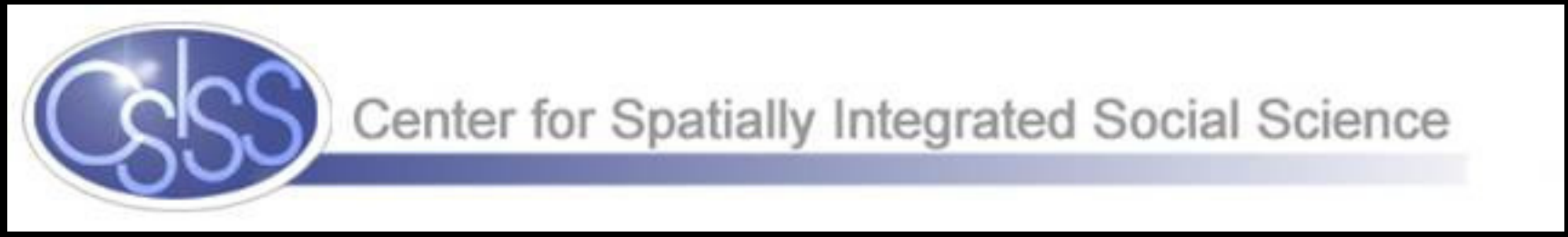


Rate Maps and Smoothing

Luc Anselin
Spatial Analysis Laboratory
Dept. Agricultural and Consumer Economics
University of Illinois, Urbana-Champaign
http://sal.agecon.uiuc.edu

Outline

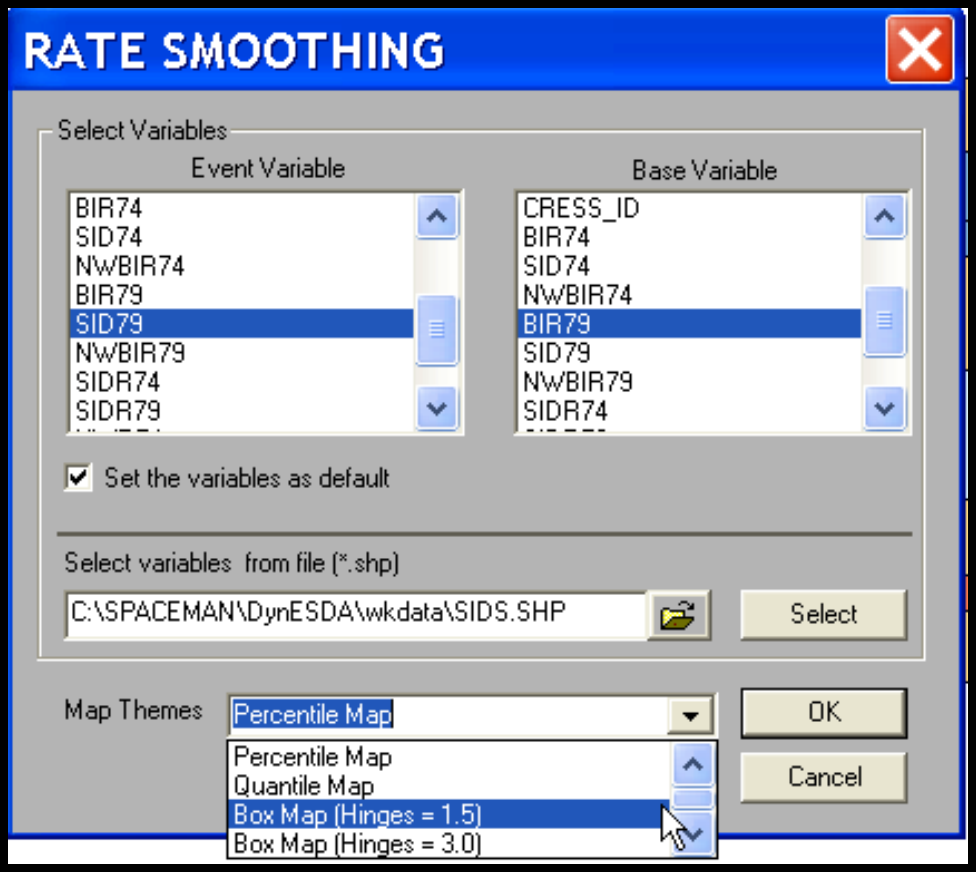
- Mapping Rates
- Risk and Excess Risk
- Empirical Bayes Smoothing
- Spatial Smoothing

Mapping Rates

Events

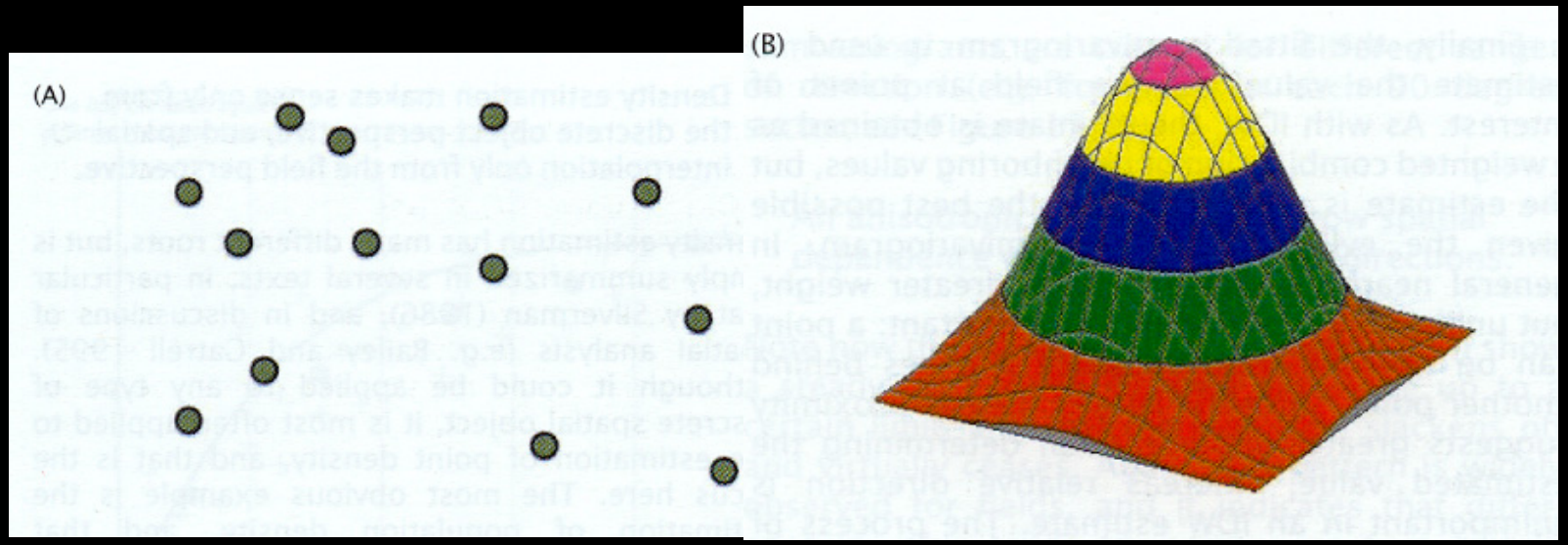
- Locations
  - individual points
    - point pattern analysis
  - areal aggregates
    - counts of events
    - rate = # events / # population at risk
    - raw rate is ML estimate of "risk"
- Probability Surface or Risk Surface
  - probability that event will happen at any given location, in any given area

Event and Base



Kernel Estimation

- Creating a Continuous Surface for Discrete Events
  - typical application = point density surface
    - e.g. location of homicides yields a crime surface
  - weighted average of values around points
    - locate a grid over data
    - center "kernel" at each grid point and compute average of points within the range
    - kernel parameters (bandwidth, slope, etc.)
  - different kernels
    - simple average, inverse distance, gaussian



data points

kernel smoother

Kernel Estimation - Density Surface

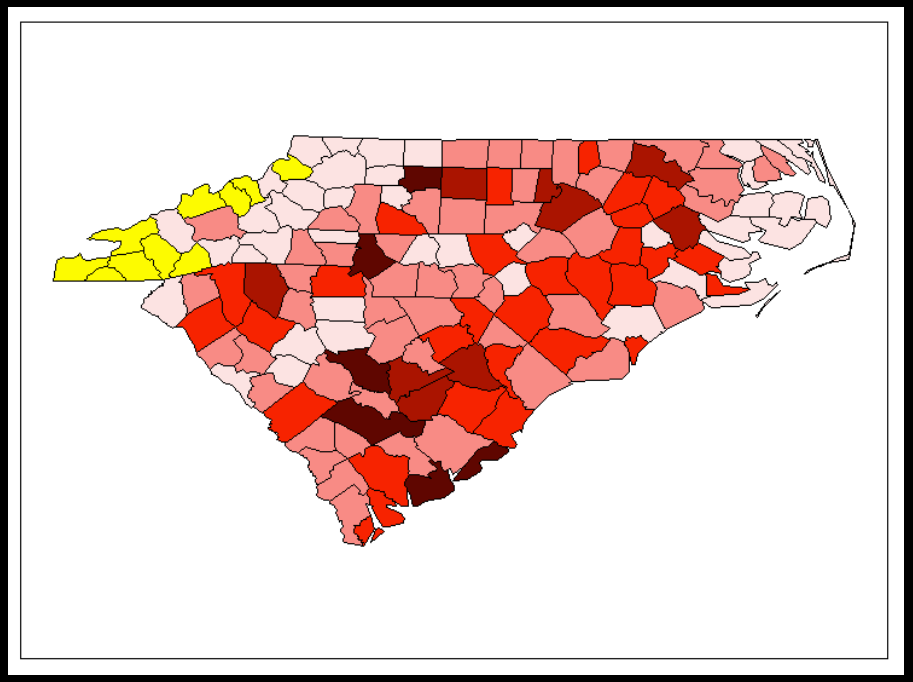
Source: Longley et al (2001)

Risk and Excess Risk

Problems with Rates

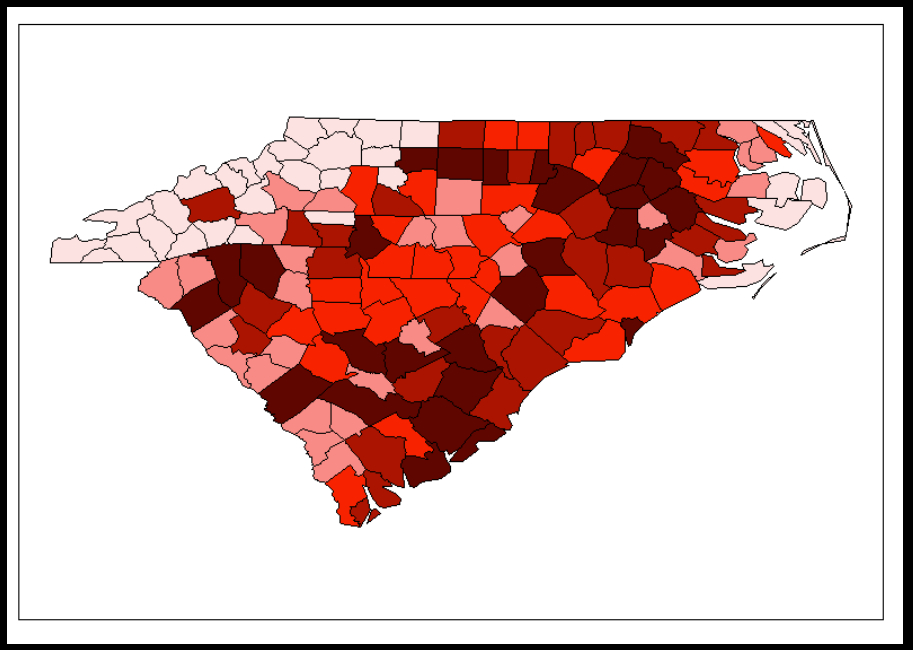
- Risk
  - rate as an estimate of underlying risk
    - expected events = risk x population at risk
- Issues
  - intrinsic heterogeneity
    - variance depends on mean
    - variance depends on base
  - variance instability
    - spurious outliers

Events and Expected Events



<< number of strokes for 65+ African-American males, NC and SC (yellow = zero)

expected # strokes using average risk for state, not average of county risks (risk = total deaths / pop at risk) >>



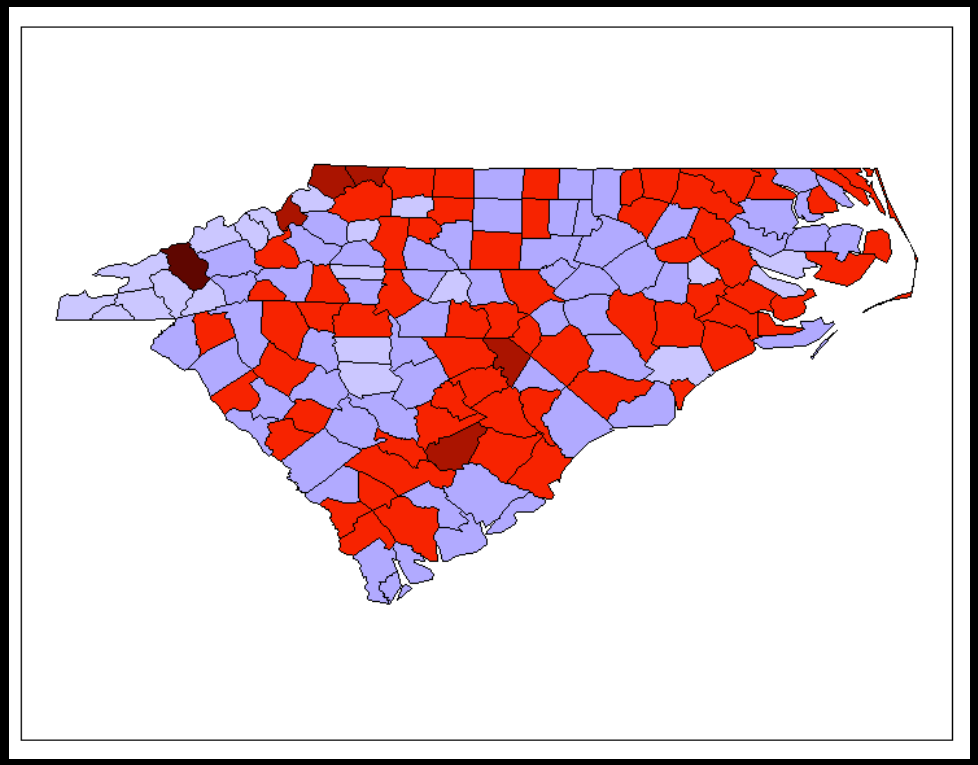
Expected Events

- Number of Events if Average Risk Applied
  - control for age/sex structure of population at risk
    - different rates for different age groups
  - apply estimated risk, a prior
    - estimated risk = average for reference population
  - expected events = Σ_k r_k x p_k for k groups
    - r_k risk for group k
    - p_k population in group k

Relative Risk

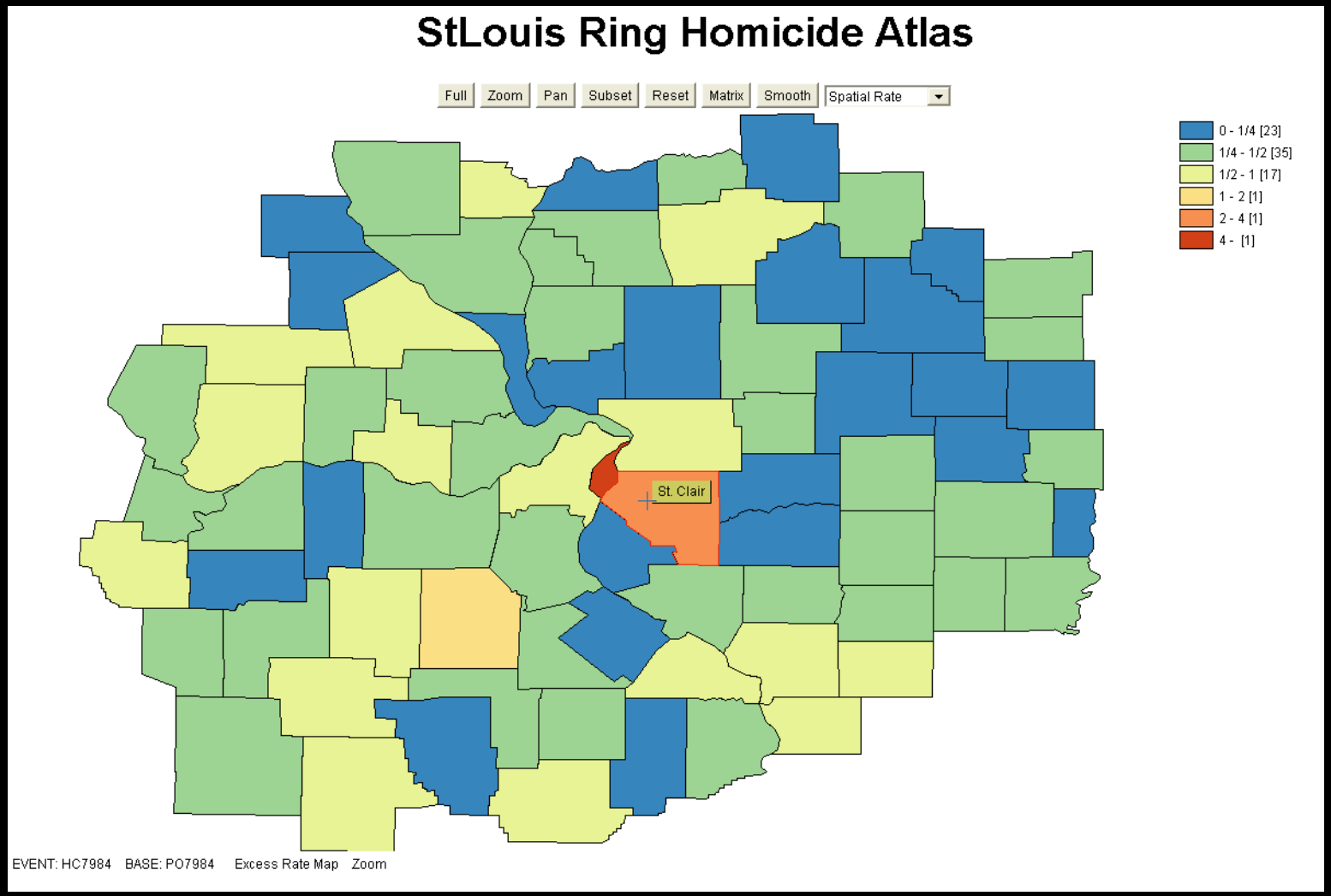
- Compare Observed to Expected
  - observed = number of events
  - expected = number of events if average risk applied to population
- Relative Risk
  - ratio of observed events over expected events
  - > 1: higher (excess) risk
  - < 1: lower risk
  - SMR = standardized mortality ratio
    - use age/gender specific rates

Excess Risk Map - Strokes



ratio of observed stroke deaths over expected, blue <1 red > 1
(darkest > 3)

Excess Risk St. Louis Homicides



Empirical Bayes Smoothing

Variance Instability in Rates

- Rates as Binomial r.v.
- x/P: x count, P population at risk
  - $E[x/P] = \pi$
  - $Var[x/P] = \pi(1-\pi)/P$ with π as unknown
- Unknown mean in variance
  - non-standard assumption
- Variance depends inversely on base P
  - smaller areas have larger variance
  - example: $\pi = 0.1$ $P_1=500$ $P_2=100,000$
    - $SD_1 = 0.013$ $SD_2 = 0.0009$

Shrinkage

- Principle
  - use "prior" information on unknown π_i to smooth (shrink) rate estimate $r_i = x_i/P_i$
  - shrinkage inverse function of variance
    - smaller variance gives higher weight to observed rate
    - larger variance gives higher weight to prior
  - prior θ_i with mean γ_i and variance ϕ_i

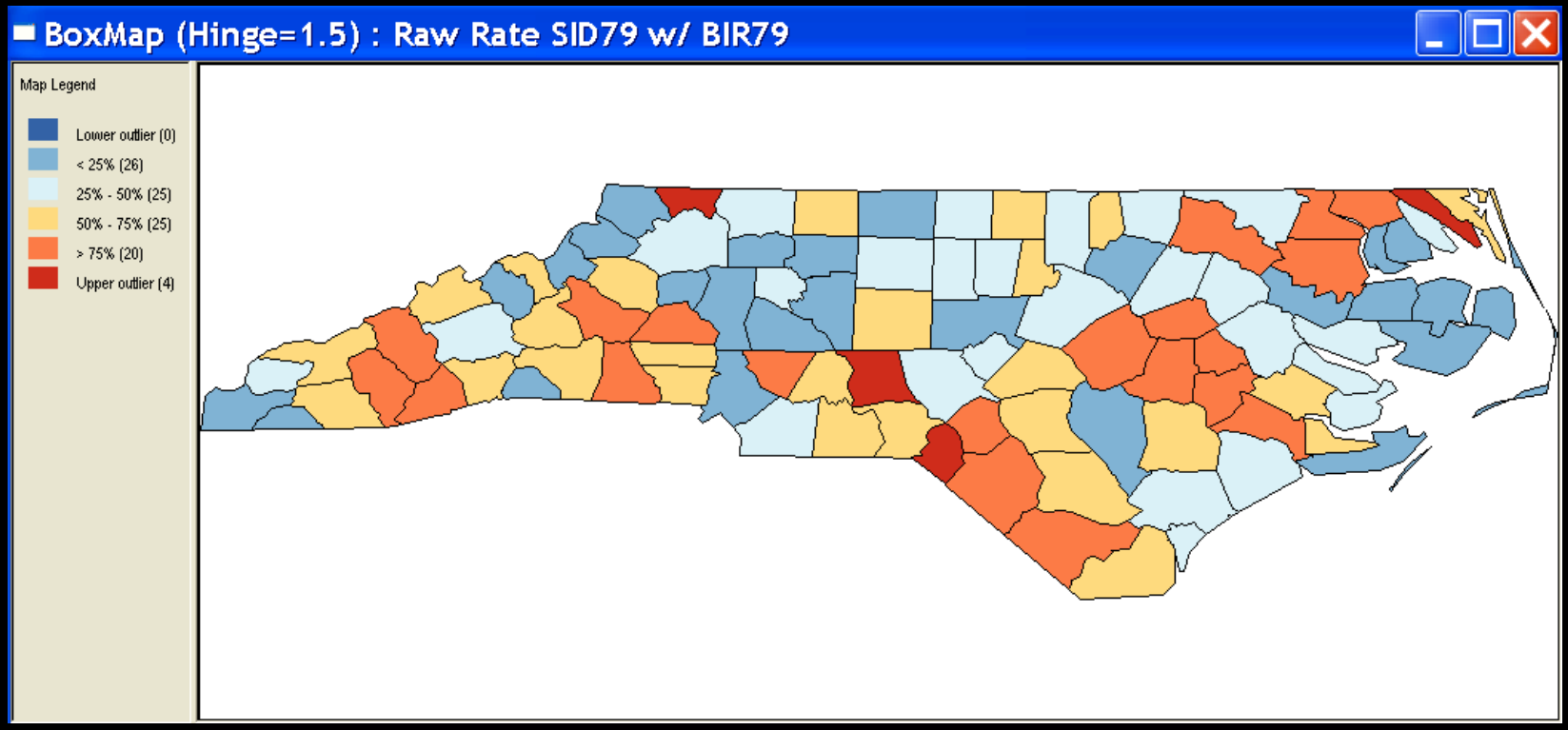
Bayesian Estimate

- Bayesian Principle
  - combine prior distribution with information provided by the data
- Empirical Bayes (EB)
  - prior comes from the data
  - est $\pi_i = w_i r_i + (1 - w_i)\theta_i$
    - is weighted average of r and prior
  - $w_i = \phi_i / [\phi_i + (\theta_i/P_i)]$
    - for large P weight w is 1 (no weight given to prior)

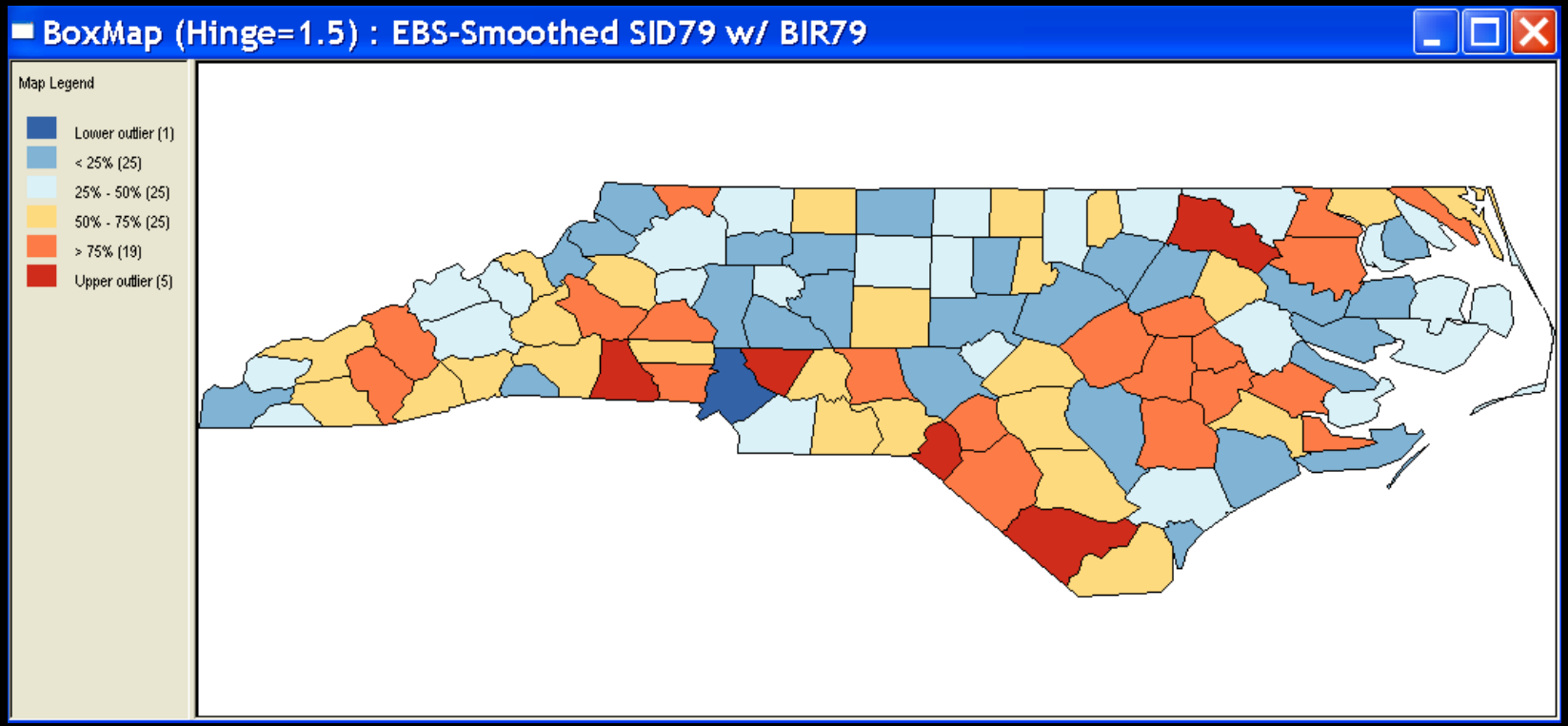
EB Transformation

- Estimate Mean (θ) and Variance (ϕ)
- Method of Moments
  - overall mean θ
    - $\theta = \Sigma_i x_i / \Sigma_i P_i$
    - $P_{av} = \Sigma_i P_i / R$ (R = # areas)
  - overall variance ϕ
    - $\phi = \{ [\Sigma_i P_i(r_i - \theta)^2] / [\Sigma_i P_i] \} - (\theta / P_{av})$

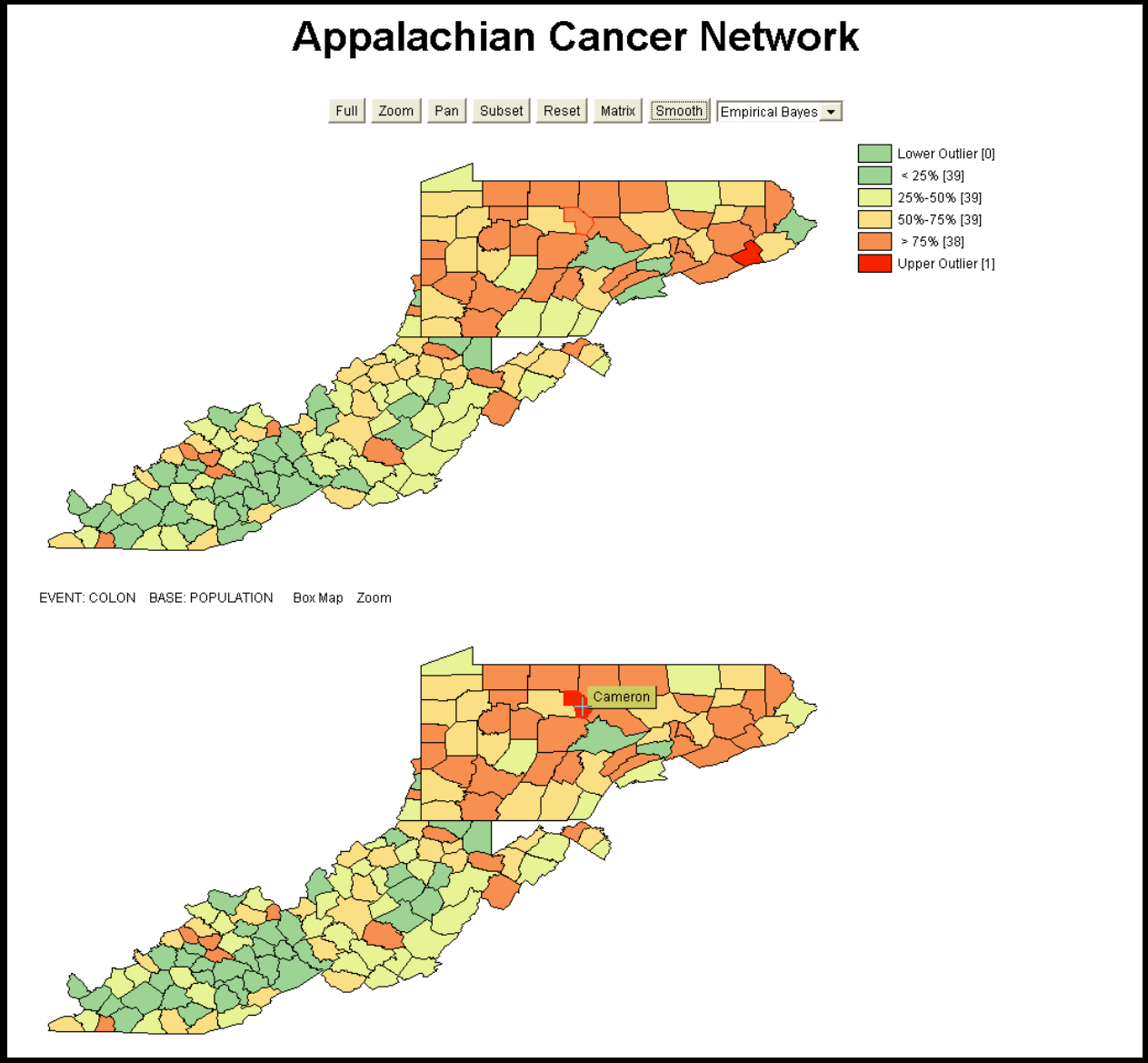
Raw Rate Map



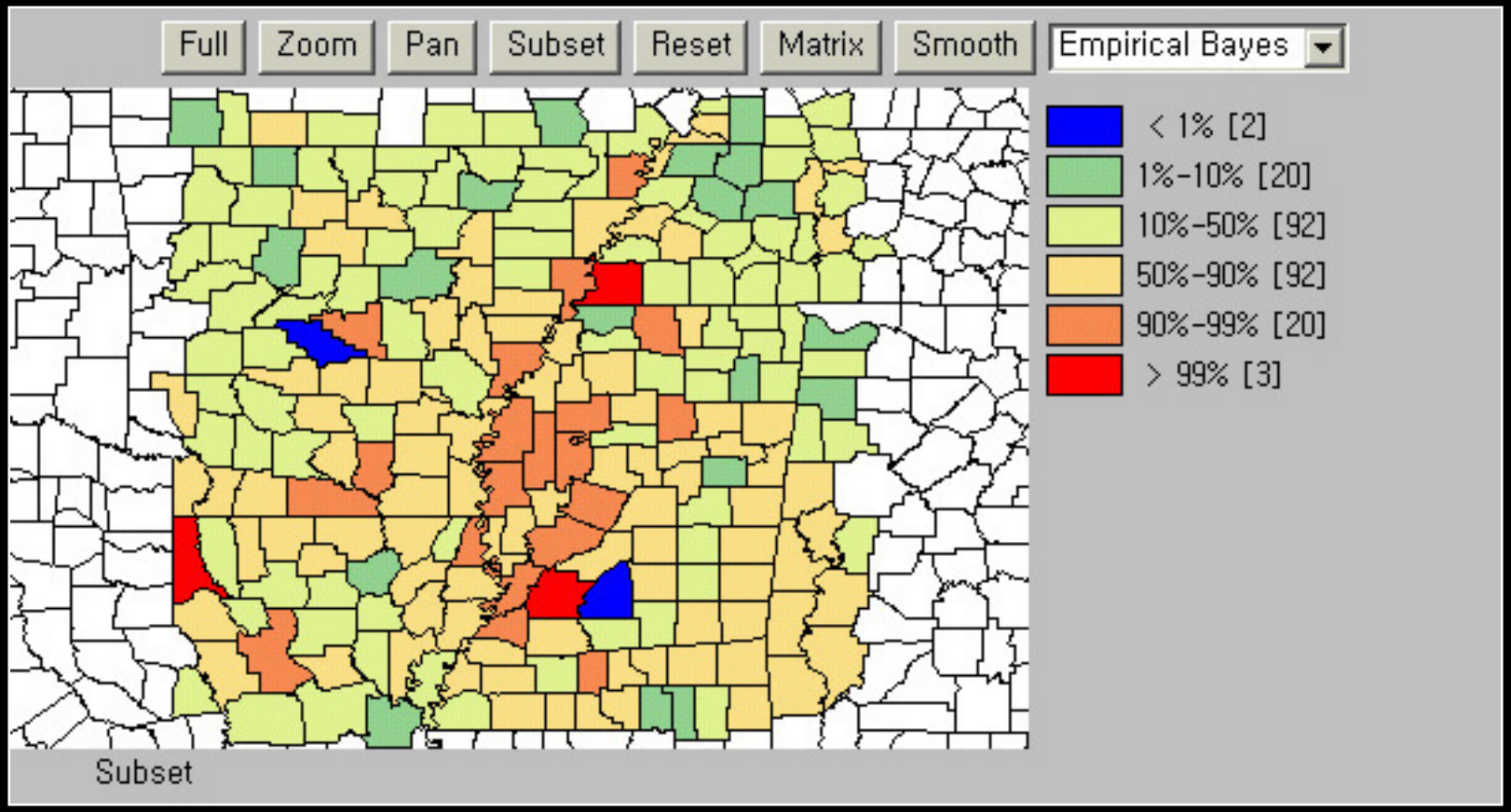
EB Smoothed Map



EB Rates vs. Raw Rates



Regional EB Smoothing

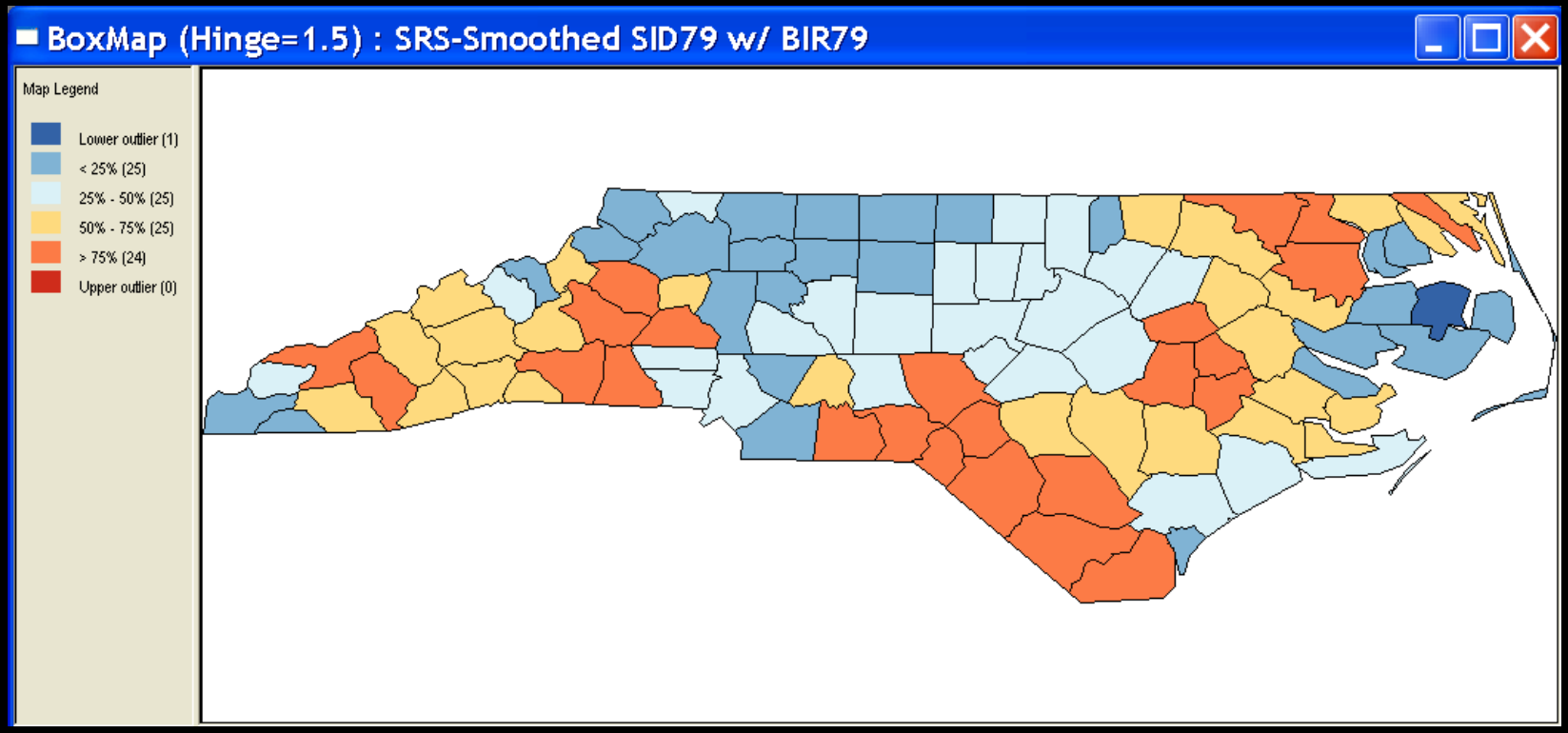


Spatial Smoothing

Spatial Window Smoother

- Replace raw rate by window average
  - define "reference" neighbors S_i
    - use window around area
  - spatial window average
    - $\Sigma_j x_j / \Sigma_j P_j =$ average for window
    - smoothes out the peaks in the data
    - fewer outliers
  - highlights broad spatial trends
    - the highs and the lows

Spatial Rate Smoother



Spatial Rate Smoothing