outcome.

• Solve real problems. You will participate in hands-on, interactive case study sessions to design real-world solutions to challenges in the sector, such as risk management, funding, and information technologies.

"Definitely interesting. Well done! Hope there are more events like this in the future. I'm positively impressed with Birkbeck."
- 1st year MSc student

• Open new vistas. Expand networks, meet partners, and learn from people with a wide range of quant backgrounds and industry experiences. You will learn about the latest computational developments, including current trends, key players, and opportunities.

These sessions and experiences will no doubt be useful in your chosen future professional environment.

Why data science?

We live in a world where the way industries compete is increasingly defined by the proliferation of data and increasing technological complexities. Organisations of all kinds are awash with data. All sorts of data, often coming from disparate sources, including financial data, customer data, call centre data, social media connections, and images, are collected by banks, insurance companies, and retailers. But for many in the industry, extracting insights from a data warehouse that contains structured and unstructured data is a daunting task.

This is where data science / computational statistics - the interface of numerical analysis, computer science, and statistics - can help. This interface includes computer arithmetic, algorithms, database methodology, languages and other aspects of the user interface, and computer graphics.

Many areas of applications can only be addressed effectively using computationally intensive statistical methods. This is often because the input datasets are so large, but it may also be because the problem requires consideration of a large number of possible alternatives. One area of application of such computational methods is finance and economics, in which heavy-tailed distributions or models with nonconstant variance are important. There is no doubt that computationally-intensive methods are becoming more commonly used in various areas of application of statistics.

Developments in other areas, such as computer science and numerical analysis, are also often directly relevant to data science / computational statistics. However, the research worker in this field must scan a wide range of literature. Numerical methods are often developed in an ad hoc way, and may be reported in the literature of any of a variety of discipines. Other developments important for statistical computing may also be reported in a wide range of journals that many statisticians are unlikely to read.

Keeping abreast of relevant developments in statistical computing is difficult not only because of the diversity of the literature, but also because of the interrelationships between statistical computing and computer hardware and software. Our aim is to map these cross currents of computational statistics and to encourage students to get involved with practical programming!

Charles Shaw
EMS representative

The team for 2017-18 academic year

Charles Shaw
Department of Economics, Mathematics and Statistics representative

E: cshaw11[at]mail.bbk.ac.uk
Charles is responsible for communications within the Department of EMS
BSc(Econ) Financial Economics, Birkbeck
MSc Financial Risk Management (Birkbeck, ongoing)

Matteo Lanzini
Department of Computer Science and Information Systems representative

E: jlanz01[at]dcs.bbk.ac.uk
Matteo is responsible for communications within the Department of CSIS
Bachelor's degree, Economics and Management, Bocconi
MSc Finance, Birkbeck
MSc Data Science (Birkbeck, ongoing)

Josue Jure
External outreach

E: jjure01[at]mail.bbk.ac.uk
Josue seeks and develops collaborative relationships with industry partners.
BSc Mathematics (Birkbeck, ongoing)

Using Python To Analyse Financial Markets

Thu 25 January, 6:30-8:00pm

In this talk we demonstrate the benefits of using Python to analyse financial market data, particularly time-series data. We discuss the parallels between the stages involved in solving a generalised data science problem, and the specific case of developing a trading strategy. We describe how open source Python libraries finmarketpy, findatapy, and chartpy aim to tackle these specific stages. In particular, we discuss how abstraction can be used to generate clean code for developing trading strategies, without the low level details of data collection and data visualisation. Further, we give Python code examples to show how we can download market data, analyse it, and how to present the results using visualisations. We also give an example of how to implement a backtest for a strategy using finmarketpy.

About the speaker: Saeed Amen is the founder of Cuemacro. He has a decade of experience creating and successfully running systematic trading models at Lehman Brothers (where he co-developed MarQCuS which had $2bn AUM), Nomura, the Thalesians and now at Cuemacro. Independently, he runs a systematic trading model with proprietary capital. He is the author of Trading Thalesians – What the ancient world can teach us about trading today (Palgrave Macmillan). He graduated with a first class honours master's degree from Imperial College in Mathematics & Computer Science. He is also the co-founder of the Thalesians.

Location:
Room 404, 30 Russell Square
Birkbeck
Bloomsbury
London
WC1E 7HX

Event summary: 32 people booked (of which 25 turned up). Audience: 11 students from Department of Economics, Mathematics, Statistics. 9 Students from Computer Science Rest from other departments. 1 phd student.

Careers in Risk, Insurance, and Wealth Management

Thu 30 November, 6:30-8:00pm

This event is designed for those with a deep understanding of statistics, an historical perspective, and a willingness to work with data. Representatives from Chartered Insurance Institute (CII), Allianz, Accenture, and other organisations were presenting at this event.

Presentations will cover:

InsureTech: the role of technology in risk and insurance, job market trends in this area incl. "data science"
Standards of technical competence, certs (finance, actuarial, risk, etc),
Career paths, incl. for mature students and career changers
Job market post-Brexit
What specific skills are needed to be successful in the sector
Bridging the gap between academic theory and industry practice

Why attend? Many good reasons. One such reason is that if you study economics, mathematics, statistics, or computer science then these companies would simply love to meet you. Worth noting that the last financial crisis was primarily a banking crisis, and as insurance industry representatives have regularly emphasised, the solvency of the insurance sector as a whole does not appear to be threatened. Why not? The answer is risk management.

Location
Room 624, Malet Street
Birkbeck
Bloomsbury
London
WC1E 7HX

Event summary: 46 people booked (of which 37 turned up). Audience: 18 MSc students from Department of Economics, Mathematics, Statistics. 8 Students from Computer Science (Mainly from MSc Data Science as we made a point of advertising the event to them). 4 Phds. Rest from other departments. Plus, unexpectedly, we had 2 MSc students (course unknown) that turned up from City and 1 from Brunel.

Audience

Below is a sample of the courses our members are studying:

BSc Mathematics
BSc Economics
BSc Financial Economics
MSc Business Innovation
Statistics (Graduate Diploma),
Mathematics (Graduate Diploma),
MSc Advanced Analytics,
MSc Advanced Computing,
MSc Advanced Computing Technologies,
MSc Advanced Computing Technologies (Data Analytics stream),
MSc Applied Statistics,
MSc Business Innovation,
MSc Computer Science,
MSc Computer Science (Intelligent Technologies stream),
MSc Data Science,
MSc Finance,
MSc Information Technology,
MSc Management,
MSc Mathematics,
MSc Computational Neuroscience,
and more

Our sessions are designed for advanced undergraduate and early graduate students who are comfortable with mathematical notation and formality. Some may need to review their knowledge of mathematical concepts alongside the algorithms under consideration.

Some sessions will be aimed at a more advanced audience, where are expected to have some background in the field. If you find the learning curve too steep, then do not worry. Advanced topics will be introduced gradually but rigorously while introducing numerical methods alongside mathematical background and motivating examples from modern computer science. It is hoped that the practical nature of this approach will help develop the intuition and comfort needed to understand more extensive literature in each subtopic.

Other activities not listed above included:

Google Hash Code competition
CV Clinic
Internship brokerage

Our Partners

MINERVA STATISTICAL CONSULTING

Minerva Statistical Consulting ffers free intensive courses in: Python, C++, Matlab, Java, R, SQL and VBA. Data and numerical analysis are considered, along with statistical, and financial applications. The courses blend practical applications and theoretical concepts. In doing so, the participant gains both of a deep understanding of the software as well as the capability of programming within familiar contexts.

http://minervastatisticalconsulting.co.uk

Birkbeck Institute for Data Analytics

The Birkbeck Institute for Data Analytics (BIDA) was founded to develop interdisciplinary research in Data Analytics and Data Science. BIDA provides the focus and a platform for research that combines data from computing and pervasive technologies, Statistics, Finance, Business, Biology, Physics, Medicine, Healthcare, Education, Transport, Politics, Geography, Archaeology, Linguistics, Psychology, Social networks, Social science, Law and Humanities.

http://bida.bbk.ac.uk

We extend our sincere appreciation to our sponsors, without whom we would have to rely on College funds to organise our activities. We would also like to thank Kevin Lau at Birkbeck Student's Union for his help in booking rooms. We would also like to thank Dr Alessandro Provetti and everyone at Thalesians for providing many constructive comments.

Birkbeck, University of London
Bloomsbury
London
WC1E 7HX

www.computationalmethods.net
info@computationalmethods.net