Data 589: Special Topics: Spatial Statistics Study Guide

M. J. Noonan

April 20, 2023

The majority of the material in the lecture slides is testable, however, this guide is intended to focus your studying efforts. The exam will consist of 10 multiple choice questions, 3 figure interpretation questions, 1 code interpretation question, and 1 equation interpretation question.

Module 1: Working with Spatial Data

- Understand the distinction between spatial data and other data structures, and the general need for specific spatial statistics tools.

- Understand and differentiate between the two different data types we covered during this course (i.e., point patterns vs continuous, but spatially autocorrelated data).

Module 2: Point Pattern Analysis

- Understand that the primary focus of point pattern analysis is on understanding the drivers of the spatial arrangement of the points and define a point pattern dataset mathematically (i.e., a point pattern is a set $\mathbf{x} = \{x_1, x_2, ...\}$ of points in a two-dimensional space, which has a finite number of points in any bounded region B (i.e., $n(\mathbf{x} \cap B)$ is finite)).

- Understand and describe the ingredients in point pattern analysis (i.e., points, window, marks, covariates).

- Define the intensity of a spatial point pattern under an assumption of homogeneity:

$$\mathbb{E}[n\mathbf{X} \cap B] = \lambda|B| \tag{1}$$

- Define the intensity of a spatial point pattern under an assumption of inhomogeneity:

$$\mathbb{E}[n(\mathbf{X} \cap B)] = \int_B \lambda(u)du \tag{2}$$

- Understand and interpret the estimators of λ and $\lambda(u)$, and know their assumptions.

  1. $\hat{\lambda} = \frac{n(\mathbf{x})}{|B|}$
  2. $\hat{\lambda(j)} = \frac{n(\mathbf{x} \cap B_j)}{|B_j|}$ for $j = 1 \ldots, m$
  3. Kernel density estimation

- Understand the objective of hot spot analysis (you do not need to know the equations for these estimators).

- Understand when/where to apply quadrat tests.

- Understand and interpret the estimators of second moment statistics, and know their assumptions and limitations, as well as why edge corrections are necessary (you do not need to memorise these equations, but should understand how they function).

  1. $M = m\frac{\Sigma_j n_j(n_j-1)}{n(n-1)}$
  2. $\hat{K}(r) = \frac{|W|}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} 1\{d_{ij} \leq r\} e_{ij}(r)$
  3. $\hat{g}(r) = \frac{K(r)\prime}{2\pi r}$

- Understand the assumption underlying the Poisson point process (homogeneity, independence, Poisson distributed random variable).

- Be able to interpret and write down the equation for a point process model (i.e., $\lambda(u) = e^{\beta_0+\beta_1 Z_1(u)+\beta_2 Z_2(u)\ldots+\beta_i Z_i(u)}$).

- Know the model validation steps used in the course.

- Know the core `R` functions used throughout the course (i.e., `intensity()`, `marks()`, `density()`, `quadratcount()`, `quadrat.test()`, `envelope()`, `rhohat()`, `miplot()`, `Kest()`, `Kinhom()`, `pcf()`, `pcfinhom()`, `ppm()`, `effectfun()`, `residuals()` `parres()`, `lurking()`), and be able to interpret them as well as mention when they would be used.

- Be able to read and interpret the key figures used in the course (i.e., intensity maps, hotspot maps, rho vs covariate, Morisita's index, Ripley's K, pair correlation function).

- You do not need to know about the tools/models we covered in passing (i.e., quadrature schemes, Cox processes, Gibbs processes).

Module 3: Spatially Autocorrelated Data

- Understand the impacts of autocorrelation (i.e., bias, over-confidence in predictions, misrepresentative confidence intervals).

- Understand that Moran's I is a measure of spatial correlation (you do not need to memorise the equation).

- Understand how to estimate, read, and interpret semi-variograms (you do not need to memorise the equation).

- Understand how Kriging, co-Kriging, regression-Kriging, and regression with correlated errors function (i.e., use cases, data requirements, limitations, general behaviour of the techniques). Note: you do not need to memorise these equations, but should understand how they function.

- Understand the relationship(s) between autocorrelation and study design.

- You do not need to know the `R` functions from this portion of the course.