	UNIVERSITÀ DEGLI STUDI FIRENZE DiSIA DIPARTIMENTO DI STATISTICA, INFORMATICA, APPLICAZIONI "GIUSEPPE PARENTI"	INSIDE project	INdividual air pollution exposure, extracellular vesicle SIgnaling and hypertensive disorder DEvelopment in pregnancy, PRIN 20152T74ZL

Course "Spatial statistical methods for epidemiological data"
January 30-February 3

Faculty:
Veronica Berrocal[1], Annibale Biggeri[2], Dolores Catelan[2], Corrado Lagazio[3]

Affiliations:
1Department of Biostatistics, University of Michigan, Ann Arbor, USA
2Department of Statistics, Computer Science, Applications "G.Parenti", University of Florence, Florence, Italy.
3Department of Economics, University of Genoa, Genoa, Italy

Course description
With the increasing availability of geographic information systems, spatial data have become more frequent in many disciplines, including public health and epidemiology. This course aims to provide an introduction to spatial statistical methods for epidemiological data, covering modeling approaches proposed in the literature for the different types of spatial support, e.g. point-referenced data, where the geographical coordinates of the observations have been recorded; and areal-averaged data, where summary statistics are reported for each areal unit. Topics covered include: exploratory analysis for spatial data, covariance functions, kriging, spatial regression, spatially-varying coefficients; spatial analysis for areal data, spatial smoothing and disease mapping; point processes, assessment of clustering, and cluster detection. Each lecture will feature a lab component, during which spatial analyses of datasets, made available to the participants, will be performed using the different statistical software, including STATA, R (downloadable at www.r-project.org) and WinBugs. Participants will be able to read and understand written English scientific material.

Location: Department of Statistics, Computer Science, Application "G.Parenti",
University of Florence, Florence, Italy.
Viale Morgagni 59, 50134 Florence

Tuition fees are not requested. A maximum of 20 participants is allowed.
Curriculum Vitae must be sent to catelan@disia.unifi.it within January 22.

Dipartimento di Statistica, Informatica, Applicazioni "*Giuseppe Parenti*"
Viale G.B. Morgagni 59, 50134 Firenze
+39 05542371, e-mail: catelan@disia.unifi.it

Detailed Program:

Day 1: Monday January 30, 2017
10am-1pm :

- Introduction to spatial statistics
- Geostatistical or point referenced data and first law of geostatistics
- Stationarity
- Semi-variograms; Covariance functions
- Kriging: simple, ordinary and universal kriging

Tutorial:

- Introduction to R packages for spatial mapping, exploratory data analysis (variogram computation) and kriging.

2:30-5pm:

- Case study on universal kriging within frequentist framework: presentation and lab exercise.

Day 2: Tuesday January 31, 2017
10am-1pm:

- Introduction to Bayesian statistics
- Bayesian prediction
- Hierarchical model for Gaussian spatial data
- Bayesian kriging
- Model-based geostatistics
- Spatial regression and spatially-varying regression coefficients

Tutorial:

- Introduction to R packages for Bayesian analysis, and Bayesian spatial data analysis.

2:30-5pm:

- Case study on downscaling air quality model output: presentation and lab exercise.

Day 3: Wednesday February 1, 2017
10-1pm:

- Introduction to areal data: spatial autocorrelation, Moran's I, Geary's C, correlograms
- Hierarchical models for areal data: conditionally autoregressive (CAR) models, proper CAR
- Simultaneous autoregressive models
- Disease mapping: Poisson-Gamma models, CAR models, and spatial convolution (Besag-York-Mollie' models)

Tutorial:

- Introduction to R packages for Bayesian hierarchical models for areal data
- Introduction to WinBUGS

2:30-5pm:

- Profiling high risk areas: frequentists and Bayesian approaches
- Case study on areal data analysis: presentation and lab

Dipartimento di Statistica, Informatica, Applicazioni "*Giuseppe Parenti*"
Viale G.B. Morgagni 59, 50134 Firenze
+39 05542371, e-mail: catelan@disia.unifi.it

Day 4, Thursday February 2, 2017:
10-1pm:

- Change of support discussion
- Introduction to point patterns: Complete spatial randomness (CSR) or homogeneous vs inhomogenous Poisson process
- First and second moment statistics: intensity function and K, F, G, J, L functions
- Clustered point patterns:
    - non-parametric and MLE estimation of intensity function of an inhomogenous Poisson process
    - Potthoff-Whittinghill test
    - finding clusters via Spatial scan test

Tutorial:

- Introduction to R packages for analysis of point pattern data

2:30-5pm:

- Approaches to the analysis of point pattern data arising from case-control studies: marked point processes
    - Cuzick-Edwards test of general clustering for case-control data
- Case study on case-control

Day 5, Friday February 3, 2017:
9-1pm:

- Introduction to point source analysis
    - Stone test
    - tests for case control data
    - GLM and Gam modelling
- Bayesian approaches to analysis of risk around putative sources
- Case study on point source problem: presentation and lab

2:30-3.30pm:

- Tying it all together: preferential sampling

Dipartimento di Statistica, Informatica, Applicazioni "*Giuseppe Parenti*"
Viale G.B. Morgagni 59, 50134 Firenze
+39 05542371, e-mail: catelan@disia.unifi.it

References:

S. Banerjee, B.P. Carlin, A.E. Gelfand (2014). "Hierarchical modeling and analysis for spatial data", CRC Press, Chapters 1-5 and Chapter 10.

R. Beelen, G. Hoek, E. Pebesma, D. Vienneau, K. de Hoogh, D. J. Briggs (2009). "Mapping of background air pollution at a fine spatial scale across the European Union", Science of the Total Environment 407, 1852-1867.

V.J. Berrocal, A. E. Gelfand, D. M. Holland (2010). "A spatio-temporal downscaler for output from numerical models". Journal of Agricultural, Biological and Environmental Statistics 15, 176-197.

V.J. Berrocal, A. E. Gelfand, D. M. Holland (2012). "Space-time data fusion under error in computer model output". Biometrics 68, 837-848.

A. Biggeri, C. Lagazio C. (1999), Case-control analysis of risk around putative sources, in A. Lawson, A. Biggeri, D.Böhning, E.Lesa re, J.Viel and R.Bertollini (eds.), Disease Mapping and Risk Assessment for Public Health, John Wiley & Sons, pp. 271-286.

A. Biggeri, D. Catelan (2012) Disease Mapping (Ch 9) in Frederick Faltin, Ron Kenett and Fabrizio Ruggeri (eds) Statistical Methods in Healthcare, Wiley, New York.

J.R. Bradley, C. K. Wikle, and S. H. Holan (2016). "Bayesian spatial change of support for count-valued survey data with application to the American Community Survey", Journal of the American Statistical Association 111, 472-487.

D. Catelan, C. Lagazio, A. Biggeri (2010) A hierarchical Bayesian approach to Multiple Testing in Disease Mapping. *Biometrical Journal* **52** (6):784-97.

L. Cecconi, L. Grisotto, D. Catelan, C. Lagazio, V. Berrocal, A. Biggeri (2016). Preferential sampling and Bayesian geostatistics: Statistical modeling and examples. Stat Methods Med Res. 25(4):1224-43.

A. E. Gelfand, L. Zhu, and B. P. Carlin (2001). "On the change of support problem for spatio-temporal data", Biostatistics 2, 31-45.

Dipartimento di Statistica, Informatica, Applicazioni "*Giuseppe Parenti*"
Viale G.B. Morgagni 59, 50134 Firenze
+39 05542371, e-mail: catelan@disia.unifi.it