Biostatistics 285
Advanced Topics: Recent Developments in Spatial Biostatistics

Instructor: Professor Sudipto Banerjee

Course Topics: Various topics in the statistical modeling and analysis for spatially referenced datasets including basics of Geographical Information Systems (GIS) data types and structures, exploratory data analysis and different methods for visualization of spatial data, different types of spatial regression models, Bayesian hierarchical modeling and analysis for spatial data, multivariate spatial analysis and spatial-temporal modeling.

Prerequisites: Prerequisites will include exposure to mathematical statistics and linear regression at the MS level as seen in Biostatistics 202A,B and 200A,B and C. Those wishing to enroll without these prerequisites will need to obtain consent from the instructor. No prior knowledge of spatial statistics will be assumed.

Learning Objectives:

COURSE LEARNING OBJECTIVES	***HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS***			
	Competencies for MPH in Biostatistics	***Competencies for MS in Biostatistics***	***Competencies for DrPH in Biostatistics***	***Competencies for PhD in Biostatistics***
1. Fit and compare different statistical models to analyze spatially referenced datasets from health and environmental sciences.	G3. Develop analytical and computational skills for the management, modeling and analysis of public health datasets with several variables that may be dependent on one another using statistically rigorous methods and models. G7. Describe preferred methodological alternatives to commonly used statistical methods when assumptions are not met.	A2. Formulate a public health question in statistical terms. A5. Identify and implement steps necessary to insure the quality of data collected in a study. A6. Conduct appropriate statistical analyses of study data and interpret the results. A8. Use statistical software to answer research questions and communicate the results to other research professionals.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B8. Research biostatistical methods and computational resources for collaborative research.	A2. Formulate a public health question in statistical terms. A4. Identify and implement steps necessary to insure the quality of data collected in a study. A5. Conduct appropriate statistical analyses of study data and interpret the results. A8. Research biostatistical methods and computational resources for collaborative research.

COURSE LEARNING OBJECTIVES	***HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS***			
	Competencies for MPH in Biostatistics	***Competencies for MS in Biostatistics***	***Competencies for DrPH in Biostatistics***	***Competencies for PhD in Biostatistics***
2. Use sophisticated statistical algorithms for fitting classical and Bayesian models and various software packages such as R and BUGS to analyze geographically referenced public health and environmental datasets.	G3. Develop analytical and computational skills for the management, modeling and analysis of public health datasets with several variables that may be dependent on one another using statistically rigorous methods and models. G6. Learn statistical programming and computational skills for conducting statistical simulation experiments, designing studies and analyzing public health datasets with several variables and potentially complex relationships.	A5. Identify and implement steps necessary to insure the quality of data collected in a study. A6. Conduct appropriate statistical analyses of study data and interpret the results. A8. Use statistical software to answer research questions and communicate the results to other research professionals. C4. Develop ability to use new and evolving computational and digital technologies into biostatistical work.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team. D3. Develop software and digital tools as necessary to apply statistical methodology. E5. Develop ability to evaluate and incorporate new and evolving computational and digital technologies into biostatistical work.	A4. Identify and implement steps necessary to insure the quality of data collected in a study. A5. Conduct appropriate statistical analyses of study data and interpret the results. B5. Develop ability to evaluate and incorporate new and evolving computational and digital technologies into biostatistical work. C6. Develop software and digital tools to implement novel biostatistical methodologies.
3. Understand the distinctions between different types of spatial data, the nature of their dependence, the underlying design by which they are collected, and the appropriate statistical methods for their analysis.	G2. Develop analytical skills and obtain broad insights involving the design and analysis of experiments to understand and model the dependence between different variables (e.g. regression), handle missing or incomplete data, and carry out rigorous statistical modeling for data obtained from a variety of public health study designs.	A2. Formulate a public health question in statistical terms. A3. Identify the strengths and weaknesses of different study designs to address public health and scientific questions; communicate these issues to public health researchers A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	A2. Formulate a public health question in statistical terms. A3. Identify the strengths and weaknesses of different study designs to address public health and scientific questions; communicate these issues to public health researchers. A5. Conduct appropriate statistical analyses of study data and interpret the results. A6. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.

COURSE LEARNING OBJECTIVES	***HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS***			
	Competencies for MPH in Biostatistics	***Competencies for MS in Biostatistics***	***Competencies for DrPH in Biostatistics***	***Competencies for PhD in Biostatistics***
4. Examine graphs, plots and other quantitative metrics to understand the nature and strength of spatial dependence in the data, what assumptions are tenable and possible inconsistencies between fitted statistical models and patterns in the data.	G6. Learn statistical programming and computational skills for conducting statistical simulation experiments, designing studies and analyzing public health datasets with several variables and potentially complex relationships. G7. Describe preferred methodological alternatives to commonly used statistical methods when assumptions are not met.	A2. Formulate a public health question in statistical terms. A6. Conduct appropriate statistical analyses of study data and interpret the results. A8. Use statistical software to answer research questions and communicate the results to other research professionals. C4. Develop ability to use new and evolving computational and digital technologies into biostatistical work.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team. D3. Develop software and digital tools as necessary to apply statistical methodology. E5. Develop ability to evaluate and incorporate new and evolving computational and digital technologies into biostatistical work.	A2. Formulate a public health question in statistical terms. A5. Conduct appropriate statistical analyses of study data and interpret the results. B5. Develop ability to evaluate and incorporate new and evolving computational and digital technologies into biostatistical work. C6. Develop software and digital tools to implement novel biostatistical methodologies
5. Construct and execute Bayesian hierarchical models to analyze spatial data at multiple scales and select the best model for the data based on principled statistical methods.	G3. Develop analytical and computational skills for the management, modeling and analysis of public health datasets with several variables that may be dependent on one another using statistically rigorous methods and models. G6. Learn statistical programming and computational skills for conducting statistical simulation experiments, designing studies and analyzing public health datasets with several variables and potentially complex relationships.	A2. Formulate a public health question in statistical terms. A6. Conduct appropriate statistical analyses of study data and interpret the results. A8. Use statistical software to answer research questions and communicate the results to other research professionals. C4. Develop ability to use new and evolving computational and digital technologies into biostatistical work.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	A2. Formulate a public health question in statistical terms. A5. Conduct appropriate statistical analyses of study data and interpret the results. B5. Develop ability to evaluate and incorporate new and evolving computational and digital technologies into biostatistical work. C6. Develop software and digital tools to implement novel biostatistical methodologies.

COURSE LEARNING OBJECTIVES	***HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS***			
	Competencies for MPH in Biostatistics	***Competencies for MS in Biostatistics***	***Competencies for DrPH in Biostatistics***	***Competencies for PhD in Biostatistics***
6. Understand the theory, principles and computational algorithms underlying statistical models to carry out spatial data spatial interpolation and prediction in the presence of misaligned or missing spatial data.	G3. Develop analytical and computational skills for the management, modeling and analysis of public health datasets with several variables that may be dependent on one another using statistically rigorous methods and models. G7. Describe preferred methodological alternatives to commonly used statistical methods when assumptions are not met.	A3. Identify the strengths and weaknesses of different study designs to address public health and scientific questions; communicate these issues to public health researchers. A6. Conduct appropriate statistical analyses of study data and interpret the results. A7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.	B6. Conduct appropriate statistical analyses of study data and interpret the results. B7. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team. D3. Develop software and digital tools as necessary to apply statistical methodology. E5. Develop ability to evaluate and incorporate new and evolving computational and digital technologies into biostatistical work.	A3. Identify the strengths and weaknesses of different study designs to address public health and scientific questions; communicate these issues to public health researchers. A5. Conduct appropriate statistical analyses of study data and interpret the results. A6. Effectively communicate the assumptions and results of analyses through oral and written communication to the collaborative team.

COURSE LEARNING OBJECTIVES	HOW THESE LEARNING OBJECTIVES ALIGN WITH COMPETENCIES FOR SPECIFIC DEGREE PROGRAMS			
	Competencies for MPH in Biostatistics	Competencies for MS in Biostatistics	Competencies for DrPH in Biostatistics	Competencies for PhD in Biostatistics
7. Participate in a spatial data analysis project to formulate scientific questions, develop rigorous statistical methods and computer programs to analyze the data and write a report describing the analysis and conclusions.	G3. Develop analytical and computational skills for the management, modeling and analysis of public health datasets with several variables that may be dependent on one another using statistically rigorous methods and models. G6. Learn statistical programming and computational skills for conducting statistical simulation experiments, designing studies and analyzing public health datasets with several variables and potentially complex relationships. G8. Develop written and oral presentations based on statistical analyses for public health professionals as well as lay audiences.	A1. Collaborate with researchers to formulate the aims of a public health research project. A2. Formulate a public health question in statistical terms. B2. Effectively communicate statistical concepts and reasoning to public health collaborators. B3. Learn to write and disseminate substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. B5. Be able to articulate interdisciplinary approaches to solving public health problems.	B1. Collaborate with researchers to formulate the aims of a public health research project. B2. Formulate a public health question in statistical terms. C4. Effectively communicate statistical concepts and reasoning to public health collaborators. C5. Learn to write and disseminate substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. D5. Understand and be able to effectively communicate the public health significance of the problems being addressed.	A1. Collaborate with researchers to formulate the aims of a public health research project. A2. Formulate a public health question in statistical terms. B1. Develop ability to critically read statistical methodological literature. C3. Learn to write and disseminate substantive field publications and communicate the statistical portion of the methodology to a substantive field audience. C5. Be able to articulate interdisciplinary approaches to solving public health problems.

Homework Assignments: You will be required to hand in several assignments throughout the quarter based on a combination of problems in the textbook and data analyses from the computer lab. Homework assignments will include a blend of theoretical, computational and data analytic questions.

Midterm and Final Examinations: There will be one midterm and a final examination held in class. The final examination will be a take-home examination and comprise a real-world spatial data analysis project in addition to some theoretical modeling questions.

Grading: Grades will be based on the homework assignments (30%), the midterm (30%) and the final (40%).

Data Sets: The data sets that are used for homework assignments will be available for downloading on the internet. Some of the analyses in the course will use data sets described in the course text and others will be available from within freely available sources. Supplementary materials describing this data set will be available on the course web site.

Computing: Statistical computing will be performed within the R statistical software environment. While familiarity with the R statistical computing environment will be helpful, it is not essential and all spatial packages in R will be discussed and demonstrated in the class.

Course Text: The required textbook for the class is *Hierarchical Modeling and Analysis for Spatial Data*, 2nd Ed., by S. Banerjee, B.P. Carlin and A.E. Gelfand, Chapman & Hall/CRC, Boca Raton, FL: 2014. It is available in the Health Sciences Bookstore.