1-2014

STAT 544.01: Topics - Spatial Statistics

Jon M. Graham
University of Montana - Missoula, jonathan.graham@umontana.edu

STAT 544 Spatial Statistics Spring, 2014

Instructor: Jon Graham, Math 204 — jgraham@mso.umt.edu, 243-2561

Time/Room: Mon, Wed, Fri, 2:10pm-3:00pm, in room Math 211

Textbook: None required, Readings will be given when appropriate.

References:

1. Interactive Spatial Data Analysis / Bailey & Gatrell (1995)
2. Statistics for Spatial Data / Cressie (1993)
3. Stat. Methods for Spatial Data Analysis / Schabenberger & Gotway (2005)
4. Introduction to Geostatistics / Kitanidis (1997)
5. Applied Geostatistics / Isaaks & Srivastava (1989)
6. Statistical Methods in Spatial Epidemiology / Lawson (2006)
7. Spatial Autocorrelation / Cliff & Ord (1973)
8. Statistical Analysis of Spatial Point Patterns / Diggle (1983)

Office Hours: To be announced, By appointment

Course Webpage: http://www.math.umt.edu/graham/stat544/

Grading: Homework: 35% Midterm: 25% Final Project: 40%

Prerequisites: One year of statistical methods or consent of instructor

Homework will be assigned approximately biweekly, and will generally require more than a few days work. The intent is that you will work on homework assignments throughout the period they are assigned. NO LATE HOMEWORK WILL BE ACCEPTED FOR ANY REASON, and the lowest homework grade will be dropped. Homework is not only a fairly substantial portion of your grade, but is vital to your success in this class. Working with other students on homework is strongly encouraged, as long as you hand in your own work, and do not simply copy someone else's work. Solutions to all problems will be provided.

The **midterm exam** will be cumulative and closed book. More about the exam, including the exact date will be given later. If you cannot make it to the midterm exam, you must let me know BEFORE the exam is given. No make-up exams will be given without a documentable reason for missing the exam.

A **final project** is required for this course. The project will entail both a written and oral portion to be conducted at the end of the semester. Please come see me by the midterm exam to discuss a topic for your project. You may choose your own topic, or come see me for help selecting one. Final projects will be presented in a poster session to be held from 3:20-5:20pm, Tuesday, May 13, and final papers will be due by 5:00pm on Friday, May 16.

Course Material and Objectives: This course is intended as an introduction to the topic of spatial statistics for graduate students with at least one year of introductory statistics. After an introduction to spatial exploratory methods and spatial correlation, the course will focus on the three primary areas of spatial analysis methods: geostatistics (variograms, isotropy, kriging, cross-validation, spatial correlation models), spatial point processes (quadrat methods, complete spatial randomness, Ripley's K-function), and area or lattice processes (Moran's I, Geary's C, Markov random fields, CAR models). Throughout the course, the software package R will be used for classroom demonstrations as well as for homework assignments. **R** is available on all computers in Math 206 (near my office) and is free for download as explained on the **Introduction to R** handout.

My goal in this course is absolutely **NOT** to teach you how to use software for working with Geographic Information Systems (GIS) such as ArcInfo or other specialized software. My goal is to teach you to think about problems involving spatial data and how to address spatial correlations which may be present in your data. You will learn a number of techniques for exploring and analyzing spatial data throughout this course, but the most important thing you should get from this course is an understanding of spatial continuity and its implications on analysis.

Additional Course Information: The last day to add this course through Cyberbear is Tuesday, February 4. The last day to drop this course or change the grading option through Cyberbear is Friday, February 14. Between Saturday, February 15 and Monday, April 7, you can drop the course with your advisor's and instructor's signature using a Drop form only. I will not recommend approval of late drops except in EXTREME circumstances (see the UM online catalog). You can add the course or change the grading option with an Add/Change form up until Friday, May 9.

Academic Misconduct: All students must practice academic honesty. Academic misconduct is subject to an academic penalty by the course instructor and/or a disciplinary sanction by the University. All students need to be familiar with the Student Conduct Code. The Code is available for review at http://life.umt.edu/vpsa/student_conduct.php.

Disability Services: The University of Montana assures equal access to instruction through collaboration between students with disabilities, instructors, and Disability Services for Students (DSS). If you think you may have a disability adversely affecting your academic performance, and you have not already registered with DSS, please contact DSS in Lommasson 154. I will work with you and DSS to provide an appropriate accommodation.

Questions are strongly encouraged during class & office hours. If you are at all lost, please let me know.

Summary of Topics

1. Introduction to Spatial Statistics
2. SPATIALLY CONTINUOUS (GEOSTATISTICAL) DATA (Chapters 5,6 - Bailey & Gatrell)
   - (a) Exploratory Analysis (Ch. 1 - Bailey & Gatrell, Ch. 2,3,4 - Isaaks & Srivastava)
     - i. Univariate: histograms, summary statistics
     - ii. Bivariate: scatterplots, correlation, linear regression
     - iii. Moving window statistics, h-scatterplots, greyscale maps, contour plots
   - (b) Spatial continuity, stationarity, isotropy (Ch. 1,5 - Bailey & Gatrell)
   - (c) Mean, variance, covariance, correlogram, (semi)variogram (Ch. 1 - B & G)
   - (d) Describing geostatistical data (5.4 - Bailey & Gatrell, Chap. 4,7,10,11 - Isaaks & Srivastava)
     - i. Spatial moving averages
     - ii. Tesselation methods
     - iii. Kernel estimation
     - iv. Covariograms, variograms, sill, range, nugget
     - v. Variogram estimators
     - vi. Anisotropic semivariograms, testing for geometric anisotropy, relative variograms
   - (e) Modelling geostatistical data (5.5 - Bailey & Gatrell, Chap.11-16 - Isaaks & Srivatava)
     - i. Ordinary, weighted, generalized least squares
     - ii. Variogram models, fitting variogram models, confidence intervals for semivariograms

iii. Simple kriging
iv. Effects of variogram on predictions, best linear unbiased predictors, crossvalidation
v. Ordinary and universal kriging

(f) Kriging extensions (Section 6.2 - Bailey & Gatrell, Chap. 4,13,17 - I & S)

i. Block kriging, indicator kriging
ii. Cokriging, cross-covariances, cross-variograms

3. POINT PROCESS DATA (Chapters 3,4 - Bailey & Gatrell, Chapter 3 - Schabenberger & Gotway)

(a) Exploratory Analysis (Ch. 3 - B & G, Ch.1,2 - Diggle, 3.3 - S&G)

i. Dot maps, intensity measures
ii. Quadrat methods, kernel estimation, nearest neighbor methods
iii. Ripley's K-function, L-function

(b) Point process models (3.5 - B & G, 4 - Diggle, 3.1,3.2 - S&G)

i. Complete spatial randomness (CSR), Poisson point models
ii. Testing for CSR via quadrat, NN, Ripley's methods

(c) Relationships between multiple point patterns (χ^2, K_{12}, L_{12}) (Ch. 4 - B & G)

4. AREA OR LATTICE DATA (Chapters 7,8 - Bailey & Gatrell, Chapter 6 - Cressie)

(a) Exploratory Analysis (Ch. 7 - B & G, 6.2 - Cressie)

i. Chloropleth maps, density equalized maps
ii. Spatial proximity (neighborhood) matrices

(b) Median polish, kernel estimation (7.4.3,7.4.4 - B & G, 2.2 - Cressie)

(c) Testing for spatial autocorrelation (7.4,7.5 - B & G, Upton & Fingleton (1985), Cliff & Ord (1973))

i. Randomization test, Monte Carlo approach, normal approximations
ii. Join count statistics (Moran's I, Geary's C)

(d) Regression with Spatially Correlated Errors (SAR, CAR models) (7.5.4 - B & G)

(e) Relationships between multiple lattices (Mantel tests) (Other sources)

This is a lot of material for a 3-credit semester course and could easily be taught over two or three courses. As a result, we will explore some topics in depth and others with the goal to merely expose you to them.