Stat 534 Project: Extrapolation from Poisson Process Intensity Surface Models

Kenny Flagg

Introduction

- Goals
  - Estimate inhomogeneous intensity surface from events in a subregion
  - Infer intensity across entire region
- Applications
  - Mapping where endangered species are located
  - Mapping geomegnetic anomalies prior to an unexploded ordnace (UXO) remediation

Maximum Likelihood Intensity Surface Fitting

- General point processes
  - The theory is not too complicated but the computation is very difficult
- Poisson processes
  - Doable with numerical methods
  - Log-likelihood of Poisson with intensity $\lambda(\mathbf{s})$ on region D (note typos in Diggle (2013))

$$\begin{aligned}\ell(\lambda) &= \{-\mu + n\log(\mu) - \log(n!)\} + \sum_{i=1}^{n}\{\log(\lambda(\mathbf{s}_i)) - \log(\mu)\}\\ &= \sum_{i=1}^{n}\log(\lambda(\mathbf{s}_i)) - \int_D \lambda(\mathbf{s})d\mathbf{s} - \log(n!).\end{aligned}$$

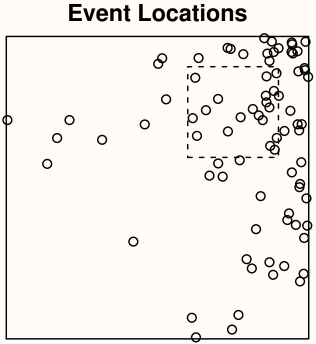
where $\mu = \int_D \lambda(\mathbf{s})d\mathbf{s}$

Poisson Process Log-Linear Model

- Assuming events are independent (conditional on the intensity function),
$$\log(\lambda(\mathbf{s})) = \mathbf{x}(\mathbf{s})^T \beta$$
where $\mathbf{x}(\mathbf{s})^T$ is a row of predictors at location $\mathbf{s}$
- Predictors can include covariates, but they must be known across the whole region
- Berman and Turner (1992) use dummy points and quadrature to set up an approximation as a weighted Poisson regression
- Their method is implemented in `spatstat`'s `ppm`, with `glm` from base R or `gam` from `mgcv` as the back-end

Simple Example

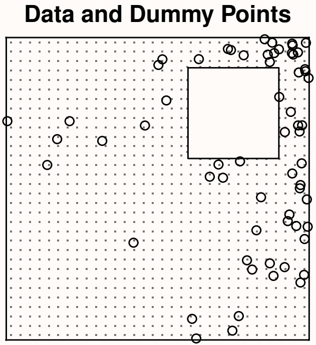
- True model
  - Poisson process on the unit square
  - $\log(\lambda(x, y)) = 5x + 2y$
- But we don't observe $0.6 < x < 0.9, 0.6 < y < 0.9$



Estimated Model

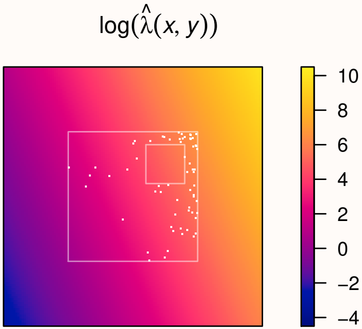
- Fit the model $\log(\lambda(x, y)) = \beta_0 + \beta_1 x + \beta_2 y$

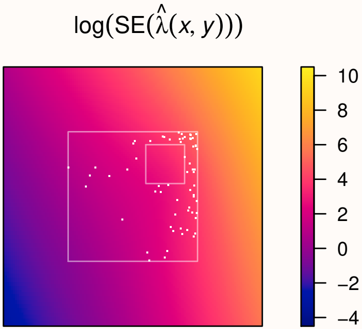
	Estimate	S.E.
$\widehat{\beta}_0$	0.20	0.56
$\widehat{\beta}_1$	4.54	0.59
$\widehat{\beta}_2$	2.00	0.44



Extrapolate the Surface

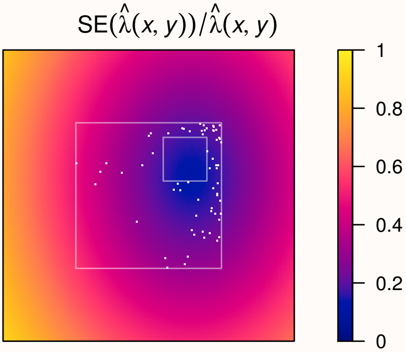
- Use the `predict` method
- Specify a new window $-0.5 < x < 1.5, -0.5 < y < 1.5$





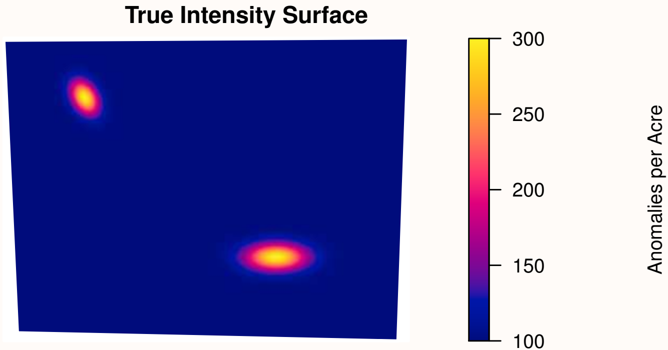
Where is the Uncertainty?

- Relative standard error is lowest where the highest intensity was observed



Simple UXO Site

- 952.38 acre region (roughly 7,625 ft by 5,709 ft)
- High density of geomagnetic anomalies around targets
- Low density of background anomalies



References I

Berman, Mark and Rolf Turner (1992). "Approximating point process likelihoods with GLIM". In: *Applied Statistics*, pp. 31–38.

Diggle, Peter J. (2013). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. 3rd ed. CRC Press.