ANNOUNCEMENTS
• Please turn in HW3 no later than Tuesday, May 23rd

TODAY
• Introduction to spatial statistics
• Types of spatial data
• Exploratory analyses of spatial data in R
• Course Project

Spatial Statistics
OVERVIEW
Main Theme:
Observations that are closer together in space are more highly correlated.

Other Themes:
• borrowing strength more heavily from surrounding areas when estimating unknown quantities
• uncertainties in predictions are greater farther away from observations

SOME AREAS OF APPLICATION
Environment
• interpolate monitor readings
• determine whether an area is in compliance with EPA mandates
• assess environmental justice issues

Biology/Ecology
• evaluate the nature of the interaction between species (competitive vs. facilitative)
• predict the spread of an invasive species
• determine the best management strategies for ecosystems

Epidemiology
• identify “hot spots”
• provide more accurate estimates of disease rates by smoothing
• relate disease rates to covariates (environmental hazards, demographic information, etc.)

Economics
• real estate prices
• locations of new stores

Sociology
• crime counts

Other: image processing, hydrology, computer experiments, geology, meteorology, etc.

HISTORY OF SPATIAL STATISTICS – OVERVIEW
• Early approaches were mostly descriptive.
• In the 1920’s and 1930’s, R.A. Fisher introduced methods to control for spatial dependence using randomization.
• Formal modeling of spatial dependence emerged in a variety of disciplines including statistics, mathematics, mining engineering, geology, and statistical physics.
• In the 1980’s and 1990’s, spatial statistics became a recognized subdiscipline of statistics.
• Currently, spatial statistics is an exciting research area and is becoming more common in standard analyses performed in applied areas.

TYPES OF ANALYSES
• Descriptive
• Formal Modeling
• Inference
• Prediction
TYPES OF SPATIAL DATA

1. Geostatistical / Point-Reference Data -- data are the values of a process at various fixed locations
   Example:

2. Lattice / Aerial Unit Data – data are summaries (aggregation) of a process taken on a regular or irregular grid
   Example:

3. Point Pattern / Point Process Data – the locations of observations are of interest
   Example:

Questions:
1. Does the intensity (number of trees per unit area) vary over the region?
2. Is there an interaction among trees of the same species or are there interactions between species?

Example: Lattice / Areal-Unit Data

Questions:
1. What is the nature of the spatial dependence?
2. What is the effect of the percentage of the population working in agricultural occupations on lip cancer rates?
Example: Geostatistical / Point-Referenced Data

PM2.5 Data in North Carolina

Questions:
1. What is the PM level in counties where there are no monitors?
2. How can we use the neighboring monitors to provide a better estimate of the average PM level in all counties?
3. How much uncertainty is there in the PM level in the different counties?

Extension: Space-Time Geostatistical Data

Pollution data from a site in northeastern Alabama (lat<30.3, lon<84.7)
January 1990 - December 2001

Extension: Multivariate Geostatistical Data

Locations of PM Monitoring Stations