



Launch Day - April 30, 2010

*What's the place for statistical consulting
and computation in the academic world ?*

Program and Abstracts





Launch Day - Program

8:30 Registration

9:00 Opening address
Benoît **Macq** - Prorector, UCL and Rainer **von Sachs** - ISBA President, IMMAQ, UCL

9:10 Bernadette **Govaerts** - ISBA, SMCS, IMMAQ, UCL
"SMCS - Past, present and future"

Morning session - Chair: Catherine Legrand - ISBA, IMMAQ, UCL

9:30 Invited presentation: Saskia **le Cessie** - Leiden University Medical Center, The Netherlands.
"Can statisticians and medical doctors speak together, work together ... make research together?"

10:20 Godelieve **Masuy-Stroobant** - IACCHOS, UCL
"Social scientists and computer scientists: an evolving team work"

10:35 Alain **Guillet** - SMCS, IMMAQ, UCL
"Implementation of a quality plan for the monitoring of blood treatment process for the Belgian Red Cross"

10:50 Coffee break

11:15 Invited presentation: Koen **Knapen** - SAS Institute, Belgium
"Statistical consulting is a collaborative endeavor"

12:05 Isabelle **Thomas** - CORE, IMMAQ, UCL
"The point of view of the geographer: Computing spatial complexities"

12:20 Annie **Robert** - IREC, UCL
"Some pitfalls in the analysis and inference from clinical studies"

12:35 Céline **Bugli** - SMCS, IMMAQ, UCL
"Data analysis to serve success at university: the « Passeport pour le bac » project"

12:50 Lunch

Afternoon session - Chair: Dominique Peeters - CORE, IMMAQ, UCL

14:00 Invited presentation : Sébastien **Lê** - University of Rennes, France
"R as a link between academic and industrial worlds: how R became a must in a decade?"

14:50 Xavier Draye and Nicolas **Schtickzelle** - ELI, UCL
"Statistics for earth and life sciences"

15:05 Invited presentation : Martin **Theus** - Telefónica o2, Germany
"There is no excuse for failing to plot and look!"

15:55 Closing address
Claude **d'Aspremont** - IMMAQ President, UCL

16:00 Cocktails



Launch Day - List of abstracts

Saskia le Cessie - Leiden University Medical Center, The Netherlands

« Working as statistician in a medical environment, what are the challenges ? »

Working as a statistician in a medical school or hospital is one of the possible job opportunities for statisticians. In such a working environment, a large amount of time will be spent doing statistical consultation, teaching statistics to medical students and collaborating as a statistician in medical research projects. In this talk I will discuss some of the problems and challenges encountered when working with medical researchers. Topics discussed are: statistical consultation, collaboration in medical research projects, communication with clinicians, time management and the combination of statistical consultation and research.

Godelieve Masuy-Stroobant - IACCHOS, UCL

“Social scientists and computer scientists: an evolving team work”

The creation of the Belgian Archives for the Social Sciences – the Bass-unit – an extension of the Department of Demography of our University in 1970, is one of the many signs that there was a need for social science researchers to collaborate with computer scientists: their main objective was to document survey data files and administrative databases in order to enhance their accessibility to further research and data analysis. Statistical software was then only in its infancy and hard to use for social scientists who were and still are more familiar with concepts and theories than with programming languages. At that time we couldn't even dream of PC's and running one of the few available multivariate analyses handling categorical variables took a lot of time. Computer scientists act then as a kind of “Deus ex machina”, translating our queries into obscure programming languages, and helping us to save time.

The picture today is completely different: statistical software packages are even more “userfriendly”, while the variety of sophisticated data analysis methods is still expanding. Do we, in this context, still need the collaboration of computer scientists? The answer is straightforward, but their job profile has changed.

Alain Guillet - SMCS, IMMAQ, UCL

« Implementation of a quality plan for the monitoring of blood treatment process for the Belgian Red Cross »

The Belgian Red Cross has collaborated with the SMCS to implement quality control procedures to monitor its blood treatment processes in order to guarantee the quality of their production and be in conformity with the Belgian legislation. This project implied the adaptation of statistical quality control techniques to the particular problematic of blood treatment and the implementation of an integrated data management and statistical analysis package accessible to laboratory technicians without advanced statistical expertise.

Koen Knapen - SAS Institute, Belgium

«Statistical Consulting is a Collaborative Endeavor»

Statistical consulting is a living paradox: the most challenging and yet ultimately, the most rewarding work a statistician can do. A consultant combines the art and science of statistics to solve a practical challenge. This presentation will focus on the “problem-solving” activities of a statistical consultant, firmly rooted in a real-life ‘context’. Because ultimately, data and statistical analysis without context remains noise to the rest of the world. The presentation will address the following topics:

- The collection, the aggregation, the modification and transformation of data as vital links in the consulting cycle (and ultimately, the learning cycle).
- The foremost qualities of a statistical consultant (eg: a genuine desire to explore new subject matters in novel environments).
- The goals of statistical consulting : To get a model that approximates and provides an explanation to the real underlying data generating mechanism.

Perhaps George E. P. Box said it best: “all models are wrong, but some are useful”. Come hear why Google’s Chief Economist is not joking when he says that the sexiest jobs of the next decade will belong to the statisticians.

Isabelle Thomas - CORE, IMMAQ, UCL

« The point of view of the geographer: Computing spatial complexities »

Geography is by essence an integrative discipline. Hence, if geographical data are often shared with different domains, their specificity is that they are geocoded and located in space. The complexity of the geographical data/problems are an enormous challenge for more « conventional » statistical methods. By one example, we here show some specificities of the spatial analyses and particularly insist on the size of the data sets, the complexity of the reality to be modelled, the numerous biases introduced in modelling (spatial aggregation (scale), spatial sampling, spatial dependency) and the links between statistical analysis and geographical information systems (G.I.S.).

Annie Robert - IREC, UCL

“Some pitfalls in the analysis and inference from clinical studies”

Céline Bugli - SMCS, IMMAQ, UCL

« Data analysis to serve success at university: the « Passeport pour le bac » project »

The « Passeport pour le bac » project aims at promoting success of students in the first-year university level by a threefold intervention: identification of studies prerequisites, assessment of students background and understanding of these prerequisites and set up of specific actions to correct the lack of knowledge.

In this kind of project, one of the important challenges is to analyze efficiently the large sets of data collected from all UCLouvain faculties. Behind this analysis there are several important data management actions like efficient data merging and data cleaning. A lot of data checks are essential and macro programs are helpful to automate the procedure and reproduce them yearly. Such type of project will gain importance in the academic world in the future as big set of data coming from course evaluations, from various institutional surveys... are more and more used to take actions to improve the quality of university teaching and research.

To the question «What’s the place for statistical consulting and computation in the academic world?», we answer that statistics has also a role to play in this kind of institutional project where a high competence is needed in data management, cleaning, reporting and modeling.

Sébastien L  - Laboratoire de Math matiques Appliqu es, AGROCAMPUS Rennes, France

« *R as a link between academic and industrial worlds: how R became a must in a decade?* »

R is a free software environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modeling, statistical tests, time series analysis, classification, clustering, etc (<http://www.r-project.org/>). It has become, within a decade, a must for those who analyze data. The purpose of this communication is to show you how R has migrated from the academic world to the industrial world and how it has linked both of them. To do so, we will first present R throughout its history, its environment, the actors who contribute to its development directly or indirectly; then, we will present and demonstrate R through two fields of applications and two R packages, one related to science education and psychology with the FactoMineR package, the other one related to sensory analysis, with the SensoMineR package.

Xavier Draye and Nicolas Schtickzelle - ELI, UCL

« *Statistics for earth and life sciences* »

Earth and life sciences are heavily dependent on the appropriate use of a large panel of statistical tools, ranging from simple ANOVA to mixed models on complex variables, genomics, geostatistics, multidimensional statistics and phylogeny. The SMCS has become a key partner at UCL by helping both young and senior scientists in earth and life sciences facing a more-than-rapid evolution of techniques and data.

Martin Theus - Telef nica o2, Germany

« *There is no excuse for failing to plot and look!* »

Visualization is everywhere, be it a subway map, stock prices or the medals table on a newspaper's frontpage. We can distinguish between the visualization of qualitative information and the visualization of quantitative information. The latter is where statistical graphics comes into play, whereas the qualitative or structural visualization is the domain of InfoVis. Statistical graphics itself can further be subdivided into exploration, diagnostics and presentation. This talk elaborates on the various aspects of data visualization and will focus on statistical graphics for data exploration in particular. Exploratory data analysis (EDA) is at the core of any analysis of data, and interactive statistical graphics one of its most effective tools. We will have a look at the R software as well as the Mondrian visualization package to illustrate the power of graphical data analysis, which has its strength in deducing new insights and easily communicating these insights to domain experts.