What is a point pattern?

For a specified, bounded region $D$, a set of locations $s_i, i = 1, 2..., n$
What is a point pattern?

- For a specified, bounded region $D$, a set of locations $s_i, i = 1, 2..., n$
- The locations are ‘random’
What is a point pattern?

- For a specified, bounded region $D$, a set of locations $s_i, i = 1, 2..., n$
- The locations are ‘random’
- Only locations, no variables at locations
What is a point pattern?

- For a specified, bounded region $D$, a set of locations $s_i, i = 1, 2..., n$
- The locations are ‘random’
- Only locations, no variables at locations
- Crude features of patterns, e.g., complete randomness, clustering, regular/systematic
What is a point pattern?

- For a specified, bounded region $D$, a set of locations $s_i, i = 1, 2..., n$
- The locations are ‘random’
- Only locations, no variables at locations
- Crude features of patterns, e.g., complete randomness, clustering, regular/systematic
- marks; marked patterns; comparison of patterns
spatial homogeneity
cluster pattern; systematic pattern
spatial heterogeneity
spatial heterogeneity
Examples

- pattern of trees in a forest, perhaps junipers and pinions
Examples

- pattern of trees in a forest, perhaps junipers and pinions
- pattern of disease cases, perhaps cases and controls
Examples

- pattern of trees in a forest, perhaps junipers and pinions
- pattern of disease cases, perhaps cases and controls
- breast cancer cases; treatment option - mastectomy or radiation
Examples

- pattern of trees in a forest, perhaps junipers and pinions
- pattern of disease cases, perhaps cases and controls
- breast cancer cases; treatment option - mastectomy or radiation
- perhaps over time, single family homes; urban development
Examples

- pattern of trees in a forest, perhaps junipers and pinions
- pattern of disease cases, perhaps cases and controls
- breast cancer cases; treatment option - mastectomy or radiation
- perhaps over time, single family homes; urban development
- again over time, bovine tuberculosis
Counting Measure

For any set \( B \subset D \), let \( N(B) \) count the number of pints in \( B \).

Given for all \( B \)’s, we call a counting measure.

Counting measure is equivalent to a point pattern.

If point pattern is random, then \( N(B) \)’s are random.

Model for the uncountable collection of sets?
Moments

- Moment measures, first order intensity -
  \[ E(N(B)) = \lambda(B) = \int_B \lambda(s) ds, \text{ second order intensity -} \]
  \[ E(N(A)N(B)) = \int_A \int_B \gamma(s, s') ds ds' \]

- Build up from \( \lambda(s) \), from \( \gamma(s, s') \)

- Scale down, i.e.,

  \[
  \lambda(s) = \lim_{|\partial s| \to 0} \frac{E(N(\partial s))}{|\partial s|}
  \]

  \[
  \gamma(s, s') = \lim_{|\partial s| \to 0, |\partial s'| \to 0} \frac{E(N(\partial s)N(\partial s'))}{|\partial s||\partial s'|}
  \]

- Obviously depends upon \( |B| \), area of \( B \)
Reweighted second order: $\gamma(s, s')/\lambda(s)\lambda(s')$
conditional: $\gamma(s, s')/\lambda(s')$, intensity at $s$ given a point at $s'$
stationary model for a point pattern:
$N(B) \sim N(B + h) \Rightarrow \gamma(s, s') = \gamma(s - s')$
isotropic model for a point pattern:
$N(B) \sim N(PB) \Rightarrow \gamma(s, s') = \gamma(||s - s'||)$
Here $P$ is an orthogonal matrix and $PB = \{s^* = Ps : s \in B\}$

cont.
Distributions

- Finally, $N(B) \sim ?$, Poisson Process
- $\lambda(s) = \lambda$ - homogeneous Poisson process, spatial homogeneity, complete spatial randomness (csr)
- $\lambda(s)$ nonconstant, fixed - nonhomogeneous Poisson process
- $\lambda(s)$ random - Cox process
- more to follow
Exploratory data analysis

- All directed at checking/criticizing the CSR assumption
Exploratory data analysis

- All directed at checking/criticizing the csr assumption

- Crude - Partition $D$ in cells of equal area (need not exhaust $D$). Compute $\bar{N}$, the mean of the cell counts. Compute $S^2_N$, the sample variance of the counts. Look at $S^2_N / \bar{N}$. 
Exploratory data analysis

- All directed at checking/criticizing the csr assumption
- Crude - Partition $D$ in cells of equal area (need not exhaust $D$). Compute $\bar{N}$, the mean of the cell counts. Compute $S^2_N$, the sample variance of the counts. Look at $S^2_N/\bar{N}$.
- Can extend to a standard $\chi^2$ test treating the cell counts as i.i.d. Poisson random variables under csr
**$F$ and $G$ functions**

Distance-based methods:
- $G(d)$ is nearest neighbor distance, event to event, i.e.,
  \[ G(d) = Pr(\text{nearest event} \leq d) \]
- $F(d)$ is nearest neighbor distance point to event, i.e.,
  \[ F(d) = Pr(\text{nearest event} \leq d) \]

Under csr $G(d) = F(d) = 1 - \exp(-\lambda \pi d^2)$

$\hat{G}$ is the empirical c.d.f. of the $n$ nearest neighbor distances (nearest neighbor distance for $s_1$, for $s_2$, etc.

$\hat{F}$ is the empirical c.d.f. arising from the $m$ nearest neighbor distances associated with a randomly selected set of $m$ points in $D$.

Edge correction if $d > b_i$ where $b_i$ is distance from $s_i$ to boundary of $D$

compare $\hat{G}$ with $G$, $\hat{F}$ with $F$ - theoretical Q-Q plot. Shorter tails - clustering, longer tails - repulsion.
The $K$ function

- The $K$ function considers the number of points within distance $d$ of an arbitrary point.
- Under isotropy, expected number is the same for any point.
- Under isotropy, can connect $K$ to $\gamma$, i.e., $\gamma(d) = \frac{\lambda^2 K'(d)}{2\pi d}$.
- We work with $K$ since it is easier to interpret, easier to estimate.
- Formally, $K(d) = (\lambda)^{-1} E(\# \text{ of points within } d \text{ of an arbitrary point})$.
Under csr, $K(d) = \lambda \pi d^2 / \lambda = \pi d^2$

\[ \hat{K}(d) = \hat{\lambda}^{-1} \sum_i \sum_j (w_{ij})^{-1} I(||s_i - s_j|| \leq d) \]

\[ \hat{\lambda} = n/|D| \]

$w_{ij}$ is an edge correction, the proportion of the circumference of the circle centered at $s_i$ with radius $||s_i - s_j||$ within $D$

Compare $\hat{K}(d)$ with $K(d) = \pi d^2$

Regularity implies $K(d) < \pi d^2$, clustering implies $K(d) > \pi d^2$
Estimating the intensity

- Crude approach: Imagine a refined grid over $D$. Then $\lambda(\partial s) = \int_{\partial s} \lambda(s) ds \approx \lambda(s)|\partial s|$

- So, for grid cell $A_l$, assume $\lambda$ is constant over $A_l$ and estimate with $N(A_l)/|A_l|$.

- Two dimensional step function, tile function; like a two-dimensional histogram
More sophisticated: A kernel intensity estimate (like a kernel density estimate)

\[ \hat{\lambda}_\tau(s) = \sum_i h(||s - s_i||/\tau)/\tau^2, \ s \in D \]

- \( h \) is a radially symmetric bivariate pdf (usually a bivariate normal),
- \( \tau \) is “bandwidth” (controls smoothness of \( \hat{\lambda} \)).
- Don’t divide by \( n \); we cumulate intensity
Generating point patterns

- Generating a realization from a HPP with intensity \( \lambda \)
  - Sample \( n \sim \text{Po}(\lambda|D|) \). Given \( n \), sample \( n \) locations uniformly over \( D \)

- Generating a realization from a NHPP given \( \lambda(s) \)
  - Compute \( \lambda_{max} = \max_{s \in D} \lambda(s) \). Sample \( n \sim \text{Po}(\lambda_{max}|D|) \). Given \( n \), sample \( n \) locations uniformly over \( D \). “Thin” by retaining \( s_i \) with probability \( \frac{\lambda(s_i)}{\lambda_{max}} \).

- What if \( \lambda(s) = \exp z(s) \) where \( z(s) \) is a process realization? Why would we simulate in this situation? Such a \( \lambda(s) \) is for model fitting!
NHPP likelihood

- Two views
- Given $N(D) = n$,
  
  $$f(s_1, s_2, \ldots s_n | N(D) = n) = \prod_i \frac{\lambda(s_i)}{(\lambda(D))^n}$$

- So, “joint density”,
  
  $$f(s_1, s_2, \ldots s_n, N(D) = n) = \prod_i \frac{\lambda(s_i)}{(\lambda(D))^n} (\lambda(D))^n \exp(-\lambda(D))/n!$$

  “likelihood” becomes

  $$L(\lambda(s), s \in D; s_1, \ldots s_n) = \prod_i \lambda(s_i) \exp(-\lambda(D))$$

- Alternatively, partition $D$ into a fine grid. From the Poisson assumption, the likelihood will be a product over the grid cells, i.e. $\prod_l \exp(-\lambda(A_l))(\lambda(A_l))^{N(A_l)}$.

- Product of the exponential terms is $\exp(-\lambda(D))$, regardless of the grid. As the grid becomes finer, $N(A_l) = 1$ or 0 if $s_i$ in $A_l$ or not.
Modeling $\lambda(s)$

- $\lambda(s) = \sigma \lambda_0(s)$, $\sigma$ unknown
Modeling $\lambda(s)$

- $\lambda(s) = \sigma \lambda_0(s)$, $\sigma$ unknown
- $\lambda(s)$ is a tiled surface over a grid requiring $\lambda_l$ for $A_l$
Modeling $\lambda(s)$

- $\lambda(s) = \sigma \lambda_0(s)$, $\sigma$ unknown
- $\lambda(s)$ is a tiled surface over a grid requiring $\lambda_l$ for $A_l$
- $\lambda(s; \theta)$, a parametric function
Modeling $\lambda(s)$

- $\lambda(s) = \sigma \lambda_0(s)$, $\sigma$ unknown
- $\lambda(s)$ is a tiled surface over a grid requiring $\lambda_l$ for $A_l$
- $\lambda(s; \theta)$, a parametric function
- $\lambda(s)$ a process realization, e.g., $\lambda(s) = \exp(z(s))$ where $z(s)$ is a realization from a spatial Gaussian process
Modeling $\lambda(s)$

- $\lambda(s) = \sigma \lambda_0(s)$, $\sigma$ unknown
- $\lambda(s)$ is a tiled surface over a grid requiring $\lambda_l$ for $A_l$
- $\lambda(s; \theta)$, a parametric function
- $\lambda(s)$ a process realization, e.g., $\lambda(s) = \exp(z(s))$ where $z(s)$ is a realization from a spatial Gaussian process
- From a Bayesian viewpoint, these are all random; these are all Cox processes
Bayesian model fitting

For parametric cases, we write likelihood as
\[ L(\theta; s_1, \ldots, s_n) \]
with a prior on \( \theta \) as usual. But we will need to calculate
\[ \int_D \lambda(s; \theta) ds \]

For the nonparametric case, we write the likelihood as
\[ L(\lambda_D; s_1, \ldots, s_n) \]
where \( \lambda_D = \{\lambda(s) : s \in D\} \). We have a GP prior on \( \log \lambda_D \)
We have an uncountable dimensional likelihood and we will need to calculate
\[ \int_D \exp(z(s)) ds \]

Can’t do this exactly. Numerical or Monte Carlo integration will reduce to a finite set of \( z(s) \)’s, hence to a multivariate normal distribution.
The “poor man’s” version. Overlay a grid on \( D \), work with the Poisson counts associated with the grid. Treat \( \lambda_i \)’s as realizations from a CAR model. Reduce to a disease mapping problem. Choice of, sensitivity to the grid?

In fact, can introduce covariates into \( \lambda(s) \), i.e.,
\[
\log \lambda(s) = X^T(s)\beta + \phi(s)
\]

With marks, an intensity, hence a process for each mark. Dependent or independent processes?
Some special Cox processes

- Poisson cluster process (Neyman-Scott process)
  - Generate parent events from a NHPP with $\lambda(s)$
  - Each parent produces a random number of offspring, e.g., $N(s_i)$ are i.i.d. $\text{Po} \circ (\delta)$

- Positions of offspring are determined by i.i.d. realizations of a bivariate density

- Only offspring are retained as point pattern

- An inhibition process: Generate a realization of a csr. “Thin” by deletion of all pairs with distance less than $\delta$ apart where $\delta$ is the minimum permissible distance
Space-time point patterns

- What is time resolution?
- Discrete time, e.g., pattern of annual plants, annual pattern of disease cases
- Time is continuous, e.g., when a house is built, when a bovine TB case occurred
- In either case, we need to think about \( \lambda(s, t) \)
- If time is discrete, then \( \lambda_t(s) \) can be modeled dynamically, parametrically or nonparametrically
- If time is continuous, then again parametric \( \lambda(s, t; \theta) \) or nonparametric logGaussian spatio-temporal Cox process model, i.e., \( z(t, s) = \log \lambda(s, t) \) is a realization from a space-time Gaussian process.
- Note that to see a pattern, we must integrate \( \lambda(s, t) \) over a region in space and an interval in time.
References

- Cressie book is still useful
- Diggle book - more modern but not very complete
- Diggle review papers - available from his website
- Møller and Waagepetersen - technical, likelihood based
- Gotway and Waller - easy reading, less modeling, more EDA style